Algebra & Number Theory

Volume 9
2015
No. 3

msp

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2015 is US $255/year for the electronic version, and $440/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 9:3 (2015)
dx.doi.org/10.2140/ant.2015.9.511

Hurwitz monodromy and full number fields

David P. Roberts and Akshay Venkatesh

We give conditions for the monodromy group of a Hurwitz space over the configuration space of branch points to be the full alternating or symmetric group on the degree. Specializing the resulting coverings suggests the existence of many number fields with surprisingly little ramification — for example, the existence of infinitely many A_m or S_m number fields unramified away from $\{2, 3, 5\}$.

1. Introduction 511
2. Hurwitz covers 513
3. Braid groups 515
4. Lifting invariants 517
5. The full-monodromy theorem 523
6. Proof of I $\Longrightarrow$ II 529
7. Proof of II $\Longrightarrow$ I 536
8. Full number fields 541
Acknowledgements 544
References 544

1. Introduction

1A. ***Overview.*** Hurwitz spaces are defined as moduli spaces of branched covers of the complex projective line P^1 satisfying certain conditions. A given Hurwitz space is canonically presented as a finite-degree covering of the configuration space of possible branching divisors. An important problem is to characterize those Hurwitz spaces for which the monodromy group of this covering is the full alternating or symmetric group on the fiber. Our main result, Theorem 5.1, gives such a characterization in an asymptotic setting when the covers of P^1 being parametrized have suitably many branch points.

Our interest in fullness of Hurwitz monodromy arises from applications to constructing number fields with large Galois group and little ramification, and in particular from an open problem posed in [Malle and Roberts 2005]: Say that a degree-m number field K is *full* if its associated Galois group is either A_m or S_m.

MSC2010: primary 14D05; secondary 20F36, 11R21.
Keywords: number fields, Hurwitz spaces.

For a given fixed set of primes $\mathcal{P}$, are there infinitely many full fields K for which the discriminant of K is divisible only by primes in $\mathcal{P}$? Our Theorem 5.1, together with experimental data to be presented in a sequel paper [Roberts ≥ 2015], strongly suggests that the answer to the question is *yes*, whenever $\mathcal{P}$ contains the set of primes dividing the order of a finite nonabelian simple group. This expectation is particularly interesting because the mass heuristic of [Bhargava 2007] predicts *no* for all $\mathcal{P}$.

Sections 2, 3 and 4 provide short summaries of large theories and serve to establish our setting. Section 5 states our main theorem, which we call the full-monodromy theorem. It has the form that two statements, I and II, about data (G, C) defining a multiindexed collection of Hurwitz covers are equivalent. Statement I is an explicit condition on (G, C) and Statement II is an asymptotic statement about the monodromy of the covers in the collection. Sections 6 and 7 prove the theorem by establishing I $\Longrightarrow$ II and II $\Longrightarrow$ I, respectively. Section 8 concludes the paper with a discussion of the application to the construction of full number fields.

1B. *The full-monodromy theorem.* This subsection provides an introductory description of the full-monodromy theorem. Define a *Hurwitz parameter* to be a triple $h = (G, C, \nu)$, where G is a finite group, $C = (C_1, \ldots, C_r)$ is a list of conjugacy classes whose union generates G, and $\nu = (\nu_1, \ldots, \nu_r)$ is a list of positive integers, with ν *allowed* in the sense that $\prod [C_i]^{\nu_i} = 1$ in the abelianization G^{ab}. A Hurwitz parameter determines an unramified covering of complex algebraic varieties

$$\pi_h : \mathsf{Hur}_h \to \mathsf{Conf}_\nu. \tag{1-1}$$

Here, the cover Hur_h is a Hurwitz variety parameterizing certain covers of the complex projective line P^1, where the coverings are "of type h". The base Conf_ν is the variety whose points are tuples $(D_1, \ldots, D_r)$ of disjoint divisors D_i of P^1, with $\deg(D_i) = \nu_i$. The map π_h sends a cover to its branch locus.

In complete analogy with the use of the term for number fields, we say that a cover of connected complex algebraic varieties $\mathsf{X} \to \mathsf{Y}$ is *full* if its monodromy group is the entire alternating or symmetric group on the degree. There are two relatively simple obstructions to (1-1) being full. One is associated to G having a nontrivial outer automorphism group, and we deal with it by replacing Hur_h by a quotient variety Hur_h^* also covering Conf_ν. The other is associated to G having a nontrivial Schur multiplier, and we deal with it by a decomposition $\mathsf{Hur}_h^* = \coprod_\ell \mathsf{Hur}_{h,\ell}^*$. Here ℓ runs over the Schur multiplier modulo a certain equivalence relation, and each $\mathsf{Hur}_{h,\ell}^*$ is a union of connected components of Hur_h^*.

The more important direction of the full-monodromy theorem is I $\Longrightarrow$ II. When G is nonabelian and simple, this direction is as follows:

Fix a nonabelian simple group G and a list $C = (C_1, \ldots, C_r)$ of conjugacy classes whose union generates G. Consider varying allowed ν and thus varying Hurwitz parameters $h = (G, C, \nu)$. Then as soon as $\min_i \nu_i$ *is sufficiently large, the covers* $\mathsf{Hur}^*_{h,\ell} \to \mathsf{Conf}_\nu$ *are full and pairwise nonisomorphic.*

The complete implication I $\Longrightarrow$ II is similar, but G is allowed to be "pseudo-simple", and therefore groups such as S_d are included. There are considerable complications arising from nontrivial abelianizations G^{ab}, even in the case $|G^{\mathrm{ab}}| = 2$. The extra generality is required for obtaining the natural converse II $\Longrightarrow$ I.

Our proof of I $\Longrightarrow$ II in general starts from the Conway–Parker theorem about connectivity of Hurwitz covers [Conway and Parker 1988; Ellenberg et al. 2013; Fried and Völklein 1991; Malle and Matzat 1999]. We deal with complications from nontrivial G^{ab} in the framework of comparing two Hochschild–Serre five-term exact sequences. We upgrade connectivity to fullness by using a Goursat lemma adapted to our current situation and the explicit classification of finite 2-transitive groups. Our general approach has much in common with the proof of Theorem 7.4 in [Dunfield and Thurston 2006], which is in a different context.

While there is a substantial literature on Hurwitz covers, our topic of asymptotic fullness has not been systematically pursued before. In related directions there are the papers [Eisenbud et al. 1991; Kluitmann 1988; Magaard et al. 2003]. We will indicate relations with some of this literature at various points in the present paper.

2. Hurwitz covers

In this section we summarize the theory of Hurwitz covers, taking the purely algebraic point of view necessary for the application to number field construction. We consider Hurwitz parameters $h = (G, C, \nu)$, with G assumed centerless to avoid technical complications. The central focus is an associated cover $\pi_h : \mathrm{HUR}_h \to \mathrm{CONF}_\nu$ and related objects. A more detailed summary can be found in [Romagny and Wewers 2006], and a comprehensive reference in [Bertin and Romagny 2011]. Note that throughout this paper we use a sans serif font for complex analytic spaces, as in $\mathbb{P}^1(\mathbb{C}) = \mathsf{P}^1$ or $\mathrm{CONF}_\nu(\mathbb{C}) = \mathsf{Conf}_\nu$.

2A. ***Configuration spaces*** $\mathbf{CONF}_\nu$**.** Let $\nu = (\nu_1, \ldots, \nu_r)$ be a vector of positive integers; we write $|\nu| = \sum \nu_i$. For k a field, let $\mathrm{CONF}_\nu(k)$ be the set of tuples $(D_1, \ldots, D_r)$ of disjoint k-rational divisors on $\mathbb{P}^1_k$ with D_i consisting of ν_i distinct geometric points.

Explicitly, we may regard

$$\mathrm{CONF}_\nu \subseteq \mathbb{P}^{\nu_1} \times \cdots \times \mathbb{P}^{\nu_r},$$

where we regard $\mathbb{P}^{\nu_i}$ as the projectivized space of binary homogeneous forms $q(x, y)$ of degree ν_i, and $\mathrm{C{\small ONF}}_\nu$ is then the open subvariety defined by nonvanishing of the discriminant $\mathrm{disc}(q_1 \cdots q_r)$. The divisor D_i associated to an r-tuple $(q_1, \dots, q_r)$ of such forms is simply the zero locus of q_i.

2B. ***Standard Hurwitz varieties*** $\mathbf{H{\small UR}}_{\boldsymbol{h}}$**.** Let k be an algebraically closed field of characteristic zero. Consider pairs (Σ, f) consisting of a proper smooth connected curve Σ over k together with a Galois covering $f : \Sigma \to \mathbb{P}^1$.

Such a pair has the following associated objects:

- An automorphism group $\mathrm{Aut}(\Sigma/\mathbb{P}^1)$ of size equal to the degree of f.
- A branch locus $Z \subset \mathbb{P}^1(k)$.
- For every $t \in Z$, a local monodromy element $g_t \in \mathrm{Aut}(\Sigma/\mathbb{P}^1)$ defined up to conjugacy. (To define this requires a compatible choice of roots of unity, i.e., an element of $\varprojlim_n \mu_n(k)$; we assume such a choice has been made.)

Consider triples (Σ, f, ι) with $\iota : G \to \mathrm{Aut}(\Sigma/\mathbb{P}^1)$ a given isomorphism. We say that such a triple has type h if $\sum \nu_i = |Z|$ and for each i there are exactly ν_i elements $t \in Z$ such that $g_t \in C_i$. The branch locus Z then defines an element of $\mathrm{C{\small ONF}}_\nu(k)$ in a natural way.

The theory of Hurwitz varieties implies that there exists a $\overline{\mathbb{Q}}$-variety $\mathrm{H{\small UR}}_h$, equipped with an étale map

$$\pi_h : \mathrm{H{\small UR}}_h \to \mathrm{C{\small ONF}}_\nu, \tag{2-1}$$

with the following property holding for all k: For any $u \in \mathrm{C{\small ONF}}_\nu(k)$, the fiber $\pi_h^{-1}(u)$ is $\mathrm{Aut}(k/\mathbb{Q}(\mu_\infty))$-equivariantly in bijection with the set of isomorphism classes of covers of $\mathbb{P}^1$ of type h, with branch locus equal to u.

2C. ***Quotiented Hurwitz varieties*** $\mathbf{H{\small UR}}^*_{\boldsymbol{h}}$**.** If (Σ, f, ι) is as above, we can modify ι by an element $\alpha \in \mathrm{Aut}(G)$, to obtain a new triple $(\Sigma, f, \iota \circ \alpha^{-1})$. If α is inner, the resulting triple is actually isomorphic to (Σ, f, ι). As a results we obtain actions by groups of outer automorphisms.

Let $\mathrm{Aut}(G, C)$ be the subgroup of $\mathrm{Aut}(G)$ consisting of those elements which fix every C_i. Then $\mathrm{Out}(G, C) = \mathrm{Aut}(G, C)/G$ acts naturally on $\mathrm{H{\small UR}}_h$, giving a quotient

$$\mathrm{H{\small UR}}^*_h = \mathrm{H{\small UR}}_h / \mathrm{Out}(G, C),$$

still lying over $\mathrm{C{\small ONF}}_\nu$. This quotient parameterizes pairs (Σ, f) equipped with an element $(D_1, \dots, D_r)$ of $\mathrm{C{\small ONF}}_\nu(k)$ so that the branch locus is precisely $\bigsqcup D_i$, and there exists an isomorphism $\iota : G \to \mathrm{Aut}(\Sigma/\mathbb{P}^1)$ so that the monodromy around each point of D_i is of type $\iota(C_i)$. Our main theorem focuses on $\mathrm{H{\small UR}}^*_h$ rather than $\mathrm{H{\small UR}}_h$.

2D. ***Descent to*** $\mathbb{Q}$. The discussion that follows is not used in the body of the paper, but it is relevant to the application to full number fields, sketched in Section 8.

The abelianized absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})^{\mathrm{ab}} = \hat{\mathbb{Z}}^{\times}$ acts on the set of conjugacy classes in any finite group by raising representing elements to powers. In particular, one can talk about rational classes, i.e., conjugacy classes fixed by this action. We say that h is *strongly rational* if all C_i are rational. In this case, (2-1) and its starred version $\pi_h^*: \mathrm{H{\small UR}}_h^* \to \mathrm{C{\small ONF}}_\nu$ canonically descend to covers over $\mathbb{Q}$. This statement can be deduced from the corresponding statement for the "large" Hurwitz space, parameterizing coverings without any restrictions on branch monodromy; for that statement see [Fried and Völklein 1991, Theorem 1] and [Romagny and Wewers 2006, Theorems 2.1 and 4.11]. The rationality of the C_i enters because of the dependence on choice of element of $\varprojlim \mu_n$, as above.

More generally, we say that h is *rational* if conjugate classes appear with equal associated multiplicities. In the main case when all the classes are different, this just means $\nu_i = \nu_j$ whenever C_i and C_j lie in the same Galois orbit. Rationality is a substantially weaker condition than strong rationality. For example, any finite group G has rational h, but only when G^{ab} is trivial or of exponent 2 can G have strongly rational h.

For rational h, there is again canonical descent to $\mathbb{Q}$, although now the maps take the form $\mathrm{H{\small UR}}_h \to \mathrm{H{\small UR}}_h^* \to \mathrm{C{\small ONF}}_\nu^\rho$, with ρ indicating a suitable Galois twisting. The subtlety of twisting is not seen in the rest of this paper. Our purpose in briefly discussing twisting here is to make clear that many Hurwitz covers are useful for the construction of full number fields.

3. Braid groups

In this section we switch to a group-theoretic point of view, describing the monodromy of Hurwitz covers $\pi_h : \mathsf{Hur}_h \to \mathsf{Conf}_\nu$ and $\pi_h^* : \mathsf{Hur}_h^* \to \mathsf{Conf}_\nu$ in terms of braid groups and their actions on explicit sets. General references for braid groups and their monodromy actions include [Malle and Matzat 1999, Chapter 3] and [Eisenbud et al. 1991, §2].

Our main theorem concerns these monodromy representations only, i.e., it is a theorem in pure group theory. The map of $\mathbb{Q}$-varieties $\mathrm{H{\small UR}}_h \to \mathrm{C{\small ONF}}_\nu$ underlying the map of complex analytic spaces $\mathsf{Hur}_h \to \mathsf{Conf}_\nu$ will return in Section 8.

3A. ***Braid groups*** $\mathbf{Br}_\nu$. The Artin braid group on n strands is defined by the generators and relations

$$\mathrm{Br}_n = \left\langle \sigma_1, \dots, \sigma_{n-1} \; : \; \begin{array}{rl} \sigma_i\sigma_j = \sigma_j\sigma_i & \text{if } |i-j| > 1 \\ \sigma_i\sigma_j\sigma_i = \sigma_j\sigma_i\sigma_j & \text{if } |i-j| = 1 \end{array} \right\rangle.$$

The rule $\sigma_i \mapsto (i, i+1)$ extends to a surjection $\mathrm{Br}_n \twoheadrightarrow S_n$. For every subgroup of S_n, one gets a subgroup of Br_n by pullback. In particular, from the last component $\nu = (\nu_1, \dots, \nu_r)$ of a Hurwitz parameter one gets a subgroup $S_\nu := S_{\nu_1} \times \cdots \times S_{\nu_r}$. We denote its pullback by Br_ν. The extreme Br_n above and the other extreme Br_{1^n} play particularly prominent roles in the literature, the latter often being called the colored or pure braid group.

3B. *Fundamental groups.* Let $\star = (1, \dots, n) \in \mathsf{Conf}_{1^n}$. We will use it as a basepoint. We use the same notation $\star$ for its image in Conf_ν for any ν. There is a standard surjection $\mathrm{Br}_n \twoheadrightarrow \pi_1(\mathsf{Conf}_n, \star)$, with kernel the smallest normal subgroup containing $\sigma_1 \cdots \sigma_{n-2}\sigma_{n-1}^2\sigma_{n-2} \cdots \sigma_1$ [Malle and Matzat 1999, Theorem III.1.4]. This map identifies σ_i with a small loop in Conf_n that swaps the points i and $i+1$. Because of this very tight connection, the group $\pi_1(\mathsf{Conf}_n, \star)$ is often called the spherical braid group or the Hurwitz braid group.

Similarly, we have surjections

$$\mathrm{Br}_\nu \twoheadrightarrow \pi_1(\mathsf{Conf}_\nu, \star). \tag{3-1}$$

Let $\mathcal{F}_h$ and $\mathcal{F}_h^*$ be the fibers of Hur_h and Hur_h^* over $\star$. To completely translate into group theory, we need group-theoretical descriptions of these fibers as Br_ν-sets. The remainder of this section accomplishes this task.

3C. *Catch-all actions.* We use the standard notational convention $g^h = h^{-1}gh$. If G is any group then Br_n acts on G^n by means of a braiding rule, whereby σ_i substitutes $g_i \to g_{i+1}$ and $g_{i+1} \to g_i^{g_{i+1}}$:

$$(\dots, g_{i-1}, g_i, g_{i+1}, g_{i+2}, \dots)^{\sigma_i} = (\dots, g_{i-1}, g_{i+1}, g_i^{g_{i+1}}, g_{i+2}, \dots). \tag{3-2}$$

Also any $\alpha \in \mathrm{Aut}(G)$ acts on G^n diagonally by

$$(g_1, \dots, g_n)^\alpha = (g_1^\alpha, \dots, g_n^\alpha). \tag{3-3}$$

The braiding action and the diagonal action commute, so one has an action of the product group $\mathrm{Br}_n \times \mathrm{Aut}(G)$ on G^n.

3D. *The* $\mathbf{Br}_\nu$*-sets* $\mathcal{F}_h$ *and* $\mathcal{F}_h^*$. Next we replace G^n by a smaller set appropriate to a given Hurwitz parameter h. This smaller set is

$$\mathcal{G}_h = \big\{(g_1, \dots, g_n) \in G^n : g_1 \cdots g_n = 1, \langle g_1, \dots, g_n \rangle = G, \\ \text{first } \nu_1 \text{ of the } g_i \text{ lie in } C_1, \text{ next } \nu_2 \text{ lie in } C_2, \text{ etc.}\big\}. \tag{3-4}$$

The subset $\mathcal{G}_h$ is not preserved by all of $\mathrm{Br}_n \times \mathrm{Aut}(G)$, but it is preserved by $\mathrm{Br}_\nu \times \mathrm{Aut}(G, C)$. The fibers then have the following group-theoretic description:

$$\mathcal{F}_h = \mathcal{G}_h / \operatorname{Inn}(G) \simeq (\text{fiber of } \mathsf{Hur}_h \to \mathsf{Conf}_\nu \text{ above } \star), \tag{3-5}$$

$$\mathcal{F}_h^* = \mathcal{G}_h / \operatorname{Aut}(G, C) \simeq (\text{fiber of } \mathsf{Hur}_h^* \to \mathsf{Conf}_\nu \text{ above } \star). \tag{3-6}$$

Here in both cases the isomorphisms $\simeq$ are isomorphisms of Br_ν-sets. A clear exposition of the relationship of Hurwitz spaces to braiding is given in [Eisenbud et al. 1991, §1], and the isomorphisms (3-5) and (3-6) are a consequence of this relationship; see also [Fried and Völklein 1991, §1]. Note that $\mathcal{F}_h^* = \mathcal{F}_h / \operatorname{Out}(G, C)$.

3E. ***The asymptotic mass formula.*** Character theory gives mass formulas [Serre 2008, Theorem 7.2.1]. These formulas, applied both to G and to subgroups intersecting all the C_i, can be used to exactly determine the degrees $\mathcal{F}_h$ and $\mathcal{F}_h^*$. We need only the asymptotic versions of the mass formulas for G, which are very simple:

$$|\mathcal{F}_h| \sim \frac{\prod_{i=1}^r |C_i|^{\nu_i}}{|G'||\operatorname{Inn}(G)|}, \qquad |\mathcal{F}_h^*| \sim \frac{\prod_{i=1}^r |C_i|^{\nu_i}}{|G'||\operatorname{Aut}(G, C)|}. \tag{3-7}$$

Here the meaning in each case is standard: the left side over the right side tends to 1 for any sequence of allowed ν with $\min_i \nu_i$ tending to ∞. The structure of the products on the right directly reflects the descriptions of the sets in Section 3D.

4. Lifting invariants

In this section we summarize the theory of lifting invariants, which plays a key role in the study of connected components of Hurwitz spaces. Group homology appears prominently, and as a standing convention we abbreviate $H_i(\Gamma, \mathbb{Z})$ by $H_i(\Gamma)$.

In brief summary, the theory being reviewed goes as follows. Let $h = (G, C, \nu)$ be a Hurwitz parameter. The group G determines its Schur multiplier $H_2(G)$. In turn, C determines a quotient group $H_2(G, C)$ of $H_2(G)$, and finally ν determines a certain torsor $H_h = H_2(G, C, \nu)$ over $H_2(G, C)$. The Conway–Parker theorem says that the natural map $\pi_0(\mathsf{Hur}_h) \to H_h$ is bijective whenever $\min_i \nu_i$ is sufficiently large.

4A. ***The Schur multiplier $H_2(G)$.*** A stem extension of G is a central extension G^* such that the kernel of $G^* \to G$ is in the derived group of G^*. A stem extension of maximal order has kernel canonically isomorphic to the cohomology group $H_2(G)$. This kernel is by definition the Schur multiplier. A stem extension of maximal order is called a Schur cover. A given group can have nonisomorphic Schur covers, but this ambiguity never poses problems for us here.

4B. ***The reduced Schur multiplier $H_2(G, C)$.*** If x, y are commuting elements of G, they canonically define an element $\langle x, y\rangle \in H_2(G)$: the commutator of lifts of x, y to a Schur cover. (In the context of this paper, there should be no confusion of this symbol with the group generated by x, y). This pairing is independent of the choice of Schur cover. In fact, a more intrinsic description is that $\langle x, y\rangle$ is the pushforward of the fundamental class of $H_2(\mathbb{Z}^2)$ under the map $\mathbb{Z}^2 \to G$ given by $(m, n) \mapsto x^m y^n$.

Fix a stem extension of maximal order $\widetilde{G} \to G$. For a conjugacy class C_i and a list of conjugacy classes $C = (C_1, \dots, C_r)$ respectively, define subgroups of the Schur multiplier

$$H_2(G)_{C_i} = \{\langle g, z\rangle : g \in C_i \text{ and } z \in Z(g)\}, \tag{4-1}$$

$$H_2(G)_C = \sum H_2(G)_{C_i}. \tag{4-2}$$

Here $Z(g)$ denotes the centralizer of g in G. The reduced Schur multiplier is then the corresponding quotient group $H_2(G, C) = H_2(G)/H_2(G)_C$.

A choice of Schur cover $\widetilde{G}$ determines a reduced Schur cover $\widetilde{G}_C = \widetilde{G}/H_2(G)_C$. The corresponding short exact sequence

$$H_2(G, C) \hookrightarrow \widetilde{G}_C \twoheadrightarrow G$$

plays an essential role in our study.

In a degree-d central extension $\pi : G^* \to G$, the preimage of a conjugacy class D consists of a certain number s of conjugacy classes, all of size $(d/s)|D|$. Always s divides d. If $s = d$ then D is called *split.* By construction, all the C_i are split in $\widetilde{G}_C$, and $\widetilde{G}_C$ is a maximal extension with this property. For more information on reduced Schur multipliers, see [Ellenberg et al. 2013, §7, v1].

4C. ***Torsors $H_2(G, C, \nu)$.*** For $i = 1, \dots, r$, let $H_2(G, C, i)$ be the set of conjugacy classes of $\widetilde{G}_C$ that lie in the preimage of the class C_i. If $\tilde{z}$ and $\tilde{g}$ are lifts to $\widetilde{G}_C$ of the identity $z = 1$ and $g \in C_i$ respectively, then one can multiply $\tilde{z} \in H_2(G, C)$ and $[\tilde{g}] \in H_2(G, C, i)$ to get $[\tilde{z}\tilde{g}] \in H_2(G, C, i)$. This multiplication operator turns each $H_2(G, C, i)$ into a torsor over $H_2(G, C)$.

One can multiply torsors over an abelian group: if T_1 and T_2 are torsors over an abelian group Z, then their product is $(T_1 \times T_2)/Z$, where all $(zt_1, z^{-1}t_2)$ have been identified. In our setting, one has a torsor

$$H_h := H_2(G, C, \nu) = \prod_i H_2(G, C, i)^{\nu_i}. \tag{4-3}$$

Note that H_h is naturally identified with the trivial torsor if all ν_i are multiples of the exponent of $H_2(G, C)$. Namely the product $\prod a_i^{\nu_i}$ is independent of choices $a_i \in H_2(G, C, i)$, and gives a distinguished element of $H_2(G, C, \nu)$. In particular, this distinguished element is fixed under $\mathrm{Aut}(G, C)$ (see Section 4E for a more detailed discussion of functoriality).

4D. ***The lifting map.*** Suppose we are given $(g_1, \dots, g_n) \in \mathcal{G}_h$. Lift each g_i to an element $\tilde{g}_i \in \widetilde{G}_C$ arbitrarily, subject to the unique condition that the product of the $\tilde{g}_i$ is the identity:

$$\tilde{g}_1 \cdots \tilde{g}_n = 1 \in \widetilde{G}_C.$$

Then each $\tilde{g}_i$ determines an element $[\tilde{g}_i] \in H_2(G, C, i)$. Their product is an element $\prod[\tilde{g}_i] \in H_2(G, C, \nu)$, independent of choices. This product is moreover unchanged if we replaced $(g_1, \dots, g_n)$ by another element in its Br_ν-orbit, or if we replace $(g_1, \dots, g_n)$ by a G-conjugate. Thus, keeping in mind the identification $\pi_0(\mathsf{Hur}_h) = \mathcal{F}_h / \mathrm{Br}_\nu$ from (3-5), we have defined a function

$$\mathrm{inv}_h : \pi_0(\mathsf{Hur}_h) \to H_h. \tag{4-4}$$

We refer to inv_h as the lifting invariant. It has been extensively studied by Fried and Serre; see [Bailey and Fried 2002; Serre 1990]. When a set decomposes according to lifting invariants, we indicate this decomposition by subscripts. Thus, e.g., $\mathcal{F}_h = \bigsqcup \mathcal{F}_{h,\ell}$ and $\mathcal{G}_h = \bigsqcup \mathcal{G}_{h,\ell}$.

The map (4-4) is equivariant with respect to the natural actions of $\mathrm{Out}(G, C)$ and so we can pass to the quotient. Writing $H_h^* = H_h / \mathrm{Out}(G, C)$, we obtain

$$\mathrm{inv}_h^* : \pi_0(\mathsf{Hur}_h^*) \to H_h^*. \tag{4-5}$$

Again we denote lifting invariants by subscripts, so that $\mathcal{F}_{h,\ell}^* = \mathcal{F}_{h,\ell} / \mathrm{Out}(G, C)_\ell$ for example, where $\mathrm{Out}(G, C)_\ell$ is the stabilizer of ℓ inside $\mathrm{Out}(G, C)$.

Note that algebraic structure is typically lost in the process of passing from objects to their corresponding starred objects. Namely, at the unstarred level one has a group $H_2(G, C)$ and its many torsors H_h. At the starred level, $H_2^*(G, C)$ is typically no longer a group, the sets H_h^* are no longer torsors, and the cardinality of H_h^* can depend on ν. Our main theorem makes direct reference only to H_h^*. However in the proof we systematically lift from H_h^* to H_h, to make use of the richer algebraic properties.

We finally note for later use that there are asymptotic mass formulas for $\mathcal{F}_{h,\ell}$ and $\mathcal{F}_{h,\ell}^*$ that are very similar to (3-7). Indeed, they are derived simply by applying (3-7) to $\widetilde{G}_C$ together with liftings of the conjugacy classes C_i:

$$|\mathcal{F}_{h,\ell}| \sim \frac{|\mathcal{F}_h|}{|H_2(G, C)|}, \qquad |\mathcal{F}_{h,\ell}^*| \sim \frac{|\mathcal{F}_{h,\ell}|}{|\mathrm{Out}(G, C)_\ell|}. \tag{4-6}$$

4E. ***Functoriality.*** Suppose we are given a surjection $f : G \to H$ of groups, together with conjugacy classes C_i in G, and set $D_i = f(C_i)$. This clearly induces a map $H_2(G, C) \to H_2(H, D)$. The functoriality of the torsors is less obvious, because of the lack of uniqueness in a Schur cover. For this, we use a more intrinsic presentation:

Amongst central extensions $\widetilde{G} \to G$ equipped with a lifting $\widetilde{C}_i$ of each C_i, there is a universal one $\widetilde{G}^*$, unique up to unique isomorphism [Ellenberg et al. 2013, Theorem 7.5.1]. Now consider the central extension $G \times \mathbb{Z}^r \to G$, where we lift C_i to $C_i \times e_i$, with e_i the i-th coordinate vector. This gives a canonical map $\alpha : \widetilde{G}^* \to G \times \mathbb{Z}^r$, and we define $H_2(G, C, \nu)_{\text{univ}}$ to be the preimage of $e \times \nu \in G \times \mathbb{Z}^r$.

This is closely related to the previous definition. Note that if we fix lifts $C_i^* \subset \widetilde{G}_C$ of each C_i, we get an induced map $\beta : \widetilde{G}^* \to \widetilde{G}_C$ from the universal property. This induces a bijection of $H_2(G, C, \nu)_{\text{univ}}$ with $H_2(G, C)$; indeed, the canonical map

$$\beta \times_G \alpha : \widetilde{G}^* \to \widetilde{G}_C \times_G (G \times \mathbb{Z}^r) \tag{4-7}$$

is an isomorphism (again, [Ellenberg et al. 2013, Theorem 7.5.1]).

So a choice of lifts C_i^* gives a distinguished element $c_\nu \in H_2(G, C, \nu)_{\text{univ}}$ — the preimage of the identity in $H_2(G, C)$. Moreover, if we replace C_i^* by $z_i C_i^*$, where $z_i \in H_2(G, C)$, then the associated map $\widetilde{G}^* \to \widetilde{G}_C$ is multiplied by the composite map $\widetilde{G}^* \to \mathbb{Z}^r \to \widetilde{G}_C$, where the second map sends $e_i \in \mathbb{Z}^r$ to z_i. Thus, with this replacement, the identification $H_2(G, C, \nu) \xrightarrow{\sim} H_2(G, C)$ has been multiplied by z^{ν_i}; in other words, the distinguished element is replaced by $\prod z_i^{-\nu_i} c_\nu$.

This construction exhibits an identification of torsors

$$H_2(G, C, \nu)_{\text{univ}} \simeq H_2(G, C, \nu)^{-1}, \tag{4-8}$$

where we write $T_1 \simeq T_2^{-1}$ for two A-torsors if there is an identification of T_1 and T_2 transferring the A-action on T_1 to the inverse of the A-action on T_2.

In fact, with respect to the identification (4-8), our lifting invariant corresponds to the lifting invariant of [Ellenberg et al. 2013]: In that paper, the authors take $(g_1, \dots, g_r)$ and associate to it the lifting invariant $\Pi = \prod \tilde{g}_i \in H_2(G, C, \nu)_{\text{univ}}$, where $\tilde{g}$ is the lift to a universal central extensions equipped with lifting. Fix $\widetilde{G}_C$, C_i^* and a morphism $\widetilde{G}^* \to \widetilde{G}_C$ as above. Choose $z_i \in H_2(G, C)$ such that the image of Π in $H_2(G, C)$ coincides with $\prod z_i^{\nu_i}$. Then $\prod \tilde{g}_i$ is carried to $\prod z_i^{\nu_i}$ multiplied by the distinguished element of $H_2(G, C, \nu)_{\text{univ}}$. On the other hand, the lifting invariant as we have defined it above equals $[C_i^* z_i^{-1}] \in H_2(G, C, \nu)$, which equals $\prod z_i^{-\nu_i}$ times the corresponding element of $H_2(G, C, \nu)$.

Now — returning to the surjection $G \to H$ — take a universal extension $\widetilde{H}^* \to H$ equipped with a lifting of the D_i, and consider $G \times_H \widetilde{H}^* \to G$; it's a central extension and it is equipped with a lifting of C_i, namely $C_i \times_H D_i^*$. There is thus a canonical map $\widetilde{G}^* \to \widetilde{H}^*$. Taking fibers above $\nu \in \mathbb{Z}^r$ gives the desired map

$$f_* : H_2(G, C, \nu)_{\text{univ}} \to H_2(H, D, \nu)_{\text{univ}},$$

and by inverting one obtains the desired map $H_2(G, C, \nu) \to H_2(H, D, \nu)$. In particular, one easily verifies that if $H = G$ and $G \to H$ is an inner automorphism, the induced map on $H_2(G, C, \nu)$ is trivial.

Finally, suppose ν is chosen to be simultaneously divisible by the order of $H_2(G, C)$ and $H_2(H, D)$ (i.e., each ν_i is so divisible). Then in fact the map $H_2(G, C, \nu) \to H_2(H, D, \nu)$ respects the natural identifications of both sides with $H_2(G, C)$ and $H_2(H, D)$ (see after (4-3)). In fact, one has natural identifications

$$H_2(G, C, \nu_1 + \nu_2) \simeq H_2(G, C, \nu_1) \times H_2(G, C, \nu_2)/H_2(G, C),$$

where the action of $z \in H_2(G, C)$ on the right is as $z : (t_1, t_2) \mapsto (t_1 z, z^{-1} t_2)$. These identifications are easily seen to be compatible with the map $H_2(G, C, \nu) \to H_2(H, D, \nu)$. Now choose C_i^* and D_i^* as above, giving rise to corresponding elements $c_\nu \in H_2(G, C, \nu)$ and $d_\nu \in H_2(H, D, \nu)$. Write $f_* c_\nu = \gamma_\nu d_\nu$ for some $\gamma_\nu \in H_2(H, D)$; then our comments show that $\gamma_{\nu_1+\nu_2} = \gamma_{\nu_1}\gamma_{\nu_2}$, and the claim follows: if ν is divisible by the order of $H_2(H, D)$, then γ_ν will be trivial.

4F. ***The Conway–Parker theorem.*** We will use a result due to Conway and Parker [1988] in the important special case where $H_2(G, C)$ is trivial. This result is also described in [Fried and Völklein 1991, Appendix] and [Malle and Matzat 1999, III.6.3]. We need the following generalization to nontrivial $H_2(G, C)$:

Proposition 4.1. *Consider Hurwitz parameters $h = (G, C, \nu)$ for (G, C) fixed and ν varying. Suppose that all the C_i are distinct. For sufficiently large* $\min_i \nu_i$, *the lifting invariant map* $\mathrm{inv}_h : \pi_0(\mathsf{Hur}_h) \to H_h$ *is bijective.*

The generalization is proved in [Ellenberg et al. 2013, §7, Theorem 7.5.1]. Because of the importance of Proposition 4.1 to this paper, we give an overview of the proof here:

Overview of proof of Proposition 4.1. First we reprise, with a few more details, the setting of Section 4E. Consider pairs $(f : G^* \to G, s)$ of a central extension of G together with a section s of f over each C_i, equivariant under conjugation, i.e.,

$$s(f(x) g f(x)^{-1}) = x s(g) x^{-1}$$

for $x \in G^*, g \in \bigcup C_i$. There is an initial object $(f^* : \widetilde{G}^* \to G, s^*)$ in the category of such pairs, i.e., a "universal central extension with section over each C_i"; in fact, we describe this initial object explicitly in the penultimate paragraph of this overview.

As discussed before (4-7), there is a natural homomorphism $\widetilde{G}^* \to G \times \mathbb{Z}^r$. Consider the sets $\mathcal{F}_h, \mathcal{G}_h$ described in Section 3D; the map sending g_i to $[g_i]$ gives a well-defined map $\mathcal{G}_h/\mathrm{Br}_\nu \to \widetilde{G}^*$, and in fact

$$\mathcal{F}_h/\mathrm{Br}_\nu \xrightarrow{I} \text{fiber of } \widetilde{G}^* \text{ above } (e, \nu).$$

As we explained in Section 4E, this map is the lifting invariant, up to the identification discussed around (4-8). We must verify that I is a bijection when all of the ν_i are large enough. The remainder of the argument is close to the argument in the appendix to [Fried and Völklein 1991]:

Consider the monoid given by $S = \bigsqcup_{n \geq 0} \big(\bigcup C_i\big)^n / \operatorname{Br}_n$, with multiplication given by concatenation. For each $g \in C_i$ let $[g]$ be the corresponding element of S (corresponding to $n = 1$). Consider inside this monoid the element

$$U = (\underbrace{x_1, x_1, \ldots, x_1}_{|G|}, \underbrace{x_2, x_2, \ldots, x_2}_{|G|}, x_3, \ldots)$$

given by taking each element of each C_i exactly $|G|$ times in succession, after fixing any ordering of such elements. Then U is central, i.e., commutes with all of S. Therefore, we may formally invert U, i.e., form the group $S[U^{-1}]$. Note that U is "divisible" by each $[g]$, and therefore each $[g]$ is invertible; consequently, $S[U^{-1}]$ is a group. Then $f : [g] \mapsto g$ defines a homomorphism $S[U^{-1}] \to G$ with central kernel; moreover, $s : g \mapsto [g]$ gives a section of this homomorphism over $\bigcup C_i$. Then it is easily verified that $(f : S[U^{-1}] \to G, s)$ is a universal central extension.

Suppose that $a = (g_1, \ldots, g_n)$, $b = (g_1', \ldots, g_n') \in \mathcal{F}_h$ have the same image in $\widetilde{G}^*$. The above construction of $\widetilde{G}^*$ shows that $(g_1, \ldots, g_n) \cdot U^k = (g_1', \ldots, g_n') \cdot U^k$ inside the semigroup S, i.e., a and b become braid-equivalent after concatenating sufficiently many copies of U. However, an elementary group-theoretic computation (see the appendix of [Fried and Völklein 1991]) shows that this implies — if $\min_i \nu_i$ is large enough — that a and b are themselves braid-equivalent. □

Various comments on Proposition 4.1 are in order. First, the condition that $\min_i \nu_i$ is sufficiently large carries on passively to many of our later considerations. We will repeat it explicitly several times but also refer to it by the word *asymptotically.*

Second, there are a number of equivalent statements. The direct translation of the bijectivity of $\pi_0(\mathsf{Hur}_h) \to H_h$ into group theory is that each fiber of $\mathcal{F}_h \to H_h$ is a single orbit of Br_ν. Alternatively, one could compose the cover $\mathsf{Hur}_h \to \mathsf{Conf}_\nu$ with the cover $\mathsf{Conf}_\nu \to \mathsf{Conf}_n$ and state the result in terms of actions of the full braid group Br_n; this is the viewpoint of both [Fried and Völklein 1991, Appendix] and [Malle and Matzat 1999, III.6.3]

Third, quotienting by $\operatorname{Out}(G, C)$ one gets a similar statement: the resulting map $\operatorname{inv}_h^* : \pi_0(\mathsf{Hur}_h^*) \to H_h^*$ is asymptotically bijective. This is the version that our full-monodromy theorem refines for certain (G, C). Note that a complication not present in Proposition 4.1 itself appears at this level: the cardinality of $H_h^* = H_h / \operatorname{Out}(G, C)$ can be dependent on ν.

5. The full-monodromy theorem

In this section, we state the full-monodromy theorem. Involved in the statement is a homological condition. We clarify the nature of this condition by giving instances when it holds and instances when it fails.

5A. ***Preliminary definitions.*** In this section, we define the notions of *pseudosimple*, *unambiguous*, and *quasifull*. All three of these notions figure prominently in the statement of the full-monodromy theorem.

We say that a centerless finite group G is *pseudosimple* if its derived group G' is a power of a nonabelian simple group and any nontrivial quotient group of G is abelian. Thus, there is an extension

$$G' \to G \to G^{\mathrm{ab}}, \tag{5-1}$$

where $G' \simeq T^w$, with T nonabelian simple, and the action of G^{ab} on T^w is transitive on the w simple factors. (Our terminology is meant to be reminiscent of similar standard terms for groups closely related to a nonabelian simple group T: *almost simple* groups are extensions $T.A$ contained in $\mathrm{Aut}(T)$ and *quasisimple* groups are quotients $M.T$ of the Schur cover $\widetilde{T}$.)

We say that a conjugacy class C_i in a group G is *ambiguous* if the G' action on C_i by conjugation has more than one orbit. If it has exactly one orbit we say that C_i is *unambiguous*. These are standard notions and for many G the division of classes into ambiguous and unambiguous can be read off from an Atlas page [Conway et al. 1985].

Essentially repeating a definition from the introduction, we say that the action of a group Γ on a set X is *full* if the image of Γ in $\mathrm{Sym}(X)$ contains the alternating group $\mathrm{Alt}(X)$. Generalizing now, we say the action is *quasifull* if the image contains $\mathrm{Alt}(X_1) \times \cdots \times \mathrm{Alt}(X_s)$, where the X_i are the orbits of Γ on X. Again we transfer the terminology to a topological setting. Thus a covering X of a connected space Y is quasifull if for any $y \in \mathsf{Y}$, the monodromy action of $\pi_1(\mathsf{Y}, y)$ on the fiber X_y is quasifull.

5B. ***Fiber powers of Hurwitz parameters.*** This subsection describes how a Hurwitz parameter $h = (G, C, \nu)$ and a positive integer k give a triple $h^k = (G^{[k]}, C^k, \nu)$. Part of this notion, in the special case $k = 2$, appears in the statement of the main theorem. The general notion plays a central role in the proof.

In general, if G is a finite group with abelianization G^{ab}, we can consider its k-fold fiber power

$$G^{[k]} = G \times_{G^{\mathrm{ab}}} \cdots \times_{G^{\mathrm{ab}}} G.$$

Note that even when $G = T^w.G^{\mathrm{ab}}$ is pseudosimple, the fiber powers $G^{[k]} = T^{wk}.G^{\mathrm{ab}}$ for $k \geq 2$ are not, because G^{ab} does not act transitively on the factors.

If C_i is a conjugacy class in a group G, we can consider its Cartesian powers $C_i^k \subseteq G^{[k]}$. In general, C_i^k is only a union of conjugacy classes. However, if C_i is unambiguous then C_i^k is a single class.

If $C = (C_1, \ldots, C_r)$ is a list of conjugacy classes, we can consider the corresponding list $(C_1^k, \ldots, C_r^k)$. Generation of G by the C_i does not imply generation of $G^{[k]}$ by the C_i^k. However, if G is pseudosimple then this implication does hold. (This can be easily deduced, for example, using the Goursat lemma, in the form of Lemma 6.1.) Thus if G is pseudosimple and C consists only of unambiguous classes, the triple h^k is a Hurwitz parameter.

Suppose, then, that G is pseudosimple and C consists of unambiguous classes. The natural map (Section 4E)

$$H_2(G^{[k]}, C^k, \nu) \to H_2(G, C, \nu)^k$$

is surjective. This surjectivity can be seen by interpreting both sides in terms of connected components (in the large ν limit) via the Conway–Parker theorem. Surjectivity can also be seen because the map is equivariant with respect to the natural map $H_2(G^{[k]}, C^k) \to H_2(G, C)^k$, which is surjective by homological algebra, as we explain after (5-2).

5C. ***Statement.*** With our various definitions in place, we can state the main result of this paper:

Theorem 5.1 (full-monodromy theorem). *Let G be a finite centerless nonabelian group, let $C = (C_1, \ldots, C_r)$ a list of distinct nonidentity conjugacy classes generating G, and consider Hurwitz parameters $h = (G, C, \nu)$ for varying allowed $\nu \in \mathbb{Z}_{\geq 1}^r$. Then the following two statements are equivalent*:

I: 1. *G is pseudosimple*,
2. *the classes C_i are all unambiguous*, *and*
3. $|H_2(G^{[2]}, C^2)| = |H_2(G, C)|^2$.

II: *All covers $\mathsf{Hur}_h^* \to \mathsf{Conf}_\nu$ are quasifull whenever $\min_i \nu_i$ is sufficiently large.*

Note that Statement II can equivalently be presented in terms of fullness: *for $\min_i \nu_i$ sufficiently large, the covers $\mathsf{Hur}_{h,\ell}^* \to \mathsf{Conf}_\nu$ are full and pairwise nonisomorphic as ℓ ranges over H_h^*.* Note also that a pseudosimple group G is simple if and only if G^{ab} is trivial. In this case, Conditions 2 and 3 of Statement I are trivially satisfied and the direction I $\Longrightarrow$ II becomes the statement highlighted in Section 1B.

For the more important direction I $\Longrightarrow$ II, the condition that $\min_i \nu_i$ is sufficiently large is simply inherited from the Conway–Parker theorem. Calculations suggest that the covers $\mathsf{Hur}_h^* \to \mathsf{Conf}_\nu$ tend to be quasifull even when all ν_i are small. We are not pursuing the important question of effectivity here, but we note that effective statements of fullness are obtained for certain classical Hurwitz parameters in [Kluitmann 1988].

Given (G, C), whether or not Conditions 1 and 2 hold is immediately determinable in practice. Evaluating Condition 3 is harder in general, and the next two subsections are devoted to giving an easily checkable reformulation applicable in many cases (Proposition 5.2) and showing (Corollary 5.3) that it sometimes fails.

5D. ***The homological condition for G of split-cyclic type.*** We say that a pseudosimple group G has *split* type if the canonical surjection $\pi : G \to G^{\mathrm{ab}}$ has a homomorphic section $s : G^{\mathrm{ab}} \to G$. Inspecting individual Atlas pages [Conway et al. 1985] shows that this a priori strong condition is actually commonly satisfied. Similarly, we say that a pseudosimple group has *cyclic* type if G^{ab} is cyclic. Again this strong-seeming condition is commonly satisfied, as indeed for a simple group T all of $\mathrm{Out}(T)$ is often cyclic [Conway et al. 1985, Chapter 1, Table 1; Chapter 3, Table 5]. When both of these conditions are satisfied, we say that G is of *split-cyclic* type.

For G of split-cyclic type, the following proposition says that Condition 3 of Theorem 5.1 is equivalent to an apparent strengthening $\hat{3}$. Moreover, these two conditions are both equivalent to a more explicit condition E which makes no reference to either fiber powers or powers. For E, we modify the notions defined in Section 4B as follows:

$$H_2'(G)_{C_i} = \{\langle g, z\rangle : g \in C_i \text{ and } z \in Z(g) \cap G'\},$$
$$H_2'(G)_C = \sum H_2'(G)_{C_i}.$$

These are straightforward variants, as indeed if one removes every $'$ one recovers the definitions (4-1) and (4-2) of the previous notions.

Proposition 5.2. *Let G be a pseudosimple group of split-cyclic type, and let $C = (C_1, \dots, C_r)$ be a list of distinct unambiguous conjugacy classes. Then the following are equivalent*:

3. $|H_2(G^{[2]}, C^2)| = |H_2(G, C)|^2$.

$\hat{3}$. $|H_2(G^{[k]}, C^k)| = |H_2(G, C)|^k$ *for all positive integers* k.

E. $H_2(G)_C = H_2'(G)_C$.

Moreover, *if* $|G^{\mathrm{ab}}|$ *is relatively prime to* $|H_2(G)|$ *then all three conditions hold.*

Proof. All three conditions involve the list C of conjugacy classes. We begin however with considerations involving G only. The k different coordinate projections $G^{[k]} \to G$ together induce a map $f_k : H_2(G^{[k]}) \to H_2(G)^k$. We first show that the assumption that G has split-cyclic type implies all the f_k are isomorphisms. We present this deduction in some detail because we will return to parts of it in Section 6E.

The map f_k is part of a morphism of five-term exact sequences (see [Eckmann and Stammbach 1970, Theorem 5.2], noting that $H_1(G')=0$)

$$\begin{array}{ccccccccc}
H_3(G^{[k]}) & \xrightarrow{\pi_3^{[k]}} & H_3(G^{\mathrm{ab}}) & \xrightarrow{\delta^{[k]}} & H_2(G'^k)_{G^{\mathrm{ab}}} & \xrightarrow{i_2^{[k]}} & H_2(G^{[k]}) & \xrightarrow{\pi_2^{[k]}} & H_2(G^{\mathrm{ab}}) \\
\downarrow & & \downarrow \Delta_3 & & \downarrow \simeq & & \downarrow f_k & & \downarrow \Delta_2 \\
H_3(G)^k & \xrightarrow{\pi_3^k} & H_3(G^{\mathrm{ab}})^k & \xrightarrow{\delta^k} & H_2(G')^k_{G^{\mathrm{ab}}} & \xrightarrow{i_2^k} & H_2(G)^k & \xrightarrow{\pi_2^k} & H_2(G^{\mathrm{ab}})^k
\end{array} \tag{5-2}$$

Each five-term sequence arises from the Hochschild–Serre spectral sequence associated to an exact sequence of groups. The top sequence comes from the k-th fiber power of $G' \overset{i}{\hookrightarrow} G \overset{\pi}{\twoheadrightarrow} G^{\mathrm{ab}}$, while the bottom sequence comes from the k-th ordinary Cartesian power.

We note that (5-2) actually shows that $H_2(G^{[k]}, C^k) \to H_2(G, C)^k$ is surjective whenever G is pseudosimple and C consists of unambiguous classes. The point is that $H_2(G)_C$ surjects onto $H_2(G^{\mathrm{ab}})$. That is because $H_2(G^{\mathrm{ab}})$ is generated by symbols $\langle \alpha, \beta \rangle$. But such a symbol belongs to the image of $H_2(G)_C$, since the $[C_i]$ generate G^{ab} and, for any $g \in C_i$, the centralizer $Z(g)$ surjects to G^{ab} because C_i is unambiguous.

The assumption that $\pi : G \to G^{\mathrm{ab}}$ has a splitting s drastically simplifies (5-2). From $\pi \circ s = \mathrm{Id}_{G^{\mathrm{ab}}}$, one obtains that $\pi_3^{[k]} \circ s_3^{[k]}$ and $\pi_3^k \circ s_3^k$ are the identity on $H_3(G^{\mathrm{ab}})$ and $H_3(G^{\mathrm{ab}})^k$, respectively. Thus $\pi_3^{[k]}$ and π_3^k are both surjective and so the boundary maps $\delta^{[k]}$ and δ^k are both 0. Thus the part of (5-2) relevant for us becomes

$$\begin{array}{ccccc}
H_2(G'^k)_{G^{\mathrm{ab}}} & \hookrightarrow & H_2(G^{[k]}) & \twoheadrightarrow & H_2(G^{\mathrm{ab}}) \\
\downarrow \simeq & & \downarrow f_k & & \downarrow \Delta_2 \\
H_2(G')^k_{G^{\mathrm{ab}}} & \hookrightarrow & H_2(G)^k & \twoheadrightarrow & H_2(G^{\mathrm{ab}})^k
\end{array} \tag{5-3}$$

We have suppressed some notation, since we have no further use for it.

The assumption that G^{ab} is cyclic is equivalent to the assumption that $H_2(G^{\mathrm{ab}})$ is the zero group. Thus exactly in this situation one gets the independent simplification of (5-2) where the last column becomes the zero map between zero groups. Applied to (5-3) it says that $f_k : H_2(G^{[k]}) \to H_2(G)^k$ is an isomorphism. We henceforth use f_k to identify $H_2(G^{[k]})$ with $H_2(G)^k$.

We now bring in the list C of conjugacy classes. We have a morphism of short exact sequences

$$\begin{array}{ccccc}
H_2(G^{[k]})_C & \hookrightarrow & H_2(G^{[k]}) & \twoheadrightarrow & H_2(G^{[k]}, C^k) \\
\cap\!| & & \| & & \downarrow \\
H_2(G)^k_C & \hookrightarrow & H_2(G)^k & \twoheadrightarrow & H_2(G, C)^k
\end{array} \tag{5-4}$$

Since the map in the right column is surjective, Conditions 3 and $\hat{3}$ become that it is an isomorphism for $k = 2$ and all k respectively. So they are equivalent to the inclusion in the left column being equality, again for $k = 2$ and all k respectively. We work henceforth with these versions of Conditions 3 and $\hat{3}$.

Trivially

$$|H_2(G)_C| = |H_2'(G)_C| \cdot |H_2(G)_C/H_2'(G)_C|. \tag{5-5}$$

But also the image of $H_2(G^{[k]})_{C^k}$ in $(H_2(G)/H_2'(G)_C)^k$ is exactly the diagonal image of $H_2(G)_C/H_2'(G)_C$. To see this, note that $H_2(G^{[k]})_{C^k}$ is generated by

$$(\langle g, z_1\rangle, \dots, \langle g, z_k\rangle),$$

where $g \in \bigcup C_i$, each $z_i \in Z(g)$, and $z_1 \equiv \cdots \equiv z_k$ modulo G'. In particular, it certainly contains the diagonal image of $H_2(G)_C$. On the other hand, the images of $\langle g, z_i\rangle$ inside $H_2(G)_C/H_2'(G)_C$ are equal to each other, since $\langle g, z_i z_j^{-1}\rangle \in H_2'(G)_C$.

Moreover, $H_2'(G)_C^k \subseteq H_2(G^{[k]})_{C^k}$. This inclusion holds because, for any $g \in C_i$ and $z \in Z(g) \cap G'$, we have

$$(\langle g, z\rangle, 0, 0, \dots) \in H_2(G^{[k]})_{C^k},$$

since we can regard the left-hand side as $(\langle g, z\rangle, \langle g, e\rangle, \langle g, e\rangle, \dots)$, and similarly for any other “coordinate”. Therefore,

$$|H_2(G^{[k]})_{C^k}| = |H_2'(G)_C|^k \cdot |H_2(G)_C/H_2'(G)_C|. \tag{5-6}$$

Dividing the k-th power of (5-5) by (5-6), one gets

$$\frac{|H_2(G)_C|^k}{|H_2(G^{[k]})_{C^k}|} = |H_2(G)_C/H_2'(G)_C|^{k-1}. \tag{5-7}$$

Condition 3 says the left side is 1 for $k = 2$. Condition $\hat{3}$ says the left side is 1 for all k. Equation (5-7) says that each of these is equivalent to $H_2(G)_C = H_2'(G)_C$, which is exactly Condition E.

For the final statement, $|H_2(G)_{C_i}/H_2'(G)_{C_i}|$ clearly divides $|H_2(G)|$. It also divides $|G^{\mathrm{ab}}|$, because $Z(g)/(Z(g) \cap G')$ surjects onto $H_2(G)_{C_i}/H_2'(G)_{C_i}$ via $z \in Z(g) \mapsto \langle g, z\rangle$, for any fixed $g \in C_i$. So, if $|H_2(G)|$ and $|G^{\mathrm{ab}}|$ are relatively prime then $H_2(G)_{C_i} = H_2'(G)_{C_i}$ always, and so Condition E holds. □

5E. ***The homological condition for G of split-p-p type.*** For p a prime, we say that a pseudosimple group G has *split-p-p* type if $G \to G^{\mathrm{ab}}$ is split and

$$|G^{\mathrm{ab}}| = |H_2(G)| = p.$$

Even this seemingly very special case is common. For example, taking $p = 2$, it includes

- all six extensions $T.A$ of sporadic groups T with A and $H_2(T.A)$ nontrivial,
- all S_d with $d \geq 5$, and
- all $\mathrm{PGL}_2(q)$ for odd $q \geq 5$.

To illustrate the tractability of Condition E of Proposition 5.2, we work it out explicitly for groups G of split-p-p type. Explicating Condition E for the full split-cyclic case would be similar but combinatorially more complicated.

For G of split-p-p type, we divide its unambiguous classes into three types. Let $\widetilde{G}$ be a Schur cover of G. An unambiguous class C is *split* if its preimage $\widetilde{C}$ consists of p conjugacy classes in $\widetilde{G}$. It is *mixed* if $\widetilde{C}$ is p different $\widetilde{G}'$ conjugacy classes but just one $\widetilde{G}$ class. Otherwise a class C is *inert*. Mixed classes are necessarily in the derived group, but split and inert classes can lie above any element of G^{ab}.

Corollary 5.3. *Let G be a pseudosimple group of split-p-p type and let $C = (C_1, \ldots, C_r)$ be a list of unambiguous classes. Then Condition E fails exactly when there are no inert classes and at least one mixed class among the C_i.*

Proof. We are considering subgroups of the p-element Schur multiplier $H_2(G)$. The subgroups have the following form:

C_i	split	mixed	inert
$H_2'(G)_{C_i}$	0	0	$H_2(G)$
$H_2(G)_{C_i}$	0	$H_2(G)$	$H_2(G)$

Thus $H_2'(G)_C = \sum_i H_2'(G)_{C_i}$ is a proper subgroup of $H_2(G)_C = \sum_i H_2(G)_{C_i}$ exactly under the conditions stated in the corollary. □

For a group $T.p$, the types of classes can be determined from an Atlas-style character table, including its lifting row and fusion column. For example, for the six sporadic T mentioned above, the mixed classes in $T.2$ are exactly as follows:

Mathieu$_{12}$	Mathieu$_{22}$	Hall–Janko	Higman–Sims	Suzuki	Fischer$_{22}$
10A	8A	8A	4A, 6A, 12A	12D, 12E, 24A	(15 classes)

In the sequences S_d and $\mathrm{PGL}_2(q)$, the patterns evident from character tables in the first few instances can be proved to hold in general. Namely for S_d, conjugacy classes are indexed by partitions of d. The type of a class C_λ can be read off from two features of the indexing partition λ, the number e of even parts and whether or not all parts are distinct:

	$e=0$	$e\in\{2,4,6,\dots\}$	$e\in\{1,3,5,\dots\}$
all distinct	ambiguous	mixed	split
not all distinct	split	inert	inert

Thus S_5 has no mixed classes while C_{42} and C_{421} are the unique mixed classes of S_6 and S_7 respectively. For $\mathrm{PGL}_2(q)$ with q odd, the division is even easier: the two classes of order the prime dividing q are ambiguous, the two classes of order 2 are inert, and all other classes are split. Thus for these $\mathrm{PGL}_2(q)$, the homological condition always holds.

6. Proof of I $\Longrightarrow$ II

In this section we prove the implication I $\Longrightarrow$ II of Theorem 5.1. Thus we consider Hurwitz parameters $h=(G,C,\nu)$ for fixed (G,C) satisfying Conditions 1–3 and varying ν. We then prove that the action of Br_ν on $\mathcal{F}_h^*$ is quasifull whenever $\min_i \nu_i$ is sufficiently large.

6A. ***A Goursat lemma.*** The classical Goursat lemma classifies certain subgroups of powers of a simple group. We state and prove a generalized version here. As usual, if one has groups G_1, G_2 endowed with homomorphisms π_1, π_2 to a third group Q, we say that G_1 and G_2 are isomorphic *over* Q if there is an isomorphism $i: G_1\to G_2$ satisfying $\pi_2 i=\pi_1$.

Lemma 6.1 (generalized Goursat lemma). *Suppose that G is pseudosimpleand $H\subseteq G^{[k]}$ is a "Goursat subgroup" in the sense that it surjects onto each coordinate factor. Then*:

1. *H is itself isomorphic over G^{ab} to $G^{[w]}$ for some $w\le k$.*
2. *There is a surjection $f:[1,k]\to[1,w]$ and automorphisms $\varphi_1,\dots,\varphi_k$ of G over G^{ab} such that H is the image of $G^{[w]}$ under*

$$(g_1,\dots,g_w)\mapsto(\varphi_1(g_{f(1)}),\dots,\varphi_k(g_{f(k)})).$$

Proof. We first prove (1) by induction, the base case $k=1$ being trivial. Note that the projection $\overline{H}=\pi_2(H)$ of H to the second factor in

$$G^{[k]}=G\times_{G^{\mathrm{ab}}}G^{[k-1]}$$

is also a Goursat subgroup. By induction, it is G^{ab}-isomorphic to $G^{[v]}$ for suitable v. The kernel $K=\ker(\pi_2)$ of the projection $H\to\overline{H}$ maps, under the first projection π_1, to a subgroup $\overline{K}\subseteq G'$ that is invariant under conjugation by G. In particular, either $\overline{K}$ is trivial, and we're done by induction, or $\overline{K}=G'$. In the latter case, we will show that $H=G\times_{G^{\mathrm{ab}}}\overline{H}$: Take any element $(m^*,\mu)\in G\times_{G^{\mathrm{ab}}}\overline{H}$. By assumption, there exists m in G such that $(m, \mu) \in H$, but then m and m^* have the same projection to G^{ab}, and so

$$(m^*, \mu) = (m^* m^{-1}, 1) \cdot (m, \mu)$$

lies in H also. This concludes the proof of the first assertion: H is isomorphic to $G^{[w]}$ over G^{ab} for some w.

Now we deduce (2) from (1). Let $\Theta = G^{[w]} \to H$ be any isomorphism and write $\Theta(g) = (\theta_1(g), \ldots, \theta_k(g))$. We need to show that, for each i, one can express $\theta_i(g)$ in the form $\varphi_i(g_{f(i)})$ as in (2). In other words, letting $\pi_j : G^{[w]} \to G$ be the j-th projection, we need to show that any surjective morphism $\theta : G^{[w]} \to G$ over G^{ab} factors as $\varphi \pi_j$ for some $j \in \{1, \ldots, w\}$ and some automorphism $\varphi : G \to G$ over G^{ab}.

So let $\theta : G^{[w]} \to G$ be any surjective morphism over G^{ab}. Its kernel K is a normal subgroup of $(G')^w$, invariant under $G^{[w]}$, and with index $|G'|$. Now, via the isomorphism $G' \simeq T^u$ for some nonabelian simple group T, the normal subgroups of $(G')^w \simeq T^{uw}$ are of the form $T_I = \prod_{(i,j) \in I} T_{(i,j)}$, where I is a subset of $P = \{1, \ldots, u\} \times \{1, \ldots, w\}$. The normal subgroups which are invariant under G^w are those for which the indexing set I is invariant under the natural action of G^{ab}. The orbits of G^{ab} on P are the sets $P_j = \{1, \ldots, u\} \times \{j\}$. So the kernel K of θ necessarily has the form T_{P-P_j} for some j. Thus K is also the kernel of the coordinate projection π_j. The unique bijection $\varphi : G \to G$ satisfying $\theta = \varphi \pi_j$ is then an automorphism of G over G^{ab}. □

6B. ***Identifying braid orbits.*** For F a set and k a positive integer we let

$$F^{\underline{k}} = \{(x_1, \ldots, x_k) : \text{all } x_i \text{ are different}\}.$$

If F has cardinality N then $F^{\underline{k}}$ has cardinality $N^{\underline{k}} := N(N-1)\cdots(N-k+1)$. In this subsection we assume Conditions 1 and 2 of Statement I in Theorem 5.1 and identify the quotient set $(\mathcal{F}_h^*)^{\underline{k}}/\operatorname{Br}_\nu$ asymptotically.

Begin with $x_1, \ldots, x_k \in \mathcal{F}_h^*$. Choose a set of representatives $\underline{g}_1, \ldots, \underline{g}_k \in \mathcal{G}_h$. Writing each $\underline{g}_i$ as a column vector, we get a matrix

$$(\underline{g}_1, \ldots, \underline{g}_k) = \begin{pmatrix} g_{11} & g_{21} & \cdots & g_{k1} \\ g_{12} & g_{22} & \cdots & g_{k2} \\ g_{13} & g_{23} & \cdots & g_{k3} \\ \vdots & \vdots & \ddots & \vdots \\ g_{1n} & g_{2n} & \cdots & g_{kn} \end{pmatrix}. \tag{6-1}$$

So, simply recalling our context:

- All the g_{ij} in a given row are in the same conjugacy class of G.
- These conjugacy classes are

$$\underbrace{C_1, \ldots, C_1}_{\nu_1}; \;\ldots\; ; \underbrace{C_r, \ldots, C_r}_{\nu_r}$$

as one goes down the rows, so that a given row is in some C_i^k.
- Each column in its given order multiplies to 1.
- Each column generates all of G.

All entries in a given row certainly have the same projection to G^{ab}, and so each row defines an element of $G^{[k]}$. Consider now the subgroup H of $G^{[k]}$ generated by the rows of this matrix. We are going to show that

$$H = G^{[k]} \iff \text{all } x_i \text{ are different.} \tag{6-2}$$

First of all, note that the condition that $H = G^{[k]}$ is independent of the choice of lifting from $\mathcal{F}_h^*$ to $\mathcal{G}_h$. For example, if we modify $\underline{g}_1$, the first column of (6-1), by an element $\alpha \in \mathrm{Aut}(G, C)$, then the subgroup generated by the rows simply changes by the automorphism $(\alpha, 1, 1, 1 \ldots, 1)$ of $G^{[k]}$. Note that α is automatically an isomorphism of G over G^{ab} because it preserves each C_i and they generate G^{ab}.

Now the $\Longrightarrow$ direction of (6-2) is easy: if $x_i = x_j$ for some $i \neq j$, then we could lift so that $\underline{g}_i = \underline{g}_j$, and then certainly $H \subsetneq G^{[k]}$.

Now suppose that $x_i \neq x_j$ for all $i \neq j$; we'll show that $H = G^{[k]}$. Since each column generates G, the subgroup H is a Goursat subgroup of $G^{[k]}$. Accordingly we may apply Lemma 6.1, and see that H can be constructed from a surjective function $f : [1, k] \to [1, w]$ together with a system of isomorphisms $\varphi_j : G \to G$ over G^{ab}, for $1 \leq j \leq k$. In particular, we may find $(y_1, \ldots, y_w) \in G^{[w]}$ which maps to the first row $(g_{11}, g_{21}, \ldots, g_{k1})$, so that

$$\varphi_j(y_{f(j)}) = g_{j1}, \quad 1 \leq j \leq k.$$

In particular, whenever $f(j) = f(j')$, the map

$$\varphi_{j'}\varphi_j^{-1}$$

carries g_{j1} to $g_{j'1}$ and so preserves C_1. By similar reasoning, applied to the second row, third row and so on, this map preserves *every* conjugacy class, so

$$\varphi_{j'}\varphi_j^{-1} \in \mathrm{Aut}(G, C)$$

whenever $f(j) = f(j')$. But $\varphi_{j'}\varphi_j^{-1}$ carries g_{ji} to $g_{j'i}$; that means that actually $x_j = x_j'$, and so $j = j'$. In other words, f is injective, and so $H \simeq G^{[k]}$, as desired.

Each matrix of the form (6-1) with H all of $G^{[k]}$ defines an element of $\mathcal{G}_{h^k}$. Now, the group $\mathrm{Aut}(G,C)^k$ acts on $G^{[k]}$; its image in the outer automorphism group will be called $\mathrm{Out}(G,C)^{[k]}$. This latter group maps onto $\mathrm{Out}(G,C)^k$, with kernel isomorphic to $(G^{\mathrm{ab}})^{k-1}$. These considerations give a bijective map

$$\mathcal{F}_{h^k}/\,\mathrm{Out}(G,C)^{[k]} \xrightarrow{\sim} \mathcal{F}_{h}^{*\underline{k}}. \tag{6-3}$$

This bijection is purely algebraic in nature and is valid for all ν.

Lifting invariants give a map $\mathcal{F}_{h^k}/\,\mathrm{Br}_\nu \to H_2(G^{[k]}, C^k, \nu)$. For any fixed k, the Conway–Parker theorem says that this map is asymptotically a bijection. Taking the quotient by $\mathrm{Out}(G,C)^{[k]}$ and incorporating the Goursat conclusion (6-3), we get the desired description of braid orbits:

$$\mathcal{F}_{h}^{*\underline{k}}/\,\mathrm{Br}_\nu \underset{a}{\xrightarrow{\sim}} H_2(G^{[k]}, C^k, \nu)/\,\mathrm{Out}(G,C)^{[k]}. \tag{6-4}$$

The map of (6-4) is defined for all allowed ν and, as indicated by the notation $\underset{a}{\xrightarrow{\sim}}$, is asymptotically a bijection.

There is, of course, a map $\mathcal{F}_{h}^{*\underline{k}}/\,\mathrm{Br}_\nu \to (\mathcal{F}_h^*/\,\mathrm{Br}_\nu)^k$; on the right-hand side of (6-4), this corresponds to the natural map

$$H_2(G^{[k]}, C^k, \nu)/\,\mathrm{Out}(G,C)^{[k]} \to (H_2(G,C,\nu)/\,\mathrm{Out}(G,C))^k. \tag{6-5}$$

Note that the action of $\mathrm{Out}(G,C)^{[k]}$ on $H_2(G^{[k]}, C^k, \nu)$ factors, under the coordinate projection $H_2(G^{[k]}, C^k, \nu) \to H_2(G,C,\nu)$, through the corresponding coordinate projection $\mathrm{Out}(G,C)^{[k]} \to \mathrm{Out}(G,C)$.

6C. ***End of the proof of I $\Longrightarrow$ II in the split-cyclic case.*** We now assume not only Conditions 1 and 2 of I, but also Condition 3. In this subsection, we complete the proof of I $\Longrightarrow$ II under the auxiliary assumption that the surjection $G \to G^{\mathrm{ab}}$ is split and G^{ab} is cyclic. Some of the notions introduced here are used again in Section 6E, where we complete the proof without auxiliary assumptions.

Consider the canonical surjections $H_2(G^{[k]}, C^k, \nu) \twoheadrightarrow H_2(G,C,\nu)^k$. Under our auxiliary assumption that G has split-cyclic type, Condition 3 and Proposition 5.2 show that

$$|H_2(G^{[k]}, C^k)| = |H_2(G,C)|^k$$

for all k. Thus, since cardinality does not change when one passes from groups to torsors, the surjections are bijections. Moreover, because inner automorphisms act trivially on $H_2(G,C,\nu)$, the action of $\mathrm{Out}(G,C)^{[k]}$ on $H_2(G,C,\nu)^k$ actually factors through $\mathrm{Out}(G,C)^k$.

Taking the quotient by $\mathrm{Out}(G,C)^{[k]}$, we can rewrite (6-4) as

$$\mathcal{F}_{h}^{*\underline{k}}/\,\mathrm{Br}_\nu \underset{a}{\xrightarrow{\sim}} H_2^*(G,C,\nu)^k. \tag{6-6}$$

Then standard group theory shows that the action of Br_ν on $\mathcal{F}_h^*$ is quasifull for sufficiently large $\min_i \nu_i$:

In general, consider a permutation group $B \subseteq \mathrm{Sym}(F)$ with orbit decomposition $F = \bigsqcup_{i=1}^s F_i$. Suppose each orbit F_i has size at least k. Then the induced action of B on $F^{\underline{k}}$ has at least s^k orbits. If equality holds, then the images $B_i \subseteq \mathrm{Sym}(F_i)$ of B are each individually k-transitive. If $k \geq 6$, then the classification of finite simple groups says that B_i contains $\mathrm{Alt}(F_i)$. Still assuming that B has exactly s^k orbits on $F^{\underline{k}}$, it is then elementary that B contains $\mathrm{Alt}(F_1) \times \cdots \times \mathrm{Alt}(F_s)$. In other words, B is quasifull, as desired.

6D. ***A lemma on 2-transitive groups.*** For the general case, Condition 3 gives us control over Br_ν-orbits only on pairs (x_1, x_2) of distinct elements in $\mathcal{F}_h^*$, not tuples of larger length. To deal with this problem, we replace the classification of multiply transitive groups by a statement derived from the classification of 2-transitive groups. The exact formulation of our lemma is inessential; its import is that full groups are clearly separated out from other 2-transitive groups in a way sufficient for our purpose.

Lemma 6.2. *Fix an odd integer $j \geq 5$ and a finite set X. Suppose a 2-transitive group $\Gamma \subseteq \mathrm{Sym}(X)$ satisfies $|X^{\underline{2j}}/\Gamma| \leq 2^{j^2-4j}$. If $|X|$ is sufficiently large, then Γ is full.*

Proof. To prove the statement, we use the classification of nonfull 2-transitive groups, as presented in [Dixon and Mortimer 1996, §7.7], thereby breaking our argument into a finite number of cases. For fixed j, we discard in each case a finite number of Γ and establish $|X^{\underline{2j}}/\Gamma| > 2^{j^2-4j}$ for all other Γ.

It suffices to restrict attention to maximal nonfull 2-transitive groups Γ. Besides a small number of examples involving seven of the sporadic groups [Dixon and Mortimer 1996, pp. 252–253], every such maximal Γ occurs in the following table:

#	type	Γ	degree $N := \lvert X\rvert$	order $\lvert\Gamma\rvert$
1	affine	$\mathrm{AGL}_d(p)$	p^d	
2	projective	$\mathrm{P\Gamma L}_d(q)$	$(q^d-1)/(q-1)$	
3	OS2	$O_{2d+1}(2)$	$2^d(2^d \pm 1)/2$	
4	unitary	$U_3(q)$	q^3+1	$q^3(q^2-1)(q^3+1)$
5	Suzuki	$Sz(q)$	q^2+1	$(q^2+1)q^2(q-1)$
6	Ree	$R(q)$	q^3+1	$(q^3+1)q^3(q-1)$

The six series are listed in the order they are treated in [Dixon and Mortimer 1996, pp. 244–252], with $d \geq 1$ and $d \geq 2$ in Cases 1 and 2 respectively. Throughout, p is a prime number and $q = p^e$ is a prime power. These numbers are arbitrary, except in Cases 5 and 6, where the base is $p = 2$ and $p = 3$ respectively and the exponent e is odd. The orders $|\Gamma|$ in Cases 1–3 are not needed in our argument and so are omitted from the table.

Cases 4–6. In these cases, the order $|\Gamma|$ grows only polynomially in the degree N, with $|\Gamma| < N^3$ holding always. One has

$$|X^{\underline{2j}}/\Gamma| \geq N^{\underline{2j}}/|\Gamma| > N^{\underline{2j}}/N^3.$$

For $j \geq 5$ fixed and $N \to \infty$, the right side tends to ∞. So, with finitely many exceptions, $|X^{\underline{2j}}/\Gamma| > 2^{j^2-4j}$.

Case 1. In this case, the affine general linear group $\mathrm{AGL}_d(\mathbb{F}_p)$ acts on the affine space $\mathbb{F}_p^d$. Let $w = \min(j, d+1)$. Fix $x_1, \ldots, x_w$ in $\mathbb{F}_p^d$ spanning an affine subspace A of dimension $w-1$. The set $A - \{x_1, \ldots, x_w\}$ has $p^{w-1} - w$ elements. There are $(p^{w-1} - w)^{\underline{2j-w}}$ ways to successively choose $x_{w+1}, \ldots, x_{2j}$ in A so that all the x_i are distinct. The tuples $(x_1, \ldots, x_{2j}) \in (\mathbb{F}_p^d)^{\underline{2j}}$ so obtained are in different $\mathrm{AGL}_d(\mathbb{F}_p)$-orbits. Thus

$$|(\mathbb{F}_p^d)^{\underline{2j}}/\,\mathrm{AGL}_d(\mathbb{F}_p)| \geq (p^{w-1} - w)^{\underline{2j-w}}.$$

For fixed $d < j$, so that $w = d+1$, the right side tends to ∞ with p, and so with finitely many exceptions $|(\mathbb{F}_p^d)^{\underline{2j}}/\,\mathrm{AGL}_d(\mathbb{F}_p)| > 2^{j^2-4j}$. For $d \geq j$, so that $w = j$, one gets no exceptions, as

$$(p^{w-1} - w)^{\underline{2j-w}} = (p^{j-1} - j)^{\underline{j}} \geq (2^{j-1} - j)^{\underline{j}} \geq (2^{j-1} - 2j + 1)^j > 2^{j^2-4j}.$$

(Case 1 is the only case where there is a complicated list of nonmaximal 2-transitive groups. Some large ones are $\mathrm{AGL}_{d/e}(\mathbb{F}_{p^e}) \subset \mathrm{A\Gamma L}_{d/e}(\mathbb{F}_{p^e}) \subset \mathrm{AGL}_d(p)$, for any e properly dividing d.)

Cases 2 and 3 are very similar to Case 1, but are sufficiently different to require separate treatments.

Case 2. Here $\Gamma = \mathrm{P\Gamma L}_d(\mathbb{F}_q) = \mathrm{PGL}_d(\mathbb{F}_q).\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$ acts on the projective space $X = \mathbb{P}^{d-1}(\mathbb{F}_q)$. Again let $w = \min(j, d+1)$. Fix $x_1, \ldots, x_w$ in $\mathbb{P}^{d-1}(\mathbb{F}_q)$ spanning a projective subspace P of dimension $w-1$. Similarly to Case 1, there are $((q^w - 1)/(q-1) - w)^{\underline{2j-w}}$ ways to successively choose $x_{w+1}, \ldots, x_{2j}$ in P so that all the x_i are distinct. The tuples $(x_1, \ldots, x_{2j}) \in \mathbb{P}^{d-1}(\mathbb{F}_q)^{\underline{2j}}$ so obtained are in different $\mathrm{PGL}_d(\mathbb{F}_q)$-orbits. However one $\mathrm{P\Gamma L}_d(\mathbb{F}_q)$-orbit can consist of up to e different $\mathrm{PGL}_d(\mathbb{F}_q)$-orbits. Thus our lower bound in this case is

$$|\mathbb{P}^{d-1}(\mathbb{F}_q)^{\underline{2j}}/\mathrm{P\Gamma L}_d(\mathbb{F}_q)| \geq \frac{1}{e}\left(\frac{q^w - 1}{q-1} - w\right)^{\underline{2j-w}}.$$

Again, the subcase $d < j$, where $w = d + 1$, is simple: the right side tends to ∞ with q and so $|\mathbb{P}^{d-1}(\mathbb{F}_q)^{\underline{2j}}/\mathrm{P\Gamma L}_d(\mathbb{F}_q)| > 2^{j^2-4j}$ holds with only finitely many exceptions. For $d \geq j$, so that $w = j$ again, one has no further exceptions since

$$\frac{1}{e}\left(\frac{q^w-1}{q-1} - w\right)^{\underline{2j-w}} > \frac{1}{e}(q^{j-1} - 2j + 1)^j > (2^{j-1} - 2j + 1)^j > 2^{j^2-4j}.$$

Case 3. Here the group in question, in its most familiar guise, is $\Gamma = \mathrm{Sp}_{2d}(\mathbb{F}_2)$ for $d \geq 2$. It is better in our context to view $\Gamma = O_{2d+1}(\mathbb{F}_2)$, as from this point of view the 2-transitive actions appear most naturally. In fact, the orbit decomposition of the natural action of $O_{2d+1}(\mathbb{F}_2)$ is

$$\mathbb{F}_2^{2d+1} - \{0\} = X_{-1} \sqcup X_1 \sqcup X_0.$$

Here X_0 is the set of isotropic vectors. The pair $(O_{2d+1}(\mathbb{F}_2), X_0)$ is a copy of the more standard pair $(\mathrm{Sp}_{2d}(\mathbb{F}_2), \mathbb{F}_2^{2d} - \{0\})$, and so in particular $|X_0| = 2^{2d} - 1$. A nonisotropic vector is in X_1 if its stabilizer is the split orthogonal group $O_{2d}^+(\mathbb{F}_2)$ and is in X_{-1} if its stabilizer is the nonsplit orthogonal group $O_{2d}^-(\mathbb{F}_2)$. From the order of the stabilizers, one gets that $|X_\epsilon| = 2^{d-1}(2^d + \epsilon)$. While the action of Γ on $X_0^{\underline{2}}$ has two orbits, the actions on the other two X_ϵ are 2-transitive. (Familiar examples for $O_{2d+1}(\mathbb{F}_2) = \mathrm{Sp}_{2d}(\mathbb{F}_2)$ come from $d = 2$ and $d = 3$. Here the groups are S_6 and $W(E_7)$, respectively. The orbit sizes on (X_{-1}, X_1, X_0) are $(6, 10, 15)$ and $(28, 36, 63)$ respectively.)

By discarding a finite number of Γ, we can assume $d \geq j$. For $\epsilon \in \{\pm 1\}$, fix $x_1, \ldots, x_j$ in X_ϵ spanning a j-dimensional vector space $V \subset \mathbb{F}_2^{2d+1}$ on which the quadratic form remains nondegenerate and with each x_i having type ϵ in this smaller space. Let $V_\epsilon = V \cap X_\epsilon$. Writing $j = 2u + 1$, one has $|V_\epsilon| = 2^{u-1}(2^u + \epsilon)$. There are $(|V_\epsilon| - j)^{\underline{j}}$ ways to successively choose $x_{j+1}, \ldots, x_{2j}$ in V_ϵ so that all the x_i are distinct. One has

$$|X_\epsilon^{\underline{2j}}/O_{2d+1}(\mathbb{F}_2)| \geq (2^{u-1}(2^u + \epsilon) - j)^{\underline{j}} \geq (2^{u-1}(2^u + \epsilon) - 2j + 1)^j > 2^{j^2-4j}.$$

Thus there are no further exceptional Γ from this case. $\square$

6E. ***End of the proof of I $\Longrightarrow$ II in general.*** We now end the proof without the split-cyclicity assumption, by modifying the standard argument of Section 6C.

Consider again the diagram (5-2) relating two five-term exact sequences. The last three terms of the top sequence and the last four terms of the bottom sequence give respectively

$$|H_2(G^{[k]})| \leq |H_2(G')_{G^{\mathrm{ab}}}|^k |H_2(G^{\mathrm{ab}})|,$$

$$|H_2(G')_{G^{\mathrm{ab}}}|^k \leq \frac{|H_3(G^{\mathrm{ab}})|^k |H_2(G)|^k}{|H_2(G^{\mathrm{ab}})|^k}.$$

Combining these inequalities and replacing $H_2(G^{[k]})$ by its quotient $H_2(G^{[k]}, C^k)$ yields

$$|H_2(G^{[k]}, C^k)| \leq |H_2(G) \times H_3(G^{\mathrm{ab}})|^k. \tag{6-7}$$

As described in Section 6C, Condition 3 implies that for $\min \nu_i$ sufficiently large, the action of Br_ν on $\mathcal{F}_h^*$ is 2-transitive when restricted to each orbit. We will use this 2-transitivity and the exponential bound (6-7) to conclude that the action of Br_ν on $\mathcal{F}_h^*$ is asymptotically quasifull.

Consider S_m in its standard full action on $Y_m = \{1, \ldots, m\}$. The induced action on $X_m = Y_m \sqcup Y_m$ is not quasifull. Let $a_{k,m}$ be the number of orbits of S_m on Y_m^k. As m increases, the sequence $a_{k,m}$ stabilizes at a number a_k. The sequence a_k appears in [Sloane 1991] as A000898. There are several explicit formulas and combinatorial interpretations. The only important thing for us is that a_k grows superexponentially, as indeed $a_k/a_{k-1} \sim \sqrt{2k}$.

From (6-7) we know that there exists an odd number j with

$$|H_2(G^{[2j]}, C^{2j}, \nu)/\operatorname{Out}(G, C)^{[2j]}| \leq |H_2(G^{[2j]}, C^{2j})| < \min(2^{j^2-4j}, a_{2j}).$$

By (6-4), the left-hand set is identified with $|\mathcal{F}_h^{*2j}/\mathrm{Br}_\nu|$ for sufficiently large $\min_i \nu_i$. Lemma 6.2 above says that, at the possible expense of making $\min_i \nu_i$ even larger, each orbit of the action of Br_ν on $\mathcal{F}_h^*$ is full. Our discussion of the action of S_m on Y_m says that the constituents are pairwise nonisomorphic, again for sufficiently large $\min_i \nu_i$. The classical Goursat lemma then says the action is quasifull. □

A consequence of the results of this section is that in fact the equivalence $3 \iff \hat{3}$ of Proposition 5.2 holds without the assumption of split-cyclicity. Condition E is also meaningful in general, and it would be interesting to identify the class of (G, C) for which the equivalence extends to include E.

7. Proof of II ⟹ I

In this section, we complete the proof of Theorem 5.1 by proving that (not I) implies (not II). Accordingly, we fix a centerless group G and a list $C = (C_1, \ldots, C_r)$ of conjugacy classes, and consider consequences of the failure of Conditions 1, 2, and 3 in turn. In all three cases, we show more than is needed for Theorem 5.1.

7A. ***Failure of Condition 1.*** The failure of the first condition requires a somewhat lengthy analysis, because it breaks into two quite different cases. The conclusion of the following lemma shows more than that asymptotic quasifullness of $\mathsf{Hur}_h^* \to \mathsf{Conf}_\nu$ fails; it shows that asymptotically each individual component $\mathsf{Hur}_{h,\ell}^* \to \mathsf{Conf}_\nu$ fails to be full.

Lemma 7.1. *Let G be a centerless group which is not pseudosimple. Let $C = (C_1, \dots, C_r)$ be a list of conjugacy classes. Consider varying allowed $\nu \in \mathbb{Z}^r_{\geq 1}$ and thus varying Hurwitz parameters $h = (G, C, \nu)$. Then for $\min_i \nu_i$ sufficiently large and any $\ell \in H^*_h$, the action of Br_ν on $\mathcal{F}^*_{h,\ell}$ is not full.*

Proof. A group is pseudosimple exactly when it satisfies two conditions: (A), it has no proper nonabelian quotients, or (B), its derived group is nonabelian. We assume first that (A) fails. Then we assume that (A) holds but (B) fails.

Assume (A) fails. Let $\overline{G}$ be a proper nonabelian quotient and $\bar{h} = (\overline{G}, (\overline{C}_1, \dots, \overline{C}_r), \nu)$ the corresponding quotient Hurwitz parameter. Consider the natural map $H_h \to H_{\bar{h}}$ from Section 4E, and let $\bar{\ell}$ be the image of ℓ.

By the definition of Hurwitz parameters, the classes C_i generate G. At least one of the surjections $C_i \to \overline{C}_i$ has to be noninjective, as otherwise the kernel of $G \to \overline{G}$ would be central in G and G is centerless. So $|C_i| \geq 2|\overline{C}_i|$ for at least one i. Similarly, since $\overline{G}$ is nonabelian and generated by the $\overline{C}_i$, one has $|\overline{C}_i| \geq 2$ for at least one i.

We now examine the induced map $\mathcal{G}_{h,\ell} \to \mathcal{G}_{\bar{h},\bar{\ell}}$. Let $\mathcal{I}_{h,\ell}$ be its image and $\phi_{h,\ell}$ the size of its largest fiber. We will use the two inequalities of the previous paragraph to show that both $\phi_{h,\ell}$ and $|\mathcal{I}_{h,\ell}|$ grow without bound with $\min_i \nu_i$.

From $|C_i| \geq 2|\overline{C}_i|$ and two applications of the asymptotic mass formula (3-7), one gets $|\mathcal{G}_{h,\ell}| \geq 1.5^{\min_i \nu_i} |\mathcal{G}_{\bar{h},\bar{\ell}}|$, and hence $\phi_{h,\ell} \geq 1.5^{\min_i \nu_i}$.

To show the growth of $|\mathcal{I}_{h,\ell}|$, we assume without loss of generality that $|\overline{C}_1| \geq 2$, and choose $y_1 \neq y_2 \in \overline{C}_1$. Let M be the exponent of a reduced Schur cover $\widetilde{G}_C$ of G. Let k be a positive integer and let $a_1, \dots, a_k$ be a sequence with $a_i \in \{1, 2\}$. Then for $\min_i \nu_i$ large enough, we claim that $\mathcal{I}_{h,\ell}$ contains an element of the form

$$(\underbrace{y_{a_1}, \dots, y_{a_1}}_{M}, \dots, \underbrace{y_{a_k}, \dots, y_{a_k}}_{M}, \underbrace{x_1, \dots, x_{\nu_1 - Mk}}_{\text{all in } \overline{C}_1}, \dots, \underbrace{x_{n-kM-\nu_r+1}, \dots, x_{n-kM}}_{\text{all in } \overline{C}_r}). \tag{7-1}$$

To see the existence of such an element, fix a lift C^*_i of the conjugacy class C_i to $\widetilde{G}_C$ and choose $\tilde{y}_1, \tilde{y}_2 \in C^*_1$ mapping (under $\widetilde{G}_C \to G \to \overline{G}$) to $y_1, y_2 \in \overline{C}_1$ respectively.

Let $z \in H_2(G, C)$ be chosen so that $z^{-1} \cdot \prod_i [C_i]^{\nu_i} = \ell$ inside $H_2(G, C, \nu)$. Consider the equation

$$(\tilde{y}^M_{a_1} \cdots \tilde{y}^M_{a_k}) \underbrace{\tilde{x}_1 \cdots \tilde{x}_{\nu_1 - kM}}_{\text{all in } C^*_1} \cdots \underbrace{\tilde{x}_{n-kM-\nu_r+1} \cdots \tilde{x}_{n-kM}}_{\text{all in } C^*_r} = z, \tag{7-2}$$

where $\tilde{x}_i \in C^*_i$. By our choice of M, the powers $\tilde{y}^M_{a_i}$ are all the identity in $\widetilde{G}_C$. One has $[C^*_1]^{\nu_1 - kM} \cdots [C^*_r]^{\nu_r} = [z]$ in $\widetilde{G}^{\mathrm{ab}}_C = G^{\mathrm{ab}}$, both sides being the identity. The asymptotic mass formula then applies to say that (7-2) in fact has many solutions $(\tilde{x}_1, \dots, \tilde{x}_{n-kM})$ where moreover the $\tilde{x}_i$ generate $\widetilde{G}_C$. Now, the image of $(\tilde{y}_{a_1}, \dots, \tilde{y}_{a_k}, \tilde{x}_1, \dots, \tilde{x}_{n-kM})$ actually defines an element of $\mathcal{G}_{h,\ell}$, and its image in $\overline{G}$ is an element of $\mathcal{I}_{h,\ell}$ of the form (7-1). Varying $(a_1, \dots, a_k)$ now, always taking $\min_i \nu_i$ sufficiently large, we conclude $|\mathcal{I}_{h,\ell}| \geq 2^k$.

For large enough $\min_i \nu_i$, the action of Br_ν on $\mathcal{G}_{h,\ell}$ is transitive, by the Conway–Parker theorem. This action preserves a partition of $\mathcal{G}_{h,\ell}$ into $b = |\mathcal{I}_{h,\ell}|$ blocks, each of size $f = \phi_{h,\ell}$. Thus the image of Br_ν on $\mathcal{G}_{h,\ell}$ is contained in the wreath product $S_f \wr S_b$. Hence the image of Br_ν on $\mathcal{F}^*_{h,\ell}$ is contained in a subquotient of $S_f \wr S_b$. But we have established that f and b increase indefinitely with $\min_i \nu_i$. Let $a = |\mathrm{Aut}(G,C)_\ell|$ and $m = |\mathcal{F}^*_{h,\ell}|$, so that $|\mathcal{G}_{h,\ell}| = ma = fb$. As soon as $\min(f,b) > a$, one has $m > \max(f,b)$ and the alternating group A_m is not a subquotient of $S_f \wr S_b$. So the action of Br_ν on $\mathcal{F}^*_{h,\ell}$ is not full.

Assume (A) holds but (B) fails. The assumptions force G' to be isomorphic to the additive group of $\mathbb{F}_p^w$ for some prime p and some power w. Moreover, consider the action of G^{ab} on G'. Now G', considered as an $\mathbb{F}_p$-vector space, is an irreducible representation of $\mathbb{F}_p[G^{\mathrm{ab}}]$. The order of G^{ab} must be coprime to p, as otherwise the fixed subspace for the p-primary part of G^{ab} would be a proper subrepresentation. So $\mathbb{F}_p[G^{\mathrm{ab}}]$ is isomorphic to a sum of finite fields and the action on $G' = \mathbb{F}_p^w$ is through a single summand $\mathbb{F}_q$. We can thus identify G' with the additive group of a finite field $\mathbb{F}_q$ and G^{ab} with a subgroup of $\mathbb{F}_q^\times$ in such a way that G itself is a subgroup of the affine group $\mathbb{F}_q.\mathbb{F}_q^\times$. Moreover, $G^{\mathrm{ab}} \subseteq \mathbb{F}_q^\times$ acts irreducibly on $\mathbb{F}_q$ as an $\mathbb{F}_p$-vector space.

We think of elements of G as affine transformations $x \mapsto mx + b$. Since braid groups act on the right in (3-2), we compose these affine transformation from left to right, so that the group law is

$$\begin{pmatrix} m_1 \\ b_1 \end{pmatrix}\begin{pmatrix} m_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} m_1 m_2 \\ m_2 b_1 + b_2 \end{pmatrix}.$$

We think of elements $(g_1, \dots, g_n) \in \mathcal{G}_h$ with $g_i = \binom{m_i}{b_i}$ in terms of the matrix

$$\begin{pmatrix} m_1 & \dots & m_i & m_{i+1} & \dots & m_n \\ b_1 & \dots & b_i & b_{i+1} & \dots & b_n \end{pmatrix}. \tag{7-3}$$

The top row is determined by C, via $m_i = [C_i]$. Thus, via the bottom row, we have realized $\mathcal{G}_h$ as a subset of $\mathbb{F}_q^n$. We can assume without loss of generality that none of the C_i are the identity class. Then the requirement $g_i \in C_i$ for membership in $\mathcal{G}_h$ gives $|G^{\mathrm{ab}}|$ choices for b_i if $m_i = 1$. If $m_i \neq 1$ then $g_i \in C_i$ allows all q choices for b_i.

Now briefly view $(g_1, \dots, g_n)$ as part of the larger catch-all set G^n of Section 3C, on which the standard braid operators σ_i act. The braiding rule (3-2) in our current setting becomes

$$\left(\dots,\binom{m_i}{b_i},\binom{m_{i+1}}{b_{i+1}},\dots\right)^{\sigma_i}=\left(\dots,\binom{m_{i+1}}{b_{i+1}},\binom{m_i}{b_{i+1}+m_{i+1}b_i-m_ib_{i+1}},\dots\right).$$

Thus the action of σ_i corresponds to the bottom row of (7-3), viewed as row vector of length n, being multiplied on the right by an n-by-n matrix in $GL_n(\mathbb{F}_q)$.

Returning now to the set $\mathcal{G}_h$ itself, any element of Br_ν can be written as a product of the σ_i and their inverses. Accordingly, image of Br_ν in $\mathrm{Sym}(\mathcal{G}_h)$ lies in $GL_n(\mathbb{F}_q)$.

To prove nonfullness, it suffices to bound the sizes of groups. On the one hand,

$$|\text{image of } \mathrm{Br}_\nu \text{ in } \mathrm{Sym}(\mathcal{F}^*_{h,\ell})| \le |\text{image of } \mathrm{Br}_\nu \text{ in } \mathrm{Sym}(\mathcal{G}_h)| \le |\,\mathrm{GL}_n(\mathbb{F}_q)| < q^{n^2}.$$

On the other hand, let $b = |H_2(G,C)||\mathrm{Out}(G,C)|+1$. Then, using (3-7), (4-6) and the fact that $|C_i| \in \{|G^{\mathrm{ab}}|, q\}$, one has

$$|\mathcal{F}^*_{h,\ell}| > \frac{\prod_i |C_i|^{\nu_i}}{|G||G'|b} \ge \frac{|G^{\mathrm{ab}}|^{n-3}}{q^2 b}$$

for all sufficiently large n. Certainly $q^{n^2} < \frac{1}{2}((a^{n-3})/(q^2b))!$ for any fixed a, b, $q>1$ and sufficiently large n. Thus the image of Br_ν in $\mathrm{Sym}(\mathcal{F}^*_{h,\ell})$ cannot contain $\mathrm{Alt}(\mathcal{F}^*_{h,\ell})$. □

The paper [Eisenbud et al. 1991] calculates monodromy in cases with $G=S_3$ and $G=S_4$, providing worked-out examples. Another illustration of the case with affine monodromy is [Malle and Matzat 1999, Proposition 10.4].

7B. ***Failure of Condition 2.*** Our next lemma has the same conclusion as the previous lemma:

Lemma 7.2. *Let G be a centerless group. Let $C=(C_1,\dots,C_r)$ be a list of conjugacy classes with at least one C_i ambiguous. Consider varying allowed $\nu\in\mathbb{Z}^r_{\ge 1}$ and thus varying Hurwitz parameters $h=(G,C,\nu)$. Then for* $\min_i \nu_i$ *sufficiently large and any $\ell\in H^*_h$, the action of* Br_ν *on $\mathcal{F}^*_{h,\ell}$ is not full.*

Proof. Introduce indexing sets B_i by writing

$$C_i = \bigsqcup\nolimits_{b\in B_i} C_{ib},$$

where each C_{ib} is a single G'-orbit. Our hypothesis says that at least one of the B_i — without loss of generality, B_1 — has size larger than 1. On the other hand, at least one of the B_i has size strictly less than C_i; otherwise G' would centralize each element of each C_i, and then all of G, which is impossible for G center-free.

Define

$$\mathcal{G}_h^{\mathrm{amb}} = \underbrace{B_1\times\cdots\times B_1}_{\nu_1}\times\cdots\times\underbrace{B_r\times\cdots\times B_r}_{\nu_r}.$$

The group G acts transitively through its abelianization G^{ab} on each B_i. For a lifting invariant $\ell \in H_h$, consider the natural map $\mathcal{G}_{h,\ell} \to \mathcal{G}_h^{\mathrm{amb}}$. The action of the braid group Br_ν on $\mathcal{G}_{h,\ell}$ descends to an action on $\mathcal{G}_h^{\mathrm{amb}}$.

Now we let $\min_i \nu_i \to \infty$ and get the following consequences, by arguments very closely paralleling those for the first case of Lemma 7.1. First, the image of the map $\mathcal{G}_{h,\ell} \to \mathcal{G}_h^{\mathrm{amb}}$ has size that goes to ∞. Second, the mass formula again shows that $|\mathcal{G}_{h,\ell}|/|\mathcal{G}_h^{\mathrm{amb}}| \to \infty$ with $\min_i \nu_i$. By the last paragraph of the first case of the proof of Lemma 7.1, the action of Br_ν on each orbit of $\mathcal{F}_{h,\ell}^*$ is forced to be imprimitive, and hence not full. $\square$

For a contrasting pair of examples, consider $h = (S_5, (C_{2111}, C_{311}, C_5), \nu)$ for $\nu = (2, 2, 1)$ and $\nu = (2, 1, 2)$. The monodromy group for the former is all of S_{125}, despite the presence of the ambiguous class C_5. The monodromy group for the latter is $S_{85} \wr S_2$ and represents the asymptotically forced nonfullness.

7C. *Failure of Condition 3.* The last lemma of this section is different in structure from the previous two, and its proof is essentially a collection of some of our previous arguments. From the discussion of surjectivity after (5-2), one always has

$$|H_2(G^{[2]}, C^2)| = a|H_2(G, C)|^2 \tag{7-4}$$

for some positive integer a. Condition 3 is that $a = 1$. The number a reappears as the cardinality of every fiber of the map of torsors

$$H_2(G^{[2]}, C^2, \nu) \overset{\pi}{\twoheadrightarrow} H_2(G, C, \nu)^2$$

considered in Section 4E.

Now suppose that ν is such that all ν_i are divisible by both the exponent of $H_2(G, C)$ and the exponent of $H_2(G^{[2]}, C^2)$. In that case, we have identifications

$$\begin{array}{ccc} H_2(G^{[2]}, C^2, \nu)_f & \xrightarrow{\ \pi\ } & H_2(G, C, \nu)_g^2 \\ \downarrow & & \downarrow \\ H_2(G^{[2]}, C^2) & \longrightarrow & H_2(G, C)^2 \end{array} \tag{7-5}$$

where the vertical bijections f, g come from Section 4C, and the fact that the diagram commutes is also explained there. The set $E := \pi^{-1} g^{-1}(0) \subseteq H_2(G^{[2]}, C^2, \nu)$ is a fiber of π. It has size ≥ 2 and $f(E) \subseteq H_2(G^{[2]}, C^2)$ is a subgroup.

The group $\mathrm{Out}(G, C)^{[2]}$, defined before (6-3), acts on $H_2(G^{[2]}, C^2, \nu)$ and also (compatibly) on $H_2(G^{[2]}, C^2)$. It preserves E and acts on it with at least two orbits, because it fixes the zero element of $f(E)$. Under the bijection (6-4), these two orbits correspond to two different braid orbits O, O' on $(\mathcal{F}_h^*)^{\underline{2}}$ which project (in both coordinates) to the same braid orbit on $\mathcal{F}_h^*$.

Summarizing, we have proved:

Lemma 7.3. *Let G be a pseudosimple group, let $C = (C_1, \dots, C_r)$ be a list of unambiguous conjugacy classes, and suppose $a > 1$ in (7-4). Consider ν with each ν_i a multiple of the exponent of both $H_2(G, C)$ and $H_2(G^{[2]}, C^2)$ so that $H_h^* = H_2(G, C, \nu)$ contains a trivial lifting invariant 0 via f from (7-5). Then for $\min_i \nu_i$ sufficiently large, the action of Br_ν on $\mathcal{F}_{h,0}^*$ is not 2-transitive and hence not full.*

8. Full number fields

The full-monodromy theorem gives us confidence in the following conjecture:

Conjecture 8.1. Suppose $\mathcal{P}$ contains the set of prime divisors of the order of a nonabelian finite simple group. Then there exist infinitely many full fields unramified outside $\mathcal{P}$.

In this final section we briefly discuss this conjecture. In particular we give a heuristic justification based on our results here. The sequel paper [Roberts $\geq$ 2015] will present a more comprehensive treatment.

8A. ***Specialization to number fields.*** Let $h = (G, C, \nu)$ be a Hurwitz parameter with each C_i rational for simplicity. Then one has (see Section 2D) the cover $\pi : \mathrm{HUR}_h^* \to \mathrm{CONF}_\nu$ of $\mathbb{Q}$-varieties. For every $u \in \mathrm{CONF}_\nu(\mathbb{Q})$, the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on the fiber $\pi^{-1}(u) \subset \mathrm{HUR}_h^*(\overline{\mathbb{Q}})$. Let $K_{h,u}^*$ be the corresponding $\mathbb{Q}$-algebra, so that $K_{h,u}^*$ factors into fields indexed by the orbits of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $\pi^{-1}(u)$.

Controlling Galois groups. The Hilbert irreducibility theorem says that the Galois groups associated to $K_{h,u}^*$ coincide with the generic Galois group of the cover for u outside a thin set. Thus, to take the example most relevant for us, if the cover is full then one has many full specializations $K_{h,u}^*$.

Controlling ramification. For $\mathcal{P}$ a set of primes, let $\mathbb{Z}[1/\mathcal{P}] = \mathbb{Z}[\{1/p\}_{p \in \mathcal{P}}]$. The variety CONF_ν comes from a scheme over $\mathbb{Z}$ and so the set of $\mathcal{P}$-integral points $\mathrm{CONF}_\nu(\mathbb{Z}[1/\mathcal{P}])$ is defined. Suppose now that $\mathcal{P}$ contains all primes dividing $|G|$. Then the cover $\mathrm{HUR}_h^* \to \mathrm{CONF}_\nu$ has bad reduction within $\mathcal{P}$. The theory of algebraic fundamental groups then implies that $K_{h,u}^*$ is ramified within $\mathcal{P}$ whenever $u \in \mathrm{CONF}_\nu(\mathbb{Z}[1/\mathcal{P}])$.

8B. ***Heuristic argument for Conjecture 8.1.*** Let $\mathcal{P}$ be as in the statement of the conjecture. Let G be a simple group with all primes dividing $|G|$ in $\mathcal{P}$. Let $C_1 \subset G$ be a class of involutions. Then $(G, (C_1))$ satisfies Statement I of the full-monodromy theorem. By Statement II, there are infinitely many Hurwitz parameters $h = (G, (C_1), (\nu_1))$ such that $\mathsf{Hur}^*_h \to \mathsf{Conf}_\nu$ has quasifull monodromy with connected components indexed by the finite set H^*_h.

There is a natural action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})^{\mathrm{ab}}$ on H^*_h. When ν_1 is divisible by the exponent of $H_{G,(C_1)}$, there is an identity element $0 \in H^*_h$ fixed by $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})^{\mathrm{ab}}$. Thus one gets infinitely many full covers $\mathrm{HUR}^*_{h,0} \to \mathrm{CONF}_\nu$ defined over $\mathbb{Q}$ and ramified within $\mathcal{P}$. It is proved in [Roberts 2014, §7] that the number $N(\nu_1)$ of $\mathrm{PGL}_2(\mathbb{Q})$-orbits represented by points in $\mathrm{CONF}_{(\nu_1)}(\mathbb{Z}[1/\mathcal{P}])$ tends to ∞ with ν_1.

Thus, for each ν_1 in an infinite arithmetic progression, one has $N(\nu_1)$ algebras $K^*_{h,u}$ ramified within $\mathcal{P}$. To prove Conjecture 8.1, one not only has to control Galois groups and ramification as in Section 8A, one has to control them simultaneously, a difficult task. However if the thin set from Hilbert irreducibility intersects the $\mathcal{P}$-integral points at random, the $K^*_{h,u}$ should have a strong tendency to be full fields. On similar grounds, one would expect the $K^*_{h,u}$ to be nonisomorphic. Direct calculations, like those summarized in the next two subsections, confirm these expectations very strongly. For Conjecture 8.1 to hold for $\mathcal{P}$, there would just have to be a subsequence of ν_1 for which one of the $N(\nu_1)$ subalgebras was full.

8C. ***Specializing a sample cover.*** To illustrate concretely how Hurwitz covers naturally lead to full number fields with controlled ramification, we summarize here the introductory example of [Roberts $\geq$ 2015]. In this example, let $h = (S_5, (C_{2111}, C_5), (4, 1))$, with C_{2111} and C_5 the class of involutions and 5-cycles respectively. Then $\mathsf{Hur}^*_h = \mathsf{Hur}_h$ is a full cover of $\mathsf{Conf}_{4,1}$ of degree 25. The fiber of $\mathsf{Hur}_h \to \mathsf{Conf}_{4,1}$ over the configuration $u = (D_1, D_2) = (\{a_1, a_2, a_3, a_4\}, \{\infty\})$ consists of all equivalence classes of quintic polynomials

$$g(y) = y^5 + by^3 + cy^2 + dy + e \tag{8-1}$$

whose critical values are a_1, a_2, a_3, a_4. Here the equivalence class of $g(y)$ consists of the five polynomials $g(\zeta y)$, where ζ runs over fifth roots of unity.

Explicitly, consider the resultant $r(t)$ of $g(y) - t$ and $g'(y)$. Then $r(t)$ equals

$$\begin{aligned}
&3125t^4 + 1250(3bc - 10e)t^3\\
&+(108b^5 - 900b^3d + 825b^2c^2 - 11250bce + 2000bd^2 + 2250c^2d + 18750e^2)t^2\\
&-2\big(108b^5e - 36b^4cd + 8b^3c^3 - 900b^3de + 825b^2c^2e + 280b^2cd^2 - 315bc^3d\\
&\qquad -5625bce^2 + 2000bd^2e + 54c^5 + 2250c^2de - 800cd^3 + 6250e^3\big)t\\
&+\big(108b^5e^2 - 72b^4cde + 16b^4d^3 + 16b^3c^3e - 4b^3c^2d^2 - 900b^3de^2 + 825b^2c^2e^2\\
&\qquad +560b^2cd^2e - 128b^2d^4 - 630bc^3de + 144bc^2d^3 - 3750bce^3 + 2000bd^2e^2\\
&\qquad +108c^5e - 27c^4d^2 + 2250c^2de^2 - 1600cd^3e + 256d^5 + 3125e^4\big).
\end{aligned}$$

For fixed $\{a_1, a_2, a_3, a_4\}$, there are generically 125 different solutions (b, c, d, e) to the equation $r(t) = 3125(t - a_1) \cdots (t - a_4)$. Two solutions are equivalent exactly if they have the same e. Whenever D_1 is rational, i.e., $\prod(t-a_i) \in \mathbb{Q}[t]$, the resulting set of e forms the set of roots of a degree-25 polynomial with rational coefficients. By taking $u \in \mathrm{CONF}_{4,1}(\mathbb{Z}[1/30])$, one gets more than 10000 different fields with Galois group A_{25} or S_{25} and discriminant of the form $\pm 2^a 3^b 5^c$.

8D. *Comparison with the mass heuristic.* Let $F_{\mathcal{P}}(m)$ be the number of full fields ramified within $\mathcal{P}$ of degree m. The mass heuristic [Bhargava 2007, (3.3)] gives an expected value $\mu_{\mathcal{P}}(m)$ for $F_{\mathcal{P}}(m)$ as an easily computed product of local masses. This heuristic has had clear success in the setting of fixed degree and large discriminant, being for example exactly right on average for $m = 5$ [Bhargava 2010]. In [Roberts 2007, §11], we considered the opposite regime of fixed ramifying primes and increasing degree. We proved there that for any $\mathcal{P}$ the sequence $\mu_{\mathcal{P}}(m)$ ultimately decreases superexponentially with m.

The convergence of $\sum_{m=1}^{\infty} \mu_{\mathcal{P}}(m)$ for any $\mathcal{P}$ argues against Conjecture 8.1. However, our calculations confirming genericity of specialization make it clear that the $K_{h,u}^*$ we are considering here simply escape the influence of the mass heuristic. For instance, one of many examples in [Roberts $\geq$ 2015] comes from the Hurwitz parameter $h = (S_6, (C_{21111}, C_{321}, C_{3111}, C_{411}), (2, 1, 1, 1))$. The covering $\mathrm{HUR}_h \to \mathrm{CONF}_{2,1,1,1}$ is full of degree 202. The specialization set $\mathrm{CONF}_{2,1,1,1}(\mathbb{Z}[1/30])$ intersects exactly 2947 different $\mathrm{PGL}_2(\mathbb{Q})$-orbits on the set $\mathrm{CONF}_{2,1,1,1}(\mathbb{Q})$ [Roberts 2014, §9.2]. The mass heuristic predicts

$$\sum_{m=202}^{\infty} \mu_{\{2,3,5\}}(m) < 10^{-16}$$

full fields in degree ≥ 202. However specialization is as generic as it could be, as the 2947 algebras $K_{h,u}$ are pairwise nonisomorphic and all full.

8E. *Concluding discussion.* There are other aspects of the sequences $F_{\mathcal{P}}(m)$ that are not addressed by Conjecture 8.1. For example, our belief is that Conjecture 8.1 still holds with the conclusion strengthened to $F_{\mathcal{P}}(m)$ being unbounded. Also notable is that fields arising from full fibers of Hurwitz covers occur only in degrees for which there is a cover. By the mass formula, these degrees form a sequence of density zero. A fundamental question is thus the support of the sequences $F_{\mathcal{P}}(m)$, meaning the set of degrees m for which $F_{\mathcal{P}}(m)$ is positive.

One extreme possibility, giving as much credence to the mass heuristic as is still reasonable, is that $F_{\mathcal{P}}(m)$ has support on a sequence of density zero in general and is eventually zero unless $\mathcal{P}$ contains the set of prime divisors of the order of a finite simple group. This would imply that the classification of finite simple groups has an unexpected governing influence on a part of algebraic number theory seemingly quite removed from general group theory. If this extreme possibility does not hold, then there would have to be a broad and as yet unknown class of number fields which is also exceptional from the point of view of the mass heuristic.

Acknowledgements

We thank Simon Rubinstein-Salzedo and John Voight for helpful comments on earlier drafts of this paper. We also thank two anonymous referees for their helpful corrections and suggestions.

Roberts was supported by grant #209472 from the Simons Foundation. Venkatesh was supported by grants from the NSF and from the Packard Foundation.

References

[Bailey and Fried 2002] P. Bailey and M. D. Fried, "Hurwitz monodromy, spin separation and higher levels of a modular tower", pp. 79–220 in *Arithmetic fundamental groups and noncommutative algebra* (Berkeley, CA, 1999), edited by M. D. Fried and Y. Ihara, Proc. Sympos. Pure Math. **70**, Amer. Math. Soc., Providence, RI, 2002. MR 2005b:14044 Zbl 1072.14026

[Bertin and Romagny 2011] J. Bertin and M. Romagny, *Champs de Hurwitz*, Mém. Soc. Math. Fr. (N.S.) **125–126**, 2011. MR 2920693 Zbl 1242.14025

[Bhargava 2007] M. Bhargava, "Mass formulae for extensions of local fields, and conjectures on the density of number field discriminants", *Int. Math. Res. Not.* **2007**:17 (2007), Art. ID rnm052. MR 2009e:11220 Zbl 1145.11080

[Bhargava 2010] M. Bhargava, "The density of discriminants of quintic rings and fields", *Ann. of Math.* (2) **172**:3 (2010), 1559–1591. MR 2011k:11152 Zbl 1220.11139

[Conway and Parker 1988] J. H. Conway and R. A. Parker, "On the Hurwitz number of arrays of group elements", unpublished preprint, 1988.

[Conway et al. 1985] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups*, Oxford University Press, Eynsham, 1985. MR 88g:20025 Zbl 0568.20001

[Dixon and Mortimer 1996] J. D. Dixon and B. Mortimer, *Permutation groups*, Graduate Texts in Mathematics **163**, Springer, New York, 1996. MR 98m:20003 Zbl 0951.20001

[Dunfield and Thurston 2006] N. M. Dunfield and W. P. Thurston, "Finite covers of random 3-manifolds", *Invent. Math.* **166**:3 (2006), 457–521. MR 2007f:57039 Zbl 1111.57013

[Eckmann and Stammbach 1970] B. Eckmann and U. Stammbach, "On exact sequences in the homology of groups and algebras", *Illinois J. Math.* **14** (1970), 205–215. MR 42 #4615 Zbl 0195.03402

[Eisenbud et al. 1991] D. Eisenbud, N. Elkies, J. Harris, and R. Speiser, "On the Hurwitz scheme and its monodromy", *Compositio Math.* **77**:1 (1991), 95–117. MR 92c:14019 Zbl 0726.14022

[Ellenberg et al. 2013] J. Ellenberg, A. Venkatesh, and C. Westerland, "Homological stability for Hurwitz spaces and the Cohen–Lenstra conjecture over function fields, II", preprint, 2013. arXiv 1212.0923

[Fried and Völklein 1991] M. D. Fried and H. Völklein, "The inverse Galois problem and rational points on moduli spaces", *Math. Ann.* **290**:4 (1991), 771–800. MR 93a:12004 Zbl 0763.12004

[Kluitmann 1988] P. Kluitmann, "Hurwitz action and finite quotients of braid groups", pp. 299–325 in *Braids* (Santa Cruz, CA, 1986), edited by J. S. Birman and A. Libgober, Contemp. Math. **78**, Amer. Math. Soc., Providence, RI, 1988. MR 90d:20071 Zbl 0701.20019

[Magaard et al. 2003] K. Magaard, S. Shpectorov, and H. Völklein, "A GAP package for braid orbit computation and applications", *Experiment. Math.* **12**:4 (2003), 385–393. MR 2005e:12007 Zbl 1068.12002

[Malle and Matzat 1999] G. Malle and B. H. Matzat, *Inverse Galois theory*, Springer, Berlin, 1999. MR 2000k:12004 Zbl 0940.12001

[Malle and Roberts 2005] G. Malle and D. P. Roberts, "Number fields with discriminant $\pm 2^a 3^b$ and Galois group A_n or S_n", *LMS J. Comput. Math.* **8** (2005), 80–101. MR 2006a:11137 Zbl 1119.11064

[Roberts 2007] D. P. Roberts, "Wild partitions and number theory", *J. Integer Seq.* **10**:6 (2007), Article 07.6.6, 34. MR 2009b:11206 Zbl 1174.11094

[Roberts 2014] D. P. Roberts, "Polynomials with prescribed bad primes", *Int. J. Number Theory* (online publication December 2014).

[Roberts ≥ 2015] D. P. Roberts, "Hurwitz number fields", in preparation.

[Romagny and Wewers 2006] M. Romagny and S. Wewers, "Hurwitz spaces", pp. 313–341 in *Groupes de Galois arithmétiques et différentiels*, edited by D. Bertrand and P. Dèbes, Sémin. Congr. **13**, Soc. Math. France, Paris, 2006. MR 2008e:14040 Zbl 1156.14314

[Serre 1990] J.-P. Serre, "Relèvements dans $\tilde{\mathfrak{A}}_n$", *C. R. Acad. Sci. Paris Sér. I Math.* **311**:8 (1990), 477–482. MR 91m:20010 Zbl 0714.20003

[Serre 2008] J.-P. Serre, *Topics in Galois theory*, 2nd ed., Research Notes in Mathematics **1**, A K Peters, Wellesley, MA, 2008. MR 2008i:12010 Zbl 1128.12001

[Sloane 1991] N. Sloane, "$a(n) = 2(a(n-1) + (n-1)a(n-2))$, $a(0) = 1$", sequence A000898 in *The online encyclopedia of integer sequences* (oeis.org), 1991.

Communicated by Barry Mazur
Received 2014-01-28 Revised 2015-01-08 Accepted 2015-02-18

roberts@morris.umn.edu *Division of Science and Mathematics, University of Minnesota, Morris, MN 56267, United States*

akshay@math.stanford.edu *Department of Mathematics, Building 380, Stanford University, Stanford, CA 94305, United States*

ALGEBRA AND NUMBER THEORY 9:3 (2015)
dx.doi.org/10.2140/ant.2015.9.547

The characteristic polynomial of the Adams operators on graded connected Hopf algebras

Marcelo Aguiar and Aaron Lauve

Dedicated to the memory of Jean-Louis Loday.

The Adams operators Ψ_n on a Hopf algebra H are the convolution powers of the identity of H. They are also called Hopf powers or Sweedler powers. We study the Adams operators when H is graded connected. The main result is a complete description of the characteristic polynomial — both eigenvalues and their multiplicities — for the action of the operator Ψ_n on each homogeneous component of H. The eigenvalues are powers of n. The multiplicities are independent of n, and in fact only depend on the dimension sequence of H. These results apply in particular to the antipode of H, as the case $n = -1$. We obtain closed forms for the generating function of the sequence of traces of the Adams operators. In the case of the antipode, the generating function bears a particularly simple relationship to the one for the dimension sequence. In the case where H is cofree, we give an alternative description for the characteristic polynomial and the trace of the antipode in terms of certain palindromic words. We discuss parallel results that hold for Hopf monoids in species and for q-Hopf algebras.

Introduction 548
1. Preliminaries on Hopf algebras 549
2. Characteristic polynomials of the Adams operators 554
3. The trace of the Adams operators 559
4. The case of cofree graded connected Hopf algebras 564
5. Examples 570
Appendix A. Hopf monoids in species 573
Appendix B. q-Hopf algebras 576
References 580

Aguiar was supported in part by NSF DMS-1001935. Lauve was supported in part by NSA #H98230-12-1-0286. This work was started at Texas A&M University and both authors are thankful to this institution.

MSC2010: primary 16T05; secondary 16T30.

Keywords: Adams operator, characteristic operation, convolution power, Hopf power, antipode, trace, graded connected Hopf algebra, Hopf monoid in species, q-Hopf algebra, Schur indicator, Eulerian idempotent.

Introduction

Let H be a Hopf algebra with antipode $\mathrm{S} : H \to H$. If H is commutative or cocommutative, then it is well known that S is an involution ($\mathrm{S}^2 = \mathrm{id}$). In particular, its eigenvalues are ± 1. Alternatively, if H is finite-dimensional, then S has finite even order (and its eigenvalues can be arbitrary even roots of unity). This paper studies the behavior of the antipode when H is graded connected. We prove in Corollary 5 that, in this case, the eigenvalues of S are always ± 1, even though S may have infinite order on any homogeneous component of H.

It is both natural and convenient to consider a more general family of operators on H: the convolution powers of the identity. These are the Adams operators Ψ_n, and the antipode is Ψ_{-1}. Our main result, Theorem 3, provides a complete description of the characteristic polynomial for the operator Ψ_n acting on the m-th homogeneous component of H. For each scalar n and nonnegative integer m, the polynomial is uniquely determined by the dimension sequence of H. The eigenvalues are powers of n. The multiplicities count samples with replacement by length and weight, with the samples taken from a weighted set that arises, numerically, as the inverse Euler transform of the dimension sequence of H, and algebraically, as a basis of the graded Lie algebra of primitive elements of a Hopf algebra canonically associated to H.

Corollaries 10 and 13 provide information on the trace of the antipode. The former provides a semicombinatorial description for its values, and the latter the following remarkable expression for the generating function:

$$\sum_{m \geq 0} \mathrm{trace}(\mathrm{S}|_{H_m})\, t^m = \frac{h(t^2)}{h(t)},$$

where $h(t)$ is the generating function for the dimension sequence of H. We derive in Corollary 17 information about the asymptotic behavior of the sequence of traces.

To put these results into perspective, consider the Hopf algebra of symmetric functions. Calculating the trace of the antipode on suitable linear bases yields a simple proof of the interesting identity (see Section 5.1)

$$c(m) = e(m) - o(m),$$

where $c(m)$ denotes the number of self-conjugate partitions of m, and $e(m)$ and $o(m)$ denote the number of partitions of m with an even number of even parts, and with an odd number of even parts, respectively. The generating function for the trace of the antipode in the preceding paragraph may be seen as an extension of this result to arbitrary graded connected Hopf algebras.

Consider now the Hopf algebra of quasisymmetric functions. Another quick calculation reveals that the trace of the antipode on the m-th component is a signed count of palindromic compositions of m; see Section 5.3. Corollary 22 provides a similar result that applies to any graded connected Hopf algebra that, as a graded coalgebra, is cofree.

Some of the main results admit extensions to certain graded connected q-Hopf algebras. Among these, we highlight Corollary 34, which describes the trace of the antipode as a polynomial in q, and Corollary 37, which involves appropriate q-generating functions and replaces the above relationship between the sequence of antipode traces and the dimension sequence.

The theory of Hopf monoids in species often runs in parallel to the theory of graded connected Hopf algebras. The main results of the paper admit variants that apply in this context. The sequence of antipode traces is now determined from the dimension sequence as follows (Corollary 29): the exponential generating function of the former is the reciprocal of that of the latter.

The paper is organized as follows. In Section 1, we discuss the necessary preliminaries from Hopf algebra theory. The proof of Theorem 3 is carried out in Section 2. Section 3 focuses on the trace of the Adams operators and the antipode particularly. In Section 4, we give alternatives to our main results about the antipode that hold in the presence of cofreeness assumptions. Section 5 provides illustrations of the results and some simple calculations. In Appendix A we present the results for Hopf monoids in species, and in Appendix B we treat the case of q-Hopf algebras.

The present paper supersedes and considerably expands on the results of our extended abstract [Aguiar and Lauve 2013].

1. Preliminaries on Hopf algebras

Throughout, all vector spaces are over a field $\Bbbk$ of characteristic zero.

The structure maps of a bialgebra H are denoted by

$$\mu : H \otimes H \to H, \quad \Delta : H \to H \otimes H,$$
$$\iota : \Bbbk \to H, \quad \varepsilon : H \to \Bbbk.$$

The *antipode* of a Hopf algebra H is denoted by

$$\mathrm{S} : H \to H.$$

1.1. ***Convolution.*** Let H be a bialgebra and let $\mathrm{End}(H)$ denote the space of linear maps $T : H \to H$. The *convolution product* of $P, Q \in \mathrm{End}(H)$ is

$$P * Q := \mu \circ (P \otimes Q) \circ \Delta.$$

This turns the space $\mathrm{End}(H)$ into an associative algebra. The unit element is

$$\iota \circ \varepsilon.$$

The bialgebra H is a *Hopf algebra* if and only if the identity of H is convolution-invertible. In this case, the antipode S is the convolution-inverse of the identity map:

$$\mathrm{S} * \mathrm{id} = \mathrm{id} * \mathrm{S} = \iota \circ \varepsilon.$$

Let H be a bialgebra. Put $\Delta^{(0)} := \mathrm{id}$, $\Delta^{(1)} := \Delta$, and

$$\Delta^{(n)} := (\Delta \otimes \mathrm{id}^{\otimes(n-1)}) \circ \Delta^{(n-1)} \quad \text{for all } n \geq 2.$$

The superscript is one less than the number of tensor factors in the codomain. Similarly, $\mu^{(n)}$ denotes the map that multiplies $n+1$ elements of H, with $\mu^{(0)} := \mathrm{id}$. The convolution powers of any $T \in \mathrm{End}(H)$ can be written as

$$T^{*0} = \iota \circ \varepsilon \qquad \text{and} \qquad T^{*n} = \mu^{(n-1)} \circ T^{\otimes n} \circ \Delta^{(n-1)} \quad \text{for } n \geq 1.$$

1.2. ***Adams operators.*** Let H be a Hopf algebra. The convolution powers id^{*n} of the identity of H are defined for any integer n. They are called *Adams operators* and are denoted by

$$\Psi_n := \mathrm{id}^{*n} : H \to H. \tag{1}$$

For $n \geq 1$, we have

$$\Psi_n = \mu^{(n-1)} \circ \Delta^{(n-1)}. \tag{2}$$

Note that $\Psi_0 = \iota \circ \varepsilon$ and $\Psi_{-1} = \mathrm{S}$. Also,

$$\Psi_{-n} = \mathrm{S}^{*n} \tag{3}$$

for all n.

This terminology is used in [Cartier 2007, §3.8] and [Loday 1992, §4.5]. Other common terminology for these operators are *Hopf powers* [Ng and Schauenburg 2008], *Sweedler powers* [Kashina et al. 2006; 2012], and *characteristic operations* [Gerstenhaber and Schack 1991; Patras 1993]. The paper [Aguiar and Mahajan 2013, §13] studies analogous operators in the context of Hopf monoids in species.

The main goal of this paper is to analyze the characteristic polynomial of these operators when the Hopf algebra H is graded connected.

1.3. ***Coradical filtration and primitive elements.*** For more details on the notions reviewed in this section, see [Montgomery 1993, Chapter 5] or [Radford 2012, Chapter 4].

Let $H^{(0)}$ denote the *coradical* of a bialgebra H, and let

$$H^{(0)} \subseteq H^{(1)} \subseteq H^{(2)} \subseteq \cdots \subseteq H$$

denote its coradical filtration. We have

$$H = \bigcup_{m \geq 0} H^{(m)}.$$

We say that H is *connected* if $H^{(0)}$ is spanned by the unit element $1 \in H$. In this case, H is a Hopf algebra; see Section 1.4. In addition, $H^{(1)} = H^{(0)} \oplus \mathcal{P}(H)$, where

$$\mathcal{P}(H) := \{x \in H \mid \Delta(x) = 1 \otimes x + x \otimes 1\}$$

is the space of *primitive elements* of H. More generally, setting $H_+ := \mathrm{Ker}(\varepsilon)$ and defining $\Delta_+ : H_+ \to H_+ \otimes H_+$ by

$$\Delta_+(x) := \Delta(x) - 1 \otimes x - x \otimes 1,$$

we have $H^{(m)} = H^{(0)} \oplus \mathrm{Ker}(\Delta_+^{(m)})$, where the iterates of Δ_+ are defined as for Δ.

Let H be a bialgebra. Let $\mathrm{gr}\, H$ denote the graded vector space associated to the coradical filtration of H:

$$\mathrm{gr}\, H := H^{(0)} \oplus (H^{(1)}/H^{(0)}) \oplus (H^{(2)}/H^{(1)}) \oplus (H^{(3)}/H^{(2)}) \oplus \cdots . \tag{4}$$

A filtration-preserving map $f : H \to K$ induces a map

$$\mathrm{gr}\, f : \mathrm{gr}\, H \to \mathrm{gr}\, K.$$

The structure maps of H are filtration-preserving (when $H \otimes H$ is endowed with the tensor product of the coradical filtrations of each factor). The induced maps turn $\mathrm{gr}\, H$ into a bialgebra. If H is a Hopf algebra, so is $\mathrm{gr}\, H$. If H is connected, so is $\mathrm{gr}\, H$. More importantly:

Lemma 1. *If the bialgebra H is connected, then the bialgebra* $\mathrm{gr}\, H$ *is commutative.*

This is an immediate consequence of [Sweedler 1969, Theorem 11.2.5.a]. A direct proof is given in [Aguiar and Sottile 2005a, Proposition 1.6]. We are thankful to Akira Masuoka, Susan Montgomery and the referee for pointing out the reference to Sweedler's text.

The passage $H \mapsto \mathrm{gr}\, H$ is functorial with respect to filtration-preserving maps. It follows that convolution products are preserved: if f and $g : H \to H$ are linear maps, then

$$\mathrm{gr}(f * g) = (\mathrm{gr}\, f) * (\mathrm{gr}\, g). \tag{5}$$

A morphism of bialgebras $f : H \to K$ preserves the coradical filtrations. The induced map $\mathrm{gr}\, f : \mathrm{gr}\, H \to \mathrm{gr}\, K$ is a morphism of bialgebras.

1.4. ***Antipode and Eulerian idempotents.*** Any connected bialgebra is a Hopf algebra with antipode

$$\mathrm{S} = \sum_{k \geq 0} (\iota \circ \varepsilon - \mathrm{id})^{*k}. \tag{6}$$

This basic result can be traced back to Sweedler [1969, Lemma 9.2.3] and Takeuchi [1971, Lemma 14]; see also [Montgomery 1993, Lemma 5.2.10] and [Radford

2012, Lemma 7.6.2]. It follows by expanding

$$x^{-1} = \frac{1}{1-(1-x)} = \sum_{k\geq 0}(1-x)^k$$

in the convolution algebra, with $x = \mathrm{id}$ and $1 = \iota\circ\varepsilon$. Connectedness guarantees that the sum in (6) is finite when evaluated on any $h \in H$. More precisely, if $h \in H^{(m)}$, then $(\mathrm{id} - \iota\circ\varepsilon)^{*k}(h) = 0$ for all $k > m$.

Assume for the remainder of the section that H is a connected Hopf algebra.

The binomial theorem yields the following expression for the n-th Adams operator for all integers n:

$$\Psi_n = \sum_{k\geq 0}\binom{n}{k}(\mathrm{id} - \iota\circ\varepsilon)^{*k}. \tag{7}$$

Moreover, the right-hand side of (7) is well-defined for all scalar values of n. In this manner, the Adams operators Ψ_n are defined for all scalars n.

Similarly, the following series expansion defines an element $\log(\mathrm{id})$ in the convolution algebra:

$$\log(\mathrm{id}) := -\sum_{k\geq 1}\frac{1}{k}(\iota\circ\varepsilon - \mathrm{id})^{*k}. \tag{8}$$

Additionally, consider the elements $\mathsf{e}^{(k)}$, for $k \geq 0$, given by

$$\mathsf{e}^{(0)} := \iota\circ\varepsilon, \qquad \mathsf{e}^{(1)} := \log(\mathrm{id}), \qquad \mathsf{e}^{(k)} := \frac{1}{k!}(\mathsf{e}^{(1)})^{*k} \quad \text{for } k > 1. \tag{9}$$

In the case where H is commutative or cocommutative, the $\mathsf{e}^{(k)}$ form a complete orthogonal system of idempotent operators on H. That is,

$$\mathrm{id} = \sum_{k\geq 0}\mathsf{e}^{(k)}, \qquad \mathsf{e}^{(k)}\circ\mathsf{e}^{(k)} = \mathsf{e}^{(k)}, \qquad \mathsf{e}^{(j)}\circ\mathsf{e}^{(k)} = 0 \quad \text{for } j \neq k. \tag{10}$$

The $\mathsf{e}^{(k)}$ are the *higher Eulerian idempotents*; $\mathsf{e}^{(1)}$ is the *first* Eulerian idempotent. It follows from (1), (9), and the identity $x^{*n} = \exp(n\log x)$ that

$$\Psi_n = \sum_{k\geq 0} n^k\,\mathsf{e}^{(k)} \tag{11}$$

for all scalars n. In particular,

$$\mathrm{S} = \sum_{k\geq 0}(-1)^k\,\mathsf{e}^{(k)}.$$

If H is cocommutative, $\mathsf{e}^{(k)}$ projects onto the subspace spanned by k-fold products of primitive elements of H. In particular, $\mathsf{e}^{(1)}$ projects onto $\mathcal{P}(H)$.

For proofs of these results, see [Loday 1992, Chapter 4], [Patras 1993] or [Schmitt 1994, §9]. Some instances of these operators in the recent literature include [Diaconis et al. 2014; Novelli et al. 2013; Pang 2014; Patras and Schocker 2006]. For references to earlier work on Eulerian idempotents, see [Aguiar and Mahajan 2013, §14].

1.5. ***Hopf–Lie theory.*** Let H be a bialgebra. The space $\mathcal{P}(H)$ is a Lie subalgebra of H under the commutator bracket. Write $\mathfrak{g}$ for the Lie algebra $\mathcal{P}(H)$. If H is connected and cocommutative, the Cartier–Milnor–Moore (CMM) theorem yields a canonical isomorphism of Hopf algebras

$$H \cong \mathcal{U}(\mathfrak{g})$$

between H and the enveloping algebra of $\mathfrak{g}$.

Let $\mathcal{S}(V)$ denote the symmetric algebra on a space V. It carries a unique Hopf algebra structure for which the elements of V are primitive. The Poincaré–Birkhoff–Witt (PBW) theorem furnishes canonical isomorphisms

$$\mathcal{U}(\mathfrak{g}) \cong \mathcal{S}(\mathfrak{g}) \quad \text{and} \quad \operatorname{gr}\mathcal{U}(\mathfrak{g}) \cong \mathcal{S}(\mathfrak{g}),$$

where $\mathfrak{g}$ is an arbitrary Lie algebra. The former is an isomorphism of coalgebras, the latter of Hopf algebras. Here $\mathcal{S}(\mathfrak{g})$ is the symmetric (Hopf) algebra on the vector space underlying $\mathfrak{g}$.

Proofs of these classical results can be found in [Cartier 2007, §3.8], [Milnor and Moore 1965, § 7] and [Quillen 1969, Appendix B]. For additional references, see [Kassel 1995, Theorem V.2.5], [Loday 1992, §3.3.4 and Appendix A], and [Montgomery 1993, Theorem 5.6.5].

Lemma 1 provides a construction of a commutative connected Hopf algebra gr H from an arbitrary connected Hopf algebra H. This will enable us to employ CMM and PBW in a wider setting than that of (co)commutative Hopf algebras.

1.6. ***Graded bialgebras.*** The bialgebra H is *graded* if there is given a vector space decomposition

$$H = \bigoplus_{m \geq 0} H_m$$

such that

$$\mu(H_p \otimes H_q) \subseteq H_{p+q} \quad \text{for all } p, q \geq 0,$$
$$\Delta(H_m) \subseteq \bigoplus_{p+q=m} H_p \otimes H_q \quad \text{for all } m \geq 0,$$
$$\operatorname{Im}(\iota) \subseteq H_0, \qquad \text{and} \qquad H_m \subseteq \operatorname{Ker}(\varepsilon) \quad \text{for all } m > 0.$$

(The condition on the unit map ι simply states that $1 \in H_0$.) If H is Hopf and the antipode satisfies

$$\mathrm{S}(H_m) \subseteq H_m$$

for all $m \geq 0$, we say that H is a graded Hopf algebra.

If H is an arbitrary bialgebra, then $\operatorname{gr} H$ is a graded bialgebra for which the component of degree n is $H^{(n)}/H^{(n-1)}$.

Let H be a graded bialgebra. The space $\mathcal{P}(H)$ is then graded with $\mathcal{P}(H)_m = \mathcal{P}(H) \cap H_m$. Moreover, $\mathcal{P}(H)$ is then a graded Lie algebra. Similarly, each subspace $H^{(n)}$ is graded with $(H^{(n)})_m = H^{(n)} \cap H_m$. Hence, $\operatorname{gr} H$ inherits a second grading for which

$$(\operatorname{gr} H)_m := (H^{(0)})_m \oplus \big((H^{(1)})_m/(H^{(0)})_m\big) \oplus \big((H^{(2)})_m/(H^{(1)})_m\big) \cdots . \tag{12}$$

Moreover, $H^{(0)} \subseteq H_0$.

We say that H is *graded connected* if $\dim H_0 = 1$. The preceding implies that in this case $H_0 = H^{(0)}$, and therefore H is indeed connected.

Assume that H is a graded connected bialgebra. One may show by induction that $H_m \subseteq H^{(m)}$ for all m. It follows that the sum (12) stops at $(H^{(m)})_m\big/(H^{(m-1)})_m$. It also follows, from (6), that H is a graded Hopf algebra and

$$\mathrm{S}|_{H_m} = \sum_{k=0}^{m} (\iota \circ \varepsilon - \mathrm{id})^{*k}\Big|_{H_m}. \tag{13}$$

More generally, it follows from (7) that

$$\Psi_n|_{H_m} = \sum_{k=0}^{m} \binom{n}{k} (\mathrm{id} - \iota \circ \varepsilon)^{*k}\Big|_{H_m}. \tag{14}$$

2. Characteristic polynomials of the Adams operators

This section contains the main result (Theorem 3), which determines the characteristic polynomials of the Adams operators on a graded connected Hopf algebra H. These only depend on the dimension sequence of H. A number of consequences about the antipode are also presented. Some preliminaries on enumeration of multisets are reviewed first.

2.1. ***Sampling with replacement and the symmetric algebra.*** Given a sequence $\boldsymbol{g} = (g_i)_{i \geq 1}$ and a partition $\lambda = 1^{k_1} 2^{k_2} \cdots r^{k_r}$, put

$$\binom{\boldsymbol{g}}{\lambda} := \binom{g_1 + k_1 - 1}{k_1} \cdots \binom{g_r + k_r - 1}{k_r}. \tag{15}$$

Let $|\lambda| := k_1 + 2k_2 + \cdots + rk_r$ denote the size of λ, and $\ell(\lambda) := k_1 + k_2 + \cdots + k_r$ denote the number of parts of λ. Given nonnegative integers k and m, set

$$\mathsf{mul}(k, m) := \sum_{\substack{|\lambda| = m \\ \ell(\lambda) = k}} \binom{\boldsymbol{g}}{\lambda}. \tag{16}$$

In particular,

$$\mathsf{mul}(0,m)=\delta(0,m) \qquad \text{and} \qquad \mathsf{mul}(k,m)=0 \quad \text{for all } k>m. \tag{17}$$

The numbers $\mathsf{mul}(k,m)$ depend on the given sequence, although this is not reflected in the notation.

If we sample with replacement from a set with g_i elements of *weight* i, then $\binom{g}{\lambda}$ counts the number of samples with weight distribution λ; in other words, the number of multisets of cardinality $\ell(\lambda)$ containing exactly k_i elements of weight i for each $i=1,\dots,r$. The numbers $\mathsf{mul}(k,m)$ then count the number of multisets of cardinality k and total weight m.

Now let W be a positively graded vector space, and let $g_i := \dim W_i$ for each $i \geq 1$. Let W^k denote the k-th symmetric power of W. In other words, W^k is the subspace of the symmetric algebra $\mathcal{S}(W)$ spanned by k-fold products of elements of W. It inherits a grading from W, where $(W^k)_m$ is spanned by products $w_1\cdots w_k$ with $\deg(w_1)+\cdots+\deg(w_k)=m$.

Fix a homogeneous basis of W, and let it be our sample set. The set of monomials of length k (that is, k-fold products of basis elements of W) is then a basis of W^k. A multiset of cardinality k and total weight m corresponds to a monomial of length k and degree m. Therefore,

$$\dim(W^k)_m = \mathsf{mul}(k,m).$$

The bivariate generating series for $\mathcal{S}(W)$ by length and degree is then

$$\sum_{k,m\geq 0} \mathsf{mul}(k,m)\, s^k t^m = \prod_{i\geq 1}(1-st^i)^{-g_i}. \tag{18}$$

This follows by expanding the right-hand side with the aid of the binomial theorem and employing (15) and (16).

The numbers $\mathsf{mul}(k,m)$ enter in Theorem 3 below.

2.2. ***Characteristic polynomial.*** We state some standard results from linear algebra.

Lemma 2. *Let V be a finite-dimensional vector space and $T \in \mathrm{End}(V)$ a linear transformation.*

(i) *Let U be a T-invariant subspace of V. If $T^{\uparrow} \in \mathrm{End}(V/U)$ denotes the linear transformation induced by T on the quotient, and $T_{\downarrow} \in \mathrm{End}(U)$ denotes the restriction of T to U, then the characteristic polynomials of these three maps satisfy*

$$\chi_T(x) = \chi_{T_\downarrow}(x)\, \chi_{T^\uparrow}(x).$$

(ii) *The characteristic polynomials of T and of the dual map $T^* \in \mathrm{End}(V^*)$ are equal.* □

We are now ready for our main result. Let H be a graded connected Hopf algebra. We assume from this point onwards that the homogeneous components H_m of H are finite-dimensional. We let $\widetilde{H}$ denote the *graded dual* of $\operatorname{gr} H$ with respect to the grading (12). It is a graded bialgebra with homogeneous components

$$\widetilde{H}_m := \big((\operatorname{gr} H)_m\big)^*.$$

If H is graded connected, $\widetilde{H}$ is graded connected and cocommutative, by Lemma 1. We let $\mathfrak{g} := \mathcal{P}(\widetilde{H})$ denote the graded Lie algebra of primitive elements of $\widetilde{H}$. For each $i \geq 1$, let

$$g_i := \dim \mathfrak{g}_i$$

denote the dimension of the homogeneous component of $\mathfrak{g}$ of degree i. Consider the corresponding numbers $\mathsf{mul}(k,m)$, as in (16).

Theorem 3. *Let H be as above. For every scalar n and nonnegative integer m, the characteristic polynomial of the restriction $\Psi_n|_{H_m}$ of the n-th Adams operator is*

$$\chi(\Psi_n|_{H_m})(x) = \prod_{k=0}^{m} (x - n^k)^{\mathsf{mul}(k,m)}. \tag{19}$$

Before the proof, a few remarks are in order. First, note that the factor indexed by $k = 0$ is nontrivial only when $m = 0$, according to (17). Second, note that the exponents $\mathsf{mul}(k,m)$ do not depend on n. Additional information on the exponents $\mathsf{mul}(k,m)$ is provided in Proposition 4.

Proof. First of all, it suffices to establish (19) when n is a nonnegative integer. Indeed, both sides depend polynomially on n — the left-hand side in view of (14).

We argue that we may replace H with $\operatorname{gr} H$. Indeed, since Ψ_n preserves both the grading and the coradical filtration of H, it preserves the filtration

$$(H^{(0)})_m \subseteq (H^{(1)})_m \subseteq \cdots \subseteq (H^{(m)})_m = H_m$$

for each m. By repeated application of Lemma 2(i), we deduce that

$$\chi(\Psi_n|_{H_m}) = \chi\big(\operatorname{gr}(\Psi_n)|_{(\operatorname{gr} H)_m}\big).$$

In addition, by (5),

$$\operatorname{gr}(\Psi_n) = \operatorname{gr}(\mathrm{id}^{*n}) = (\operatorname{gr}\mathrm{id})^{*n} = \mathrm{id}^{*n} = \Psi_n.$$

Therefore,

$$\chi(\Psi_n|_{H_m}) = \chi(\Psi_n|_{(\operatorname{gr} H)_m}).$$

Next, we may replace $\operatorname{gr} H$ with $\widetilde{H}$. Indeed, the map $T \mapsto T^*$ is an isomorphism of convolution algebras $\operatorname{End}(H) \cong \operatorname{End}(H^*)$ (where duals and endomorphisms are in the graded sense). Together with Lemma 2(ii) this implies that

$$\chi(\Psi_n|_{(\mathrm{gr}\,H)_m}) = \chi(\Psi_n|_{\widetilde{H}_m}).$$

Now let $\mathfrak{g} = \mathcal{P}(\widetilde{H})$. By CMM, $\widetilde{H} \cong \mathcal{U}(\mathfrak{g})$, and by PBW, $\mathrm{gr}\,\mathcal{U}(\mathfrak{g}) \cong \mathcal{S}(\mathfrak{g})$ as Hopf algebras. The same argument as above shows that we may replace $\widetilde{H}$ with $\mathcal{S}(\mathfrak{g})$.

As $\mathcal{S}(\mathfrak{g})$ is cocommutative, the Eulerian idempotents are available. From (11) we have that

$$\chi(\Psi_n|_{\mathcal{S}(\mathfrak{g})_m}) = \prod_{k\geq 0} \chi(n^k\,\mathsf{e}^{(k)}\,|_{\mathcal{S}(\mathfrak{g})_m}).$$

It thus suffices to calculate the characteristic polynomial of the $\mathsf{e}^{(k)}$ on $\mathcal{S}(\mathfrak{g})$.

Finally, the action of $\mathsf{e}^{(k)}$ on $\mathcal{S}(\mathfrak{g})$ is simply a projection onto $\mathfrak{g}^k$, the subspace spanned by k-fold products of elements of $\mathfrak{g}$. It follows that

$$\chi(n^k\,\mathsf{e}^{(k)}\,|_{\mathcal{S}(\mathfrak{g})_m})(x) = (x-n^k)^{\mathsf{mul}(k,m)},$$

where

$$\mathsf{mul}(k,m) = \dim\,(\mathfrak{g}^k)_m.$$

This completes the proof. □

Remark. One may easily see that the Adams operators act on $\mathcal{S}(W)$ as follows:

$$\Psi_n(w_1\cdots w_k) = n^k w_1\cdots w_k,$$

where $w_i \in W$, $i = 1,\ldots,k$. The proof of Theorem 3 can then be completed without explicit mention of the Eulerian idempotents.

On the other hand, assume that H is a graded connected Hopf algebra that is either commutative or cocommutative. The expression (11) for the Adams operators in terms of the Eulerian idempotents shows that the former are simultaneously diagonalizable. The thesis of Amy Pang [2014] contains a discussion of a common eigenbasis for the Adams operators on such H.

The exponents $\mathsf{mul}(k,m)$ are determined by the dimension sequence of $\mathfrak{g}$, through (15) and (16). In turn, this sequence is related to the dimension sequence of H by

$$1+\sum_{m\geq 1} h_m t^m = \prod_{i\geq 1}(1-t^i)^{-g_i}, \tag{20}$$

where $h_m := \dim H_m$. Indeed, the right-hand side is the generating series for $\mathcal{S}(\mathfrak{g})$, and we have from the proof of Theorem 3 that $H_m \cong \widetilde{H}_m \cong \mathcal{S}(\mathfrak{g})_m$ as vector spaces. It follows that the sequences (g_i) and (h_m) determine each other. In particular:

Proposition 4. *The exponents* $\mathsf{mul}(k,m)$ *are determined by the dimension sequence of* H.

Remark. Equation (20) says that the sequence $(h_m)_{m\geq 1}$ is the *Euler transform* of $(-g_i)_{i\geq 1}$, in the sense of [Sloane and Plouffe 1995, p. 20]. The nonnegativity of the sequence (g_i) restricts the class of sequences (h_m) that may be realized as dimension sequences of graded connected Hopf algebras.

2.3. *Eigenvalues of the antipode.* Let H be a graded connected Hopf algebra and S its antipode.

Applying Theorem 3 in the case $n = -1$ yields information about the antipode, since $\mathrm{S} = \Psi_{-1}$. We obtain

$$\chi(\mathrm{S}|_{H_m})(x) = (x-1)^{\mathsf{emul}(m)}(x+1)^{\mathsf{omul}(m)}, \tag{21}$$

where

$$\mathsf{emul}(m) := \sum_{k \text{ even}} \mathsf{mul}(k,m) \quad \text{and} \quad \mathsf{omul}(m) := \sum_{k \text{ odd}} \mathsf{mul}(k,m).$$

In particular:

Corollary 5. *The eigenvalues of the antipode are* ± 1.

Remark. Corollary 5 fails for general Hopf algebras. Let ω be a primitive cube root of unity and consider Taft's Hopf algebra $T_3(\omega)$ [Taft 1971], with generators $\{g, x\}$ and relations $\{g^3 = 1,\ x^3 = 0,\ gx = \omega\, xg\}$. The coproduct and antipode are determined by $\Delta(g) = g \otimes g$, $\mathrm{S}(g) = g^{-1}$, $\Delta(x) = 1 \otimes x + x \otimes g$, and $\mathrm{S}(x) = -xg^{-1}$. Here $x^2 + \omega\, x^2 g$ is an eigenvector of S with eigenvalue ω.

From Corollary 5 we deduce:

Corollary 6. *The antipode is diagonalizable if and only if it is an involution.*

Remark. The antipode of a graded connected Hopf algebra need not be an involution (or equivalently, diagonalizable). For example, consider the Malvenuto–Reutenauer Hopf algebra of permutations (Example 8). Its antipode is of infinite order already on the homogeneous component of degree 3; see Remark 5.6 in [Aguiar and Sottile 2005b].

2.4. *Composition powers of the antipode.* Consider S^2, the composition of the antipode S with itself. Corollary 5 implies that 1 is the only eigenvalue of S^2, so $\mathrm{S}^2 - \mathrm{id}$ is nilpotent on each homogeneous component. We have a more precise result.

Proposition 7. *The map* $(\mathrm{S}^2 - \mathrm{id})|_{H_m}$ *is nilpotent of order at most* m.

Proof. Since $\mathrm{gr}(H)$ is commutative, $\mathrm{S}^2 = \mathrm{id}$ on $\mathrm{gr}(H)$. Hence $(\mathrm{S}^2 - \mathrm{id})(H^{(m)}) \subseteq H^{(m-1)}$ for all $m \geq 1$. On $H^{(1)} = \Bbbk \oplus \mathcal{P}(H)$, we have $\mathrm{S} = \pm\mathrm{id}$, and then $\mathrm{S}^2 - \mathrm{id} = 0$. By induction, $(\mathrm{S}^2 - \mathrm{id})^m(H^{(m)}) = 0$ for all $m \geq 1$. The statement follows by recalling that $H_m \subseteq H^{(m)}$. $\square$

Remark. Let d_m be the order of nilpotency of $(\mathrm{S}^2-\mathrm{id})|_{H_m}$, so that $1 \le d_m \le m$ by Proposition 7. The lower bound $d_m = 1$ is attained by any involutory Hopf algebra for all m. Computations suggest that the Hopf algebra of *signed permutations* [Bonnafé and Hohlweg 2006] attains the upper bound $d_m = m$ for all $m \ge 1$.

Example 8. Consider the Hopf algebra $H = \mathfrak{S}\mathrm{Sym}$ introduced by Malvenuto and Reutenauer [1995]. We claim that $d_m \le m-1$ for $m \ge 2$ (and $d_1 = 1$). Let $\mathrm{id}_m = 123\cdots m$ be the identity permutation, in one-line notation, and let ω_m be the longest permutation of m elements, $\omega_m = m\cdots 321$. Let $\mathcal{F}$ and $\mathcal{M}$ denote the *fundamental* and *monomial* bases of H, in the notation of [Aguiar and Sottile 2005b]. It follows from Corollary 6.3 of the same paper that

$$H_m = (H_m \cap H^{(m-1)}) \oplus K_m,$$

where K_m is the one-dimensional space spanned by $\mathcal{F}_{\omega_m} = \mathcal{M}_{\omega_m}$. The proof of Proposition 7 shows that on $H_m \cap H^{(m-1)}$ the order of nilpotency is at most $m-1$. On the other hand, it is easy to see that

$$\mathrm{S}(\mathcal{F}_{\omega_m}) = (-1)^m \mathcal{F}_{\mathrm{id}_m} \quad \text{and} \quad \mathrm{S}(\mathcal{F}_{\mathrm{id}_m}) = (-1)^m \mathcal{F}_{\omega_m}.$$

Therefore, S^2 is the identity on K_m. The claim follows.

We turn to higher composition powers of the antipode. Since H^{op} is another graded connected Hopf algebra, it possesses an antipode, and this is the inverse of S by [Montgomery 1993, Lemma 1.5.11] or [Aguiar and Mahajan 2013, Proposition 1.23]. In particular, S is invertible, and we may consider powers S^n for any integer n.

Proposition 9. *For any integer n,*

$$\chi(\mathrm{S}^n|_{H_m})(x) = \begin{cases} \chi(\mathrm{id}|_{H_m})(x) & \text{if } n \text{ is even,} \\ \chi(\mathrm{S}|_{H_m})(x) & \text{if } n \text{ is odd.} \end{cases} \tag{22}$$

Proof. As in the proof of Theorem 3, we may assume that H is commutative. In this case, $\mathrm{S}^2 = \mathrm{id}$ and hence $\mathrm{S}^n = \mathrm{id}$ for even n and $\mathrm{S}^n = \mathrm{S}$ for odd n. □

3. The trace of the Adams operators

We study the trace of the Adams operators on H. The generating functions for their sequences of traces admit closed expressions in terms of the inverse Euler transform of the dimension sequence of H. The generating function for the trace of the antipode is particularly remarkable.

3.1. ***Generating functions for the trace.*** We return to the situation of Theorem 3. Thus, H is a graded connected Hopf algebra, and the integers $\mathrm{mul}(k,m)$ are determined by its dimension sequence through (15), (16), and (20).

As an immediate consequence of this theorem, we have:

Corollary 10. *For all scalars n and nonnegative integers m,*

$$\operatorname{trace}(\Psi_n|_{H_m}) = \sum_{k=0}^{m} n^k \operatorname{mul}(k,m). \tag{23}$$

In particular,

$$\operatorname{trace}(\mathrm{S}|_{H_m}) = \sum_{k=0}^{m} (-1)^k \operatorname{mul}(k,m) = \operatorname{emul}(m) - \operatorname{omul}(m). \tag{24}$$

We turn to generating functions for the trace. First we consider the n-th Adams operator. Recall that the sequence (g_i) is determined by the dimension sequence of H through (20).

Corollary 11. *For all scalars n,*

$$\sum_{m\geq 0} \operatorname{trace}(\Psi_n|_{H_m})\, t^m = \prod_{i\geq 1}(1-nt^i)^{-g_i}. \tag{25}$$

Proof. This follows at once from (18) and (23). □

For each scalar n, let

$$h_n(t) := \sum_{m\geq 0} \operatorname{trace}(\Psi_n|_{H_m})\, t^m$$

denote the generating function for the sequence of traces of Ψ_n. As a consequence of Corollary 11, these functions satisfy certain interesting relations. In order to state them, let

$$\mu_k := \{\omega \in \mathbb{C} \mid \omega^k = 1\}$$

denote the group of complex k-th roots of unity.

Proposition 12. *For each scalar n and positive integer k,*

$$h_{n^k}(t^k) = \prod_{\omega\in\mu_k} h_{\omega n}(t). \tag{26}$$

In particular,

$$h_{n^2}(t^2) = h_n(t)\, h_{-n}(t). \tag{27}$$

Proof. We employ the factorization

$$1 - t^k = \prod_{\omega\in\mu_k} (1-\omega t).$$

Then, from (25), we have

$$h_{n^k}(t^k) = \prod_{i\geq 1}(1-n^k t^{ki})^{-g_i} = \prod_{i\geq 1}\prod_{\omega\in\mu_k} (1-\omega n t^i)^{-g_i} = \prod_{\omega\in\mu_k} h_{\omega n}(t). \quad \square$$

The generating function for the trace of the antipode takes a special form. Let

$$h(t) := 1 + \sum_{m\geq 1} h_m\, t^m$$

denote the generating function for the dimension sequence of H.

Corollary 13.
$$\sum_{m\geq 0} \operatorname{trace}(\mathrm{S}|_{H_m})\, t^m = \frac{h(t^2)}{h(t)}. \tag{28}$$

Proof. Since $\Psi_1 = \mathrm{id}$ and $\Psi_{-1} = \mathrm{S}$, we have $h(t) = h_1(t)$ and

$$\sum_{m\geq 0} \operatorname{trace}(\mathrm{S}|_{H_m})\, t^m = h_{-1}(t).$$

Thus, (28) is the case $n = 1$ of (27). □

Remark. Corollary 13 shows that the sequence of antipode traces is determined by the dimension sequence in a simple manner. The result also shows that, conversely, the dimension sequence is determined by the sequence of antipode traces since the relation (28) can be solved for $h(t)$.

Example 14. Suppose the dimension sequence of H is given by $h_m := r^m$, where r is a fixed nonnegative integer. Then $h(t) = 1/(1-rt)$ and

$$\frac{h(t^2)}{h(t)} = \frac{1-rt}{1-rt^2} = \sum_{n\geq 0} r^n\, t^{2n} - \sum_{n\geq 0} r^{n+1}\, t^{2n+1}.$$

It follows from (28) that

$$\operatorname{trace}(\mathrm{S}|_{H_m}) = \begin{cases} r^{m/2} & \text{if } m \text{ is even,} \\ -r^{(m+1)/2} & \text{if } m \text{ is odd.} \end{cases}$$

We have computed the trace of the antipode without knowing anything other than the dimension sequence of H. A Hopf algebra with the given dimension sequence is the free algebra on r primitive generators of degree 1; a direct computation of the trace can then be carried out. We do this for the dual Hopf algebra in Example 23.

We record the generating function for the trace of the composition powers of the antipode.

Corollary 15. *For any integer n,*

$$\sum_{m\geq 0} \operatorname{trace}(\mathrm{S}^n|_{H_m})\, t^m = \begin{cases} h(t) & \text{if } n \text{ is even,} \\ h(t^2)/h(t) & \text{if } n \text{ is odd.} \end{cases} \tag{29}$$

Proof. This follows from Proposition 9 and Corollary 13. □

3.2. ***The trace of the antipode versus the dimension sequence.*** We write

$$a_m := \mathrm{trace}(\mathrm{S}|_{H_m})$$

for each nonnegative integer m. Let $a(t)$ be its generating function.

We analyze the behavior of the sequence (a_m) in relation to the dimension sequence (h_m).

Corollary 16. *If the sequence (h_m) satisfies a linear recursion with constant coefficients, then so does (a_m).*

Proof. We employ [Stanley 2012, Theorem 4.1.1]. In this situation, the series $h(t)$ is rational, and hence so is $a(t)$ by (28). □

We turn to asymptotics. We assume there exists a meromorphic function $h(z)$ of a complex variable z, holomorphic on a neighborhood of 0, and such that its Taylor expansion is the generating function $h(t)$. It then follows from Pringsheim's theorem [Flajolet and Sedgewick 2009, Theorem IV.6] that a dominant singularity occurs at a positive real number R. Moreover, $R \le 1$. Indeed, if $R > 1$, the coefficients h_m would approach 0 by the exponential growth formula [ibid., Theorem IV.7], and this would force the integers h_m to be 0 from a point on. (As we remark at the end of Section 3.3, this can only happen if H is the one-dimensional Hopf algebra, a triviality which we exclude from consideration.)

Corollary 17. *Suppose that R is the unique singularity of $h(z)$ in the disk $|z| \le R^{1/4}$. Let γ be the order of this singularity. Suppose further that $h(z)$ is nonzero in the disk $|z| \le R^{1/2}$, and*

$$h(-R^{1/2}) \ne \pm h(R^{1/2}).$$

Then

$$\frac{a_m}{h_m} \sim \frac{R^{m/2}}{2^\gamma}\left(\frac{1}{h(R^{1/2})} + (-1)^m \frac{1}{h(-R^{1/2})}\right). \tag{30}$$

Proof. The hypotheses guarantee that R is the unique dominant singularity for $h(z)$ and also that $\pm R^{1/2}$ are the unique dominant singularities for $a(z) = h(z^2)/h(z)$. We then have the standard approximations [Flajolet and Sedgewick 2009, §B.IV; Wilf 2006, §5.2]

$$h_m \sim \frac{1}{\Gamma(\gamma)} m^{\gamma-1} R^{-m} h^*(R)$$

and

$$a_m \sim \frac{2^\gamma}{\Gamma(\gamma)} m^{\gamma-1} R^{-m/2} h^*(R)\left(\frac{1}{h(R^{1/2})} + (-1)^m \frac{1}{h(-R^{1/2})}\right),$$

where Γ is the gamma function and $h^*(R) = \lim_{z\to R}(1-z/R)^\gamma h(z)$. The result follows. □

Example 18. Suppose the dimension sequence is given by

$$h_0 = h_1 = h_2 := 1 \qquad \text{and} \qquad h_m := h_{m-1} + h_{m-2} \quad \text{for all } m \geq 3.$$

Thus, for $m \geq 1$, the h_m are the Fibonacci numbers. In this case,

$$h(z) = \frac{z^2 - 1}{z^2 + z - 1} = (z^2 - 1)(z + \phi)^{-1}(z - 1/\phi)^{-1},$$

where $\phi = (1 + \sqrt{5})/2$ is the golden ratio. The hypotheses of Corollary 17 are satisfied with $R = 1/\phi$ and $\gamma = 1$. We obtain from (30) the approximation

$$\frac{a_m}{h_m} \sim \begin{cases} \phi^{-m/2} & \text{if } m \text{ is even,} \\ -\phi^{(-m+3)/2} & \text{if } m \text{ is odd.} \end{cases}$$

A Hopf algebra with this dimension sequence is discussed in Section 5.4, and the sequence a_m is computed explicitly; see (52). The above may then be seen to follow from the well-known approximation $h_m \sim \phi^m/\sqrt{5}$ for the Fibonacci numbers.

3.3. ***Schur indicators.*** A theme occurring in the recent Hopf algebra literature involves a generalization of the *Frobenius–Schur indicator function* of a finite group. If $\rho : G \to \mathrm{End}(V)$ is a complex representation of G, then the (second) indicator is

$$\nu_2(G, \rho) := \frac{1}{|G|} \sum_{g \in G} \operatorname{trace} \rho(g^2).$$

The only values this invariant can take on irreducible representations are $0, 1, -1$, and this occurs precisely when V is a complex, real, or quaternionic representation, respectively [Serre 1977, Proposition 39]. In [Linchenko and Montgomery 2000], a reformulation of the definition was given in terms of convolution powers of the *integral*[1] in $\mathbb{C}G$. This extended the notion of (higher) Schur-indicators to all finite-dimensional Hopf algebras, and has since become a valuable tool for the study of these algebras [Kashina et al. 2002; Ng and Schauenburg 2008; Sage and Vega 2012; Shimizu 2012]. In the case where ρ is the regular representation (and H is semisimple), it is shown in [Kashina et al. 2006] that the higher Schur indicators can be reformulated further, removing all mention of the integral: for all nonnegative integers n,

$$\nu_{n+1}(H) = \operatorname{trace}(\mathrm{S} \circ \Psi_n).$$

See also [Kashina et al. 2012]. Our results lead to the following formula for these invariants in the case where H is graded connected (instead of finite-dimensional).

[1]This construct, present for finite-dimensional Hopf algebras, is unavailable for general graded connected Hopf algebras.

Corollary 19. *Let H be a graded connected Hopf algebra. Then, for all scalars n and nonnegative integers m,*

$$\operatorname{trace}(\mathrm{S} \circ \Psi_n) = \sum_{k=0}^{m} (-n)^k \,\mathsf{mul}(k,m),$$

where $\mathsf{mul}(k,m)$ *is as in Theorem 3.*

Proof. As in the proof of Theorem 3, we may assume that H is commutative. Then S is a morphism of algebras, and, using (2) and (3), we have

$$\mathrm{S} \circ \Psi_n = \mathrm{S} \circ \mu^{(n-1)} \circ \Delta^{(n-1)} = \mu^{(n-1)} \circ \mathrm{S}^{\otimes n} \circ \Delta^{(n-1)} = \mathrm{S}^{*n} = \Psi_{-n}.$$

The result follows from (23). □

Remark. We mention in passing that the only Hopf algebra H that is at the same time connected and finite-dimensional is the (unique) one-dimensional Hopf algebra. Indeed, combining Lemma 1 with CMM and PBW (as in the proof of Theorem 3), we have $H \cong \mathcal{S}(V)$ as vector spaces for some space V. But the only space V for which the symmetric algebra is finite-dimensional is $V = 0$. Hence $H \cong \Bbbk$. The situation is of course different over fields of positive characteristic or for (-1)-Hopf algebras.

4. The case of cofree graded connected Hopf algebras

We study the characteristic polynomial and the trace of the antipode of a graded connected Hopf algebra that, as a graded coalgebra, is cofree. Since the former are invariant under duality, the results apply as well to graded connected Hopf algebras that are free as algebras. (We make no further mention of this point as we proceed.)

4.1. ***Cofreeness.*** A graded connected Hopf algebra H is *cofree* if, as a graded coalgebra, it is isomorphic to a *deconcatenation* coalgebra $\mathcal{T}^\vee(V)$ on a graded vector space V. The underlying space of the latter is

$$\mathcal{T}^\vee(V) := \bigoplus_{k \geq 0} V^{\otimes k}$$

(the same as that of the tensor algebra $\mathcal{T}(V)$). The coproduct on a k-fold tensor is

$$\Delta(x_1 \cdots x_k) = 1 \otimes (x_1 \cdots x_k) + \sum_{i=1}^{k-1} (x_1 \cdots x_i) \otimes (x_{i+1} \cdots x_k) + (x_1 \cdots x_k) \otimes 1.$$

For more details, see [Aguiar and Mahajan 2010, §2.6] or [Radford 2012, §4.5].

In this situation, we have

$$H^{(m)} \cong \bigoplus_{k=0}^{m} V^{\otimes k} \quad \text{and} \quad \mathcal{P}(H) \cong V$$

as graded vector spaces.

The *shuffle product* of two tensors $x_1 \cdots x_i$ and $x_{i+1} \cdots x_k$ is the following element of $\mathcal{T}^\vee(V)$:

$$\sum_\sigma x_{\sigma^{-1}(1)} \cdots x_{\sigma^{-1}(k)},$$

where the sum is over all permutations $\sigma \in S_k$ such that

$$\sigma(1) < \cdots < \sigma(i) \quad \text{and} \quad \sigma(i+1) < \cdots < \sigma(k).$$

The shuffle product turns the deconcatenation coalgebra $\mathcal{T}^\vee(V)$ into a commutative graded connected Hopf algebra. The antipode acts on a tensor by reversing the components:

$$\mathrm{S}(x_1 x_2 \cdots x_k) = (-1)^k x_k \cdots x_2 x_1. \tag{31}$$

Under the assumption of cofreeness, there exists a stronger version of Lemma 1.

Lemma 20. *Let H be a graded connected Hopf algebra that is cofree as a graded coalgebra. Then*

$$\mathrm{gr}\, H \cong \mathcal{T}^\vee(V)$$

as graded Hopf algebras, where $V = \mathcal{P}(H)$.

This result appears in [Aguiar and Sottile 2005b, Proposition 1.5].

4.2. ***Palindromes and Lyndon words.*** A (weighted) *alphabet* is a set that is graded by the positive integers and whose homogeneous components are finite. The elements of degree n are called *letters of weight n*. A *word* is a sequence of letters in the alphabet and a *palindrome* is a word that coincides with its reversal. The *length* of a word is the number of letters in the word, and the *weight* of a word is the sum of the weights of its letters.

Given an alphabet, let

$$\mathsf{pal}(m) \quad \text{and} \quad \mathsf{npal}(m)$$

denote the number of palindromes of weight m and the number of nonpalindromic words of weight m. Also, let

$$\mathsf{epal}(m) \quad \text{and} \quad \mathsf{opal}(m)$$

denote the number of palindromes of weight m of even and odd length, and let

$$\mathsf{pal}(k, m)$$

denote the number of palindromes of length k and weight m.

Assume now that a cofree graded connected Hopf algebra H is given. Thus, $H \cong \mathcal{T}^\vee(V)$ as graded coalgebras, with $V = \mathcal{P}(H)$. Let

$$v_n := \dim V_n$$

denote the dimension of the space of homogeneous primitive elements of degree n. Let h_n and g_n be as in Proposition 4, so (h_n) is the dimension sequence of H and (g_n) is the dimension sequence of $\mathcal{P}(\widetilde{H})$. Since $H \cong \mathcal{T}^\vee(V)$ as graded vector spaces, we have

$$1 + \sum_{m \geq 1} h_m t^m = \frac{1}{1 - \sum_{n \geq 1} v_n t^n}. \tag{32}$$

Together with (20), this yields

$$1 - \sum_{n \geq 1} v_n t^n = \prod_{i \geq 1} \left(1 - t^i\right)^{g_i}. \tag{33}$$

In particular, the sequences (h_n), (g_n), and (v_n) determine each other.

Fix a homogeneous basis of V, and let it be our alphabet. Thus, there are v_n letters of weight n. Equation (32) then says that h_m is the total number of words of weight m. In particular,

$$\mathsf{npal}(m) = h_m - \mathsf{epal}(m) - \mathsf{opal}(m).$$

Equation (33) says that g_i is the number of *Lyndon words* of weight i in the given alphabet. Then (16) says that $\mathsf{mul}(k, m)$ counts the number of multisets of Lyndon words of cardinality k and total weight m. *Witt's formula* [Kang and Kim 1996, Theorem 2.2] provides an explicit formula for (g_n) in terms of (v_n):

$$g_n = \sum_{d \mid n} \frac{\mu(d)}{d} \sum_{\lambda \vdash n/d} \frac{(\ell(\lambda) - 1)!}{\lambda!} v^\lambda, \tag{34}$$

where $\mu(d)$ is the classical Möbius function, the inner sum is over all partitions $\lambda = 1^{k_1} 2^{k_2} \cdots r^{k_r}$ of n/d, $\ell(\lambda) = k_1 + k_2 + \cdots + k_r$, $\lambda! = k_1! k_2! \cdots k_r!$, and

$$v^\lambda := v_1^{k_1} v_2^{k_2} \cdots v_r^{k_r}.$$

(The special case of (34) in which all letters are of weight 1 appears in [Lothaire 1997, Corollary 5.3.5] and [Reutenauer 1993, Corollary 4.14].)

4.3. ***Characteristic polynomial and trace of the antipode in the cofree case.*** As an alternative to (19) and (24), we have the following expressions for the characteristic polynomial and the trace of the antipode of a cofree graded connected Hopf algebra H.

Theorem 21. *For H as above and any nonnegative integer m,*

$$\chi(\mathrm{S}|_{H_m})(x) = (x+1)^{\mathsf{opal}(m)}(x-1)^{\mathsf{epal}(m)}(x^2-1)^{\mathsf{npal}(m)/2}. \tag{35}$$

Proof. By Lemma 20, gr H is isomorphic to the Hopf algebra $\mathcal{T}^\vee(V)$, where the latter is equipped with the shuffle product and the deconcatenation coproduct. As in Theorem 3, it suffices to analyze the antipode S of this Hopf algebra.

Now, it follows from (31) that each palindrome yields an eigenvector of S. The eigenvalue is ± 1 according to the parity of the length. This explains the first two factors in (35). The nonpalindromic words pair up with their reversals and organize in 2×2 blocks of the form

$$\pm \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

where the sign again depends on the parity of the length. This accounts for the remaining factor. □

As an immediate consequence, we have:

Corollary 22. *For H as above and any nonnegative integer m,*

$$\mathrm{trace}(\mathrm{S}|_{H_m}) = \mathsf{epal}(m) - \mathsf{opal}(m) = \sum_{k=0}^{m} (-1)^k \mathsf{pal}(k,m). \tag{36}$$

Since there are no palindromes of even length and odd weight, (35) and (36) imply

$$\chi(\mathrm{S}|_{H_m})(x) = (x+1)^{\mathsf{pal}(m)}(x^2-1)^{\mathsf{npal}(m)/2}$$

and

$$\mathrm{trace}(\mathrm{S}|_{H_m}) = -\mathsf{pal}(m)$$

for all odd m.

Example 23. Let V be an r-dimensional vector space. We view it as a graded vector space concentrated in degree 1 and consider the Hopf algebra $\mathcal{T}^\vee(V)$. Our alphabet consists of r letters of weight 1, and the palindrome distribution is

$$\mathsf{pal}(k,m) = \begin{cases} r^k & \text{if } m = 2k, \\ r^{k+1} & \text{if } m = 2k+1, \\ 0 & \text{otherwise.} \end{cases}$$

It follows from (36) that

$$\mathrm{trace}(\mathrm{S}|_{H_m}) = \begin{cases} r^{m/2} & \text{if } m \text{ is even}, \\ -r^{(m+1)/2} & \text{if } m \text{ is odd}. \end{cases}$$

We arrived at the same conclusion by different means in Example 14.

We return to the general discussion. From (24) and (36), we deduce

$$\mathsf{epal}(m) - \mathsf{opal}(m) = \mathsf{emul}(m) - \mathsf{omul}(m), \tag{37}$$

or equivalently,

$$\sum_{k=0}^{m} (-1)^k \,\mathsf{pal}(k,m) = \sum_{k=0}^{m} (-1)^k \,\mathsf{mul}(k,m). \tag{38}$$

In general, the pairs $(\mathsf{epal}, \mathsf{opal})$ and $(\mathsf{emul}, \mathsf{omul})$, as well as the triangular arrays pal and mul, are different.

Example 24. Consider again the Malvenuto–Reutenauer Hopf algebra $\mathfrak{S}$Sym. We compare the integers $\mathsf{mul}(k,m)$ and $\mathsf{pal}(k,m)$ for low values of k and m.

Now, $\mathfrak{S}$Sym is cofree and the relevant alphabet is the set of permutations with no global descents; see [Aguiar and Sottile 2005b, Corollary 6.3]. On the component of degree $m = 3$, we have the following distribution of palindromes.

length (k)	1	2	3
permutations	$123, 132, 213$	$231, 312$	321
words on alphabet	$123, 132, 213$	$12\vert 1, 1\vert 12$	$1\vert 1\vert 1$
$\mathsf{pal}(k,3)$	3	0	1

Beneath each permutation, we recorded its expression as words in the alphabet. Counting those words that are palindromic we obtained the integers $\mathsf{pal}(k,3)$.

The integer v_m is the number of permutations of m elements with no global descents. The integers (g_m) are calculated from either (20) or (33). The first few values are as follows.

m	1	2	3	4	5	6
v_m	1	1	3	13	71	461
g_m	1	1	4	17	92	572

The sequences (v_m) and (g_m) are A003319 and A112354 in [OEIS $\geq$ 2015]. Finally, the integers $\mathsf{mul}(k,m)$ are computed from (16). For $m = 3$, we find the following:

k	1	2	3
$\mathsf{mul}(k,3)$	4	1	1

Beyond $m = 3$, the integers $\mathsf{mul}(k,m)$ and $\mathsf{pal}(k,m)$ differ more drastically; see Figure 1. However, the alternating sum of the entries in each column is the same for both arrays, as predicted by (38).

$k \setminus m$	1	2	3	4	5	6
1	1	1	4	17	92	572
2		1	1	5	21	119
3			1	1	5	22
4				1	1	5
5					1	1
6						1

$k \setminus m$	1	2	3	4	5	6
1	1	1	3	13	71	461
2		1	0	1	0	3
3			1	1	4	14
4				1	0	2
5					1	1
6						1

Figure 1. The arrays $\mathsf{mul}(k,m)$ (left) and $\mathsf{pal}(k,m)$ (right) for $\mathfrak{S}\mathrm{Sym}$.

4.4. ***Generating functions.*** We continue to employ the notation of Section 4.2. Let

$$v(t) := \sum_{n\geq 1} v_n t^n$$

be the generating function for the dimension sequence of V.

We have the following generating functions for even and odd palindromes. The functions are bivariate to account for length and weight.

Proposition 25.

$$\sum_{k,m\geq 0} \mathsf{pal}(2k,m)\, s^k t^m = \frac{1}{1-sv(t^2)}, \tag{39}$$

$$\sum_{k,m\geq 0} \mathsf{pal}(2k+1,m)\, s^k t^m = \frac{v(t)}{1-sv(t^2)}. \tag{40}$$

Proof. Consider a palindrome of even length $2k$. Removing the first and last letters (which are equal) yields a palindrome of length $2k-2$. The weights of the two palindromes differ by twice the weight of this letter. Therefore,

$$\mathsf{pal}(2k,m) = v_1 \mathsf{pal}(2k-2,m-2) + v_2 \mathsf{pal}(2k-2,m-4) + \cdots .$$

This recursion leads at once to (39). A similar argument establishes (40). □

Corollary 26.

$$\sum_{m\geq 0} \operatorname{trace}(\mathrm{S}|_{H_m})\, t^m = \frac{1-v(t)}{1-v(t^2)}. \tag{41}$$

Proof. This follows by subtracting (40) from (39), letting $s=1$, and using (36). □

Remark. Corollary 26 is a special case of Corollary 13, in view of (32). The above may be regarded as a *semicombinatorial* proof of this result, which is possible under the cofreeness assumption.

5. Examples

We carry out explicit calculations for the Hopf algebra of symmetric functions and a few related Hopf algebras, focusing on the trace of the antipode. They offer no difficulty, as explicit formulas for the antipodes of these Hopf algebras are known. Our purpose here is simply to illustrate some of the results from the preceding sections.

The paper [Aguiar et al. 2006] contains a concise description of each of the Hopf algebras discussed in this section. Other references are given as we proceed.

5.1. ***Symmetric functions.*** Consider the Hopf algebra of symmetric functions $H = \mathrm{Sym}$. See [Macdonald 1995, Chapter I] for the results used below. On the basis of Schur functions, the antipode acts by $\mathrm{S}(s_\lambda) = (-1)^{|\lambda|} s_{\lambda'}$, where λ' is the partition conjugate to λ. Therefore,

$$\mathrm{trace}(\mathrm{S}|_{H_m}) = (-1)^m c(m), \tag{42}$$

where $c(m)$ is the number of self-conjugate partitions of m.

We turn to Corollary 10. For this Hopf algebra, $g_i = 1$ for all $i \geq 1$. Hence $\binom{g}{\lambda} = 1$ for all λ, and $\mathsf{mul}(k, m) = p_k(m)$, the number of partitions of m into k parts. From (24) we deduce

$$(-1)^m c(m) = \sum_{k=0}^{m} (-1)^k p_k(m). \tag{43}$$

(Note that $p_o(m) = 0$ for $m > 0$.) The number of odd parts in a partition of m has the same parity as m. Hence, the previous identity is equivalent to

$$c(m) = e(m) - o(m), \tag{44}$$

where $e(m)$ and $o(m)$ denote the number of partitions of m with an even number of even parts and with an odd number of even parts. This identity appears in [Aigner 2007, Exercise 1.60] and [Stanley 2012, Chapter 1, Exercise 22(b)].

It is possible to obtain this result more directly as follows. Consider the power sum basis of Sym. Since $\mathrm{S}(p_\lambda) = (-1)^{\ell(\lambda)} p_\lambda$, we have

$$\mathrm{trace}(\mathrm{S}|_{H_m}) = p_e(m) - p_o(m), \tag{45}$$

where $p_e(m)$ and $p_o(m)$ are the number of partitions of m of even length and of odd length. Equating (42) and (45) gives (43) again.

We further illustrate Corollary 10 by deriving certain identities involving the Littlewood–Richardson coefficients $c^{\lambda}_{\mu,\nu}$. Recall that the latter are the structure constants for both the product and coproduct on the Schur basis of Sym,

$$s_\mu \cdot s_\nu = \sum_{\lambda} c^{\lambda}_{\mu,\nu}\, s_\lambda \quad \text{and} \quad \Delta(s_\lambda) = \sum_{\mu,\nu} c^{\lambda}_{\mu,\nu}\, s_\mu \otimes s_\nu.$$

Formula (23) (with $n = \pm 2$) yields the following identities, for all $m \geq 1$:

$$\sum_{\lambda,\mu,\nu \vdash m} (c^{\lambda}_{\mu,\nu})^2 = \sum_{k=1}^{m} 2^k \, p_k(m) \tag{46}$$

and

$$\sum_{\lambda,\mu,\nu \vdash m} c^{\lambda}_{\mu,\nu} c^{\lambda}_{\mu',\nu'} = \sum_{k=1}^{m} (-1)^{m-k} \, 2^k \, p_k(m). \tag{47}$$

Incidentally, the fact that the antipode preserves (co)products says that $c^{\lambda}_{\mu,\nu} = c^{\lambda'}_{\mu',\nu'}$.

5.2. ***Schur P-functions.*** A partition of an integer is *strict* if its parts are all distinct. It is *odd* if each of its parts is odd.

Let λ be a strict partition and $P_\lambda \in \mathrm{Sym}$ the corresponding *Schur P-function*, as in [Macdonald 1995, Section III.8]. Let H be the subspace of Sym spanned by the P_λ, as λ runs over all strict partitions. Then H is a Hopf subalgebra of Sym. We have

$$\mathrm{S}(P_\lambda) = (-1)^{|\lambda|} P_\lambda. \tag{48}$$

Therefore,

$$\mathrm{trace}(\mathrm{S}|_{H_m}) = (-1)^m \, p_d(m),$$

where $p_d(m)$ is the number of strict partitions of m.

It is known that H is the subalgebra of Sym generated by the odd power sums p_{2i+1} for $i \geq 0$. Therefore,

$$\mathrm{trace}(\mathrm{S}|_{H_m}) = (-1)^m \, p_o(m), \tag{49}$$

where $p_o(m)$ is the number of odd partitions of m. Equating (48) and (49) recovers the classical fact that odd and strict partitions are equinumerous [Stanley 2012, Proposition 1.8.5].

Regarding the quantities in Proposition 4, we have that $g_i = 1$ if i is odd and 0 otherwise. It follows that $\binom{g}{\lambda} = 1$ when λ is odd and $\binom{g}{\lambda} = 0$ otherwise. Therefore, $\mathsf{mul}(k, m)$ is the number of odd partitions of m of length k. In an odd partition, the parities of m and k are the same. Thus, identity (24) simply counts odd partitions according to their length.

5.3. ***Quasisymmetric functions.*** Let us turn to the Hopf algebra $H = \mathrm{QSym}$ of quasisymmetric functions. Consider the *fundamental* and *monomial* quasisymmetric functions, denoted by F_α and M_α, respectively. As α runs over the compositions of m, both $\{F_\alpha\}$ and $\{M_\alpha\}$ constitute bases of H_m.

The antipode has the following descriptions:

$$\mathrm{S}(F_\alpha) = (-1)^m F_{\tilde{\alpha}'} \quad \text{and} \quad \mathrm{S}(M_\alpha) = (-1)^{\ell(\alpha)} \sum_{\beta \leq \alpha} M_{\tilde{\beta}},$$

where $\tilde{\gamma}$ is the reversal of γ, γ' is its conjugate (obtained by reflecting the ribbon diagram of γ across the main diagonal), and $\leq$ is the refinement partial order on compositions. Note that $\alpha = \tilde{\alpha}'$ if and only if α is symmetric with respect to reflection across the antidiagonal. There are precisely $2^{(m-1)/2}$ of these when m is odd and zero when m is even. Calculating the trace on the fundamental basis we thus obtain

$$\operatorname{trace}(\mathrm{S}|_{H_m}) = \begin{cases} -2^{(m-1)/2} & \text{if } m \text{ is odd,} \\ 0 & \text{otherwise.} \end{cases} \tag{50}$$

The compositions α that contribute to the trace on the monomial basis satisfy $\tilde{\alpha} \leq \alpha$. Since reversal is an order-preserving involution, this happens if and only if $\tilde{\alpha} = \alpha$, that is, when α is palindromic. Let $\mathsf{pal}(m)$ denote the number of palindromic compositions of m. If m is even, exactly half of the palindromic compositions of m have odd length; if m is odd, all of them do. We conclude that

$$\operatorname{trace}(\mathrm{S}|_{H_m}) = \begin{cases} -\mathsf{pal}(m) & \text{if } m \text{ is odd,} \\ 0 & \text{otherwise.} \end{cases} \tag{51}$$

One may arrive at the same identity from (36).

Equating (50) and (51) we deduce that, for all odd m,

$$\mathsf{pal}(m) = 2^{(m-1)/2}.$$

It is easy to give a direct proof of this fact (and of $\mathsf{pal}(m) = 2^{\lfloor m/2 \rfloor}$ for all $m \geq 0$).

Since QSym is cofree, Theorem 21 applies. The space of primitive elements is spanned by the monomials $M_{(n)}$ for $n \geq 1$. One finds that

$$\mathsf{pal}(k,m) = \begin{cases} \binom{\lceil m/2 \rceil - 1}{\lceil k/2 \rceil - 1} & \text{if } m \text{ is even, or if } m \text{ is odd and } k \text{ is odd,} \\ 0 & \text{if } m \text{ is odd and } k \text{ is even.} \end{cases}$$

Formula (36) boils down in this case to the basic identities $2^h = \sum_{j=0}^{h} \binom{h}{j}$ and $0^h = \sum_{j=0}^{h} (-1)^j \binom{h}{j}$.

Regarding the quantities in Proposition 4, we have by (33) that g_i is the number of Lyndon words of weight i in an alphabet with one letter of weight n for each $n \geq 1$. This number is given by $g_1 = 1$ and

$$g_i = \frac{1}{i} \sum_{d \mid i} \mu(d)\, 2^{i/d}$$

for $i \geq 2$ [Kang and Kim 1996, Proposition 2.3].

5.4. ***Peak quasisymmetric functions.*** Let H be the Hopf algebra of *peak* quasisymmetric functions [Stembridge 1997]. It is a Hopf subalgebra of QSym, with a basis θ_α indexed by compositions α into odd parts (*odd compositions*). The number of odd compositions of m is the Fibonacci number f_m (with $f_0 = f_1 = f_2 = 1$).

A formula for the antipode of H is given in [Billera et al. 2003]:

$$\mathrm{S}(\theta_\alpha) = (-1)^m \theta_{\tilde{\alpha}}.$$

It follows that

$$\operatorname{trace}(\mathrm{S}|_{H_m}) = \begin{cases} f_{m/2} & \text{if } m \text{ is even,} \\ -f_{\lceil m/2 \rceil + 1} & \text{if } m \text{ is odd} \end{cases} \tag{52}$$

(as palindromic odd compositions of m arise from odd compositions of $m/2$). One may arrive at the same identity from (36).

As for QSym, H is cofree. There is one primitive element of degree n for each odd n. One finds that

$$\mathsf{pal}(k,m) = \begin{cases} \binom{(m+k)/4-1}{(m-k)/4} & \text{if } m \text{ is even and } 4 \mid (m-k), \\ \binom{\lfloor (m+k-1)/4 \rfloor}{\lfloor (m-k+1)/4 \rfloor} & \text{if } m \text{ and } k \text{ are odd,} \\ 0 & \text{otherwise.} \end{cases}$$

Information about these numbers can be found in [OEIS $\geq$ 2015, A046854 and A168561].

Formula (36) yields the following basic identities:

$$f_h = \sum_{j=0}^{\lfloor h/2 \rfloor} \binom{h-j-1}{j} \text{ for } h \geq 0 \quad \text{and} \quad f_h = \sum_{j=0}^{h-2} \binom{\lfloor \frac{h+j}{2} \rfloor - 1}{\lfloor \frac{h-j}{2} \rfloor - 1} \quad \text{for } h \geq 2.$$

Appendix A. Hopf monoids in species

The results from the earlier sections admit variants for Hopf monoids in species. We list the main ones in this section, along with indications for the proofs, which are similar to the ones for graded connected Hopf algebras.

This section assumes familiarity with the notion of Hopf monoid in species, as developed in [Aguiar and Mahajan 2010; 2013] and with the notation employed there. For the most part, the latter (shorter) reference suffices. The antipode is discussed in [Aguiar and Mahajan 2013, §5] and the Adams operators (convolution powers of the identity) appear in [loc. cit., §14.4].

All Hopf monoids $\boldsymbol{H}$ are assumed to be connected and finite-dimensional. That is, $\boldsymbol{H}[\varnothing]$ is one-dimensional, and for each finite set I, the vector space $\boldsymbol{H}[I]$ is finite-dimensional.

A.1. ***Characteristic polynomial.*** The starting point is the following result, whose proof is similar to that of [Aguiar and Sottile 2005a, Proposition 1.6].

Lemma 27. *The Hopf monoid* gr $\boldsymbol{H}$ (*associated to the coradical filtration of* $\boldsymbol{H}$) *is commutative.*

Let $h_m := \dim \boldsymbol{H}[m]$, and let

$$\boldsymbol{h}(t) := 1 + \sum_{m\geq 1} h_m \frac{t^m}{m!}$$

denote the exponential generating function for the dimension sequence of $\boldsymbol{H}$.

Theorem 28. *For every scalar n and finite set I, the characteristic polynomial of the restriction $\Psi_n\big|_{\boldsymbol{H}[I]}$ of the n-th Adams operator is of the form*

$$\chi(\Psi_n|_{\boldsymbol{H}[I]})(x) = \prod_{k=0}^{m} (x - n^k)^{\mathsf{xmul}(k,m)}, \tag{53}$$

where $m = |I|$. The nonnegative integers $\mathsf{xmul}(k,m)$ are independent of n and are determined by the dimension sequence of $\boldsymbol{H}$, as follows:

$$\sum_{k,m\geq 0} \mathsf{xmul}(k,m)\, s^k \frac{t^m}{m!} = \boldsymbol{h}(t)^s. \tag{54}$$

Proof. As in the proof of Theorem 3, a combination of Lemma 27 with the PBW and CMM theorems (for Hopf monoids in species [Aguiar and Mahajan 2013, §15]) shows that we can assume that

$$\boldsymbol{H} = \mathcal{S}(\boldsymbol{P})$$

for a certain positive species $\boldsymbol{P}$. Here, $\mathcal{S}(\boldsymbol{P})$ is the free commutative monoid on $\boldsymbol{P}$ with its canonical Hopf monoid structure [loc. cit., §7].

Let $\boldsymbol{p}(t)$ be the exponential generating function for the dimension sequence of $\boldsymbol{P}$. Since $\mathcal{S}(\boldsymbol{P}) = \boldsymbol{E} \circ \boldsymbol{P}$, where $\boldsymbol{E}$ is the exponential species, we have

$$\boldsymbol{h}(t) = \exp(\boldsymbol{p}(t)) \quad \text{and} \quad \boldsymbol{h}(t)^s = \exp(s\,\boldsymbol{p}(t)).$$

On the other hand, a direct calculation of the Adams operators on $\mathcal{S}(\boldsymbol{P})$ shows that the characteristic polynomial is as in (53) and that the integers $\mathsf{xmul}(k,m)$ are determined by

$$\sum_{k,m\geq 0} \mathsf{xmul}(k,m)\, s^k \frac{t^m}{m!} = \exp(s\,\boldsymbol{p}(t)).$$

The result follows. ∎

A.2. ***Trace of the antipode.*** Let

$$\boldsymbol{a}(t) := 1 + \sum_{m \geq 1} \operatorname{trace}(\mathrm{S}|_{\boldsymbol{H}[m]}) \, \frac{t^m}{m!}$$

denote the exponential generating function for the trace of the antipode of a Hopf monoid $\boldsymbol{H}$. This is none other than the reciprocal of the exponential generating function for the dimension sequence.

Corollary 29.
$$\boldsymbol{a}(t) = \frac{1}{\boldsymbol{h}(t)}. \tag{55}$$

Proof. This follows from Theorem 28, taking $n = -1$ in (53) and $s = -1$ in (54). □

Example 30. Let $\boldsymbol{H} = \boldsymbol{\Sigma}$ be the Hopf monoid of set compositions [Aguiar and Mahajan 2013, §11.1]. We have

$$\boldsymbol{h}(t) = \frac{1}{2 - \exp t}.$$

Therefore, $\boldsymbol{a}(t) = 2 - \exp t = 1 - \sum_{m \geq 1} t^m/m!$, and we obtain

$$\operatorname{trace}(\mathrm{S}|_{\boldsymbol{H}[m]}) = -1 \tag{56}$$

for all $m \geq 1$. This result can also be obtained by a direct calculation, starting from either of the expressions for the antipode of $\boldsymbol{\Sigma}$ given in [loc. cit., Proposition 59 or Theorem 60].

For an extension of this result, assume that there is a positive species $\boldsymbol{P}$ such that

$$\boldsymbol{H} \cong \boldsymbol{L} \circ \boldsymbol{P}$$

as species, where $\boldsymbol{L}$ is the species of linear orders. Not every Hopf monoid $\boldsymbol{H}$ is of this form, but this is the case if $\boldsymbol{H}$ is free or cofree [loc. cit., §6]. In this situation, we have

$$\operatorname{trace}(\mathrm{S}|_{\boldsymbol{H}[m]}) = -\dim \boldsymbol{P}[m]. \tag{57}$$

Note that if $\boldsymbol{P} = \boldsymbol{E}$, then $\boldsymbol{H} \cong \boldsymbol{\Sigma}$, and (57) recovers (56).

Example 31. Let $\boldsymbol{H} = \boldsymbol{\Pi}$ be the Hopf monoid of set partitions [Aguiar and Mahajan 2013, §9.3]. We have

$$\boldsymbol{h}(t) = \exp(\exp t - 1).$$

Therefore,

$$\boldsymbol{a}(t) = \exp(1 - \exp t) = 1 - t + \frac{t^3}{3!} + \frac{t^4}{4!} - 2\frac{t^5}{5!} - 9\frac{t^6}{6!} - 9\frac{t^7}{7!} + 50\frac{t^8}{8!} + 267\frac{t^9}{9!} + \cdots.$$

It follows that

$$\operatorname{trace}(\mathrm{S}|_{\boldsymbol{H}[m]}) = \Pi_e(m) - \Pi_o(m), \tag{58}$$

where $\Pi_e(m)$ and $\Pi_o(m)$ denote the number of set partitions of $[m]$ into an even and an odd number of blocks, respectively. This result can also be obtained by a direct calculation, starting from either of the expressions for the antipode of $\mathbf{\Pi}$ given in [loc. cit., Theorem 33 or Proposition 35].

Remark. The Hopf monoid $\mathbf{\Pi}$ is in many ways parallel to the Hopf algebra Sym of symmetric functions (Section 5.1). A consequence of the preceding calculation, however, is that $\mathbf{\Pi}$ does not admit a linear basis that behaves under the antipode in the same manner as the Schur basis of Sym. More precisely, there is no basis $\{s_\pi \mid \pi \vdash [m]\}$ of the space $\mathbf{\Pi}[m]$ with the property that

$$\mathrm{S}(s_\pi) = (-1)^m \, s_{\pi'}$$

for some map $\pi \to \pi'$ on the set of set partitions of $[m]$. Indeed, if this were the case, the sequence of antipode traces would alternate in sign.

For an extension of the calculation in Example 31, let $\boldsymbol{H}$ be a Hopf monoid and $\boldsymbol{P}$ a positive species such that

$$\boldsymbol{H} \cong \boldsymbol{E} \circ \boldsymbol{P}$$

as species. (The first part of the proof of Theorem 28 shows that every connected Hopf monoid is of this form.) In this situation, an $\boldsymbol{H}$-structure on a finite set I is an *assembly* of $\boldsymbol{P}$-structures, in the sense of [Bergeron et al. 1998, §1.4]. Conversely, one may regard a $\boldsymbol{P}$-structure as a *connected* $\boldsymbol{H}$-structure. We then have

$$\operatorname{trace}(\mathrm{S}|_{\boldsymbol{H}[m]}) = h_e(m) - h_o(m), \tag{59}$$

where $h_e(m)$ and $h_o(m)$ denote the number of $\boldsymbol{H}$-structures with an even and odd number of connected components, respectively. If $\boldsymbol{P} = \boldsymbol{E}$, then (59) recovers (58).

The combination of (55) and (59) provide a semicombinatorial description for the reciprocal of any power series arising as the exponential generating function for the dimension sequence of a connected Hopf monoid in species.

Appendix B. q-Hopf algebras

Fix a scalar q. A q-Hopf algebra is a Hopf monoid in the lax braided monoidal category of graded vector spaces, with lax braiding $V \otimes W \to W \otimes V$ given by

$$x \otimes y \mapsto q^{mn} \, y \otimes x,$$

where $x \in V$ and $y \in W$ are homogeneous elements of degrees m and n. If $q = 1$, a q-Hopf algebra is just a graded Hopf algebra as in Section 1.6. For information on q-Hopf algebras, see [Aguiar and Mahajan 2010, §2.3].

In this section we discuss extensions of some of the main results from earlier sections to the context of connected q-Hopf algebras.

B.1. ***Cofreeness for connected q-Hopf algebras.*** Let V be a graded vector space. When $x \in V_n$, we write $|x| = n$. The deconcatenation coalgebra on V (Section 4.1), endowed with the *q-shuffle product*, is a connected q-Hopf algebra. The q-shuffle product of two homogeneous tensors $x_1 \cdots x_i$ and $x_{i+1} \cdots x_k$ is the following element of $\mathcal{T}^\vee(V)$:

$$\sum_\sigma q^{\mathsf{inv}_x(\sigma)}\, x_{\sigma^{-1}(1)} \cdots x_{\sigma^{-1}(k)},$$

where the sum is over all permutations $\sigma \in S_k$ such that

$$\sigma(1) < \cdots < \sigma(i), \quad \sigma(i+1) < \cdots < \sigma(k),$$

and

$$\mathsf{inv}_x(\sigma) := \sum_{\substack{a<b \\ \sigma(a)>\sigma(b)}} |x_a||x_b|.$$

We denote the resulting q-Hopf algebra by $\mathcal{T}_q^\vee(V)$. The antipode is given by

$$\mathrm{S}(x_1 x_2 \cdots x_k) = (-1)^k q^{\mathsf{inv}_x(k)}\, x_k \cdots x_2 x_1, \tag{60}$$

where

$$\mathsf{inv}_x(k) := \sum_{a<b} |x_a||x_b|.$$

The following is an extension of Lemma 20. The proof of the latter result given in [Aguiar and Sottile 2005b, Propositions 1.4 and 1.5] yields its extension as well.

Lemma 32. *Let H be a connected q-Hopf algebra that is cofree as a graded coalgebra. Then*

$$\operatorname{gr} H \cong \mathcal{T}_q^\vee(V)$$

as q-Hopf algebras, where $V = \mathcal{P}(H)$.

B.2. ***Characteristic polynomial and trace of the antipode.*** Let H be a connected q-Hopf algebra that is cofree as a graded coalgebra. We lay the groundwork for a description of the characteristic polynomial of such a Hopf algebra. Let a weighted alphabet be given, with v_n letters of weight n, as in Section 4.2. The *multiweight* of a word is the sequence of letter weights. If the word has weight m, its multiweight is a composition of m, and we write $\alpha \vDash m$.

Given a composition α, let

$$\mathsf{pal}(\alpha) \quad \text{and} \quad \mathsf{npal}(\alpha)$$

denote the number of palindromes and nonpalindromes, respectively, of multi-weight α. If $\alpha = (a_1, \ldots, a_k)$, then

$$\mathsf{pal}(\alpha) = \begin{cases} \prod_{i=1}^{\lceil k/2 \rceil} v_{a_i} & \text{if } \alpha = \tilde{\alpha}, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad \mathsf{npal}(\alpha) = \Big(\prod_{i=1}^{k} v_{a_i} \Big) - \mathsf{pal}(\alpha).$$

(Recall that $\tilde{\alpha} = (a_k, \ldots, a_1)$ denotes the reversal of α.) Let $\ell(\alpha)$ denote the length k of α, and let

$$\mathsf{inv}(\alpha) := \sum_{1 \le i < j \le k} a_i a_j .$$

Let H be as above. Fix a homogeneous basis of $V = \mathcal{P}(H)$ and take it as our alphabet. Thus, $v_n = \dim V_n$.

Theorem 33. *For each nonnegative integer m, the characteristic polynomial of the antipode is*

$$\chi(\mathrm{S}|_{H_m})(x) = \prod_{\alpha \vDash m} \big(x - (-1)^{\ell(\alpha)} q^{\mathsf{inv}(\alpha)} \big)^{\mathsf{pal}(\alpha)} (x^2 - q^{2\,\mathsf{inv}(\alpha)})^{\mathsf{npal}(\alpha)/2}. \tag{61}$$

Proof. Lemma 32 allows us to assume that $H = \mathcal{T}_q^{\vee}(V)$. The result follows from (60), as in the proof of Theorem 21. □

In particular, the eigenvalues of the antipode of such a q-Hopf algebra are positive or negative powers of q. We record the resulting expression for the trace.

Corollary 34.
$$\mathrm{trace}(\mathrm{S}|_{H_m}) = \sum_{\alpha \vDash m} (-1)^{\ell(\alpha)} \mathsf{pal}(\alpha)\, q^{\mathsf{inv}(\alpha)}. \tag{62}$$

We recover (35) and (36) as the case $q = 1$ of (61) and (62).

Example 35. Let V be an r-dimensional vector space. View it as a graded vector space concentrated in degree 1 and consider the q-Hopf algebra $\mathcal{T}_q^{\vee}(V)$. Then

$$\mathsf{pal}(\alpha) = \begin{cases} r^{\lceil m/2 \rceil} & \text{if } \alpha = (1^m), \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$\mathrm{trace}(\mathrm{S}|_{H_m}) = (-1)^m r^{\lceil m/2 \rceil} q^{\binom{m}{2}}.$$

This generalizes the conclusion of Example 14.

B.3. *Generating functions.* We continue to assume that H is a connected q-Hopf algebra that is cofree as a graded coalgebra, and $V = \mathcal{P}(H)$. We also assume that $q \neq 0$.

Let

$$v_q(t) := \sum_{n \ge 1} v_n \frac{t^n}{q^{\binom{n}{2}}}.$$

All generating functions in this section will be of this form.

For each pair of nonnegative integers k and m, let

$$\mathsf{pal}_q(k,m) := \sum_{\substack{\alpha \vDash m \\ \ell(\alpha)=k}} \mathsf{pal}(\alpha)\, q^{\mathsf{inv}(\alpha)}.$$

Then (62) may be rewritten as

$$\operatorname{trace}(\mathrm{S}|_{H_m}) = \sum_{k=0}^{m} (-1)^k \mathsf{pal}_q(k,m). \tag{63}$$

We have the following q-generating functions for even and odd palindromes, generalizing Proposition 25.

Proposition 36.

$$\sum_{k,m\geq 0} \mathsf{pal}_q(2k,m)\, s^k \frac{t^m}{q^{\binom{m}{2}}} = \frac{1}{1-s v_{q^2}(t^2)}, \tag{64}$$

$$\sum_{k,m\geq 0} \mathsf{pal}_q(2k+1,m)\, s^k \frac{t^m}{q^{\binom{m}{2}}} = \frac{v_q(t)}{1-s v_{q^2}(t^2)}. \tag{65}$$

Proof. Consider a palindrome of even length $2k$ and weight $m>0$. Its multiweight α is a palindromic composition of m, necessarily of the form

$$\alpha = (a, \beta, a),$$

where a is a positive integer and β is a palindromic composition of $m-2a$. We have

$$\mathsf{inv}(\alpha) = \mathsf{inv}(\beta) + a^2 + 2a(m-2a) \quad \text{and} \quad \mathsf{pal}(\alpha) = v_a\, \mathsf{pal}(\beta).$$

The former is equivalent to

$$\mathsf{inv}(\alpha) - \binom{m}{2} = \mathsf{inv}(\beta) - \binom{m-2a}{2} - 2\binom{a}{2}.$$

Therefore,

$$\begin{aligned}
\frac{\mathsf{pal}_q(2k,m)}{q^{\binom{m}{2}}} &= \sum_{\substack{\alpha \vDash m \\ \ell(\alpha)=k}} \mathsf{pal}(\alpha)\, q^{\mathsf{inv}(\alpha)-\binom{m}{2}} \\
&= \sum_{a\geq 1} \sum_{\substack{\beta \vDash m-2a \\ \ell(\beta)=2k-2}} v_a\, \mathsf{pal}(\beta)\, q^{\mathsf{inv}(\beta)-\binom{m-2a}{2}-2\binom{a}{2}} \\
&= \sum_{a\geq 1} \frac{v_a}{q^{2\binom{a}{2}}} \frac{\mathsf{pal}_q(2k-2,m-2a)}{q^{\binom{m-2a}{2}}}.
\end{aligned}$$

This recursion leads to (64). A similar argument establishes (65). □

We arrive at a generalization of Corollary 26:

Corollary 37.
$$\sum_{m\geq 0} \mathrm{trace}(\mathrm{S}|_{H_m}) \frac{t^m}{q^{\binom{m}{2}}} = \frac{1-v_q(t)}{1-v_{q^2}(t^2)}. \tag{66}$$

Proof. This follows by subtracting (65) from (64), letting $s = 1$, and using (63). □

B.4. *q-deformations.* The results in the Sections B.2 and B.3 apply only under the assumption of cofreeness. For q-Hopf algebras, this hypothesis is less restrictive than it may seem, as we now argue.

Suppose our q-Hopf algebra is obtained by deforming the product of an ordinary Hopf algebra and leaving the unit and the coalgebra structure unchanged. Thus, we have a family of products μ_q on a coalgebra H, turning it into a connected q-Hopf algebra for each q, which we denote by $H(q)$. Assume also that μ_q depends polynomially on q. An example is $\mathcal{T}_q^\vee(V)$, which is a deformation of $\mathcal{T}^\vee(V)$. More generally, the q-Hopf algebras constructed from Hopf monoids in species by means of the functor $\mathcal{K}_{V,q}^\vee$, as in [Aguiar and Mahajan 2010, §19.7], are all of this form.

In this situation, we may consider the 0-Hopf algebra $H(0)$. A result of Loday and Ronco [2006, Theorem 2.6] (see also [Aguiar and Mahajan 2010, Theorem 2.13]) guarantees that $H(0)$ is cofree as a graded coalgebra. Since the coproduct has not been deformed, we have that our q-Hopf algebra $H(q)$ is cofree for all q.

In particular, the preceding results apply to such Hopf algebra deformations. By duality, they also apply in situations where the coproduct has been polynomially deformed while the rest of the structure has been kept.

B.5. (-1)*-Hopf algebras.* The results of Section 2 relied on the PBW and CMM theorems for graded connected Hopf algebras. While these results are not available for general q-bialgebras, they are for $q = \pm 1$. In particular, the eigenvalues of the antipode of a (-1)-Hopf algebra are still ± 1, and the characteristic polynomials of the Adams operators take the form (19). (The multiplicities $\mathsf{mul}(k, m)$ are no longer given by (16).)

References

[Aguiar and Lauve 2013] M. Aguiar and A. Lauve, "Antipode and convolution powers of the identity in graded connected Hopf algebras", pp. 1053–1064 in *25th International Conference on Formal Power Series and Algebraic Combinatorics* (Paris, 2013), Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2013. MR 3091063 Zbl 1294.05165

[Aguiar and Mahajan 2010] M. Aguiar and S. Mahajan, *Monoidal functors, species and Hopf algebras*, CRM Monograph Series **29**, Amer. Math. Soc., Providence, RI, 2010. MR 2012g:18009 Zbl 1209.18002

[Aguiar and Mahajan 2013] M. Aguiar and S. Mahajan, "Hopf monoids in the category of species", pp. 17–124 in *Hopf algebras and tensor categories*, edited by N. Andruskiewitsch et al., Contemp. Math. **585**, Amer. Math. Soc., Providence, RI, 2013. MR 3077234 Zbl 06342020

[Aguiar and Sottile 2005a] M. Aguiar and F. Sottile, "Cocommutative Hopf algebras of permutations and trees", *J. Algebraic Combin.* **22**:4 (2005), 451–470. MR 2006g:16080 Zbl 1094.16024

[Aguiar and Sottile 2005b] M. Aguiar and F. Sottile, "Structure of the Malvenuto–Reutenauer Hopf algebra of permutations", *Adv. Math.* **191**:2 (2005), 225–275. MR 2005m:05226 Zbl 1056.05139

[Aguiar et al. 2006] M. Aguiar, N. Bergeron, and F. Sottile, "Combinatorial Hopf algebras and generalized Dehn–Sommerville relations", *Compos. Math.* **142**:1 (2006), 1–30. MR 2006h:05233 Zbl 1092.05070

[Aigner 2007] M. Aigner, *A course in enumeration*, Graduate Texts in Mathematics **238**, Springer, Berlin, 2007. MR 2008f:05001 Zbl 1123.05001

[Bergeron et al. 1998] F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial species and tree-like structures*, Encyclopedia of Mathematics and its Applications **67**, Cambridge University Press, 1998. MR 2000a:05008 Zbl 0888.05001

[Billera et al. 2003] L. J. Billera, S. K. Hsiao, and S. van Willigenburg, "Peak quasisymmetric functions and Eulerian enumeration", *Adv. Math.* **176**:2 (2003), 248–276. MR 2005a:06003 Zbl 1027.05105

[Bonnafé and Hohlweg 2006] C. Bonnafé and C. Hohlweg, "Generalized descent algebra and construction of irreducible characters of hyperoctahedral groups", *Ann. Inst. Fourier* (*Grenoble*) **56**:1 (2006), 131–181. MR 2007e:20015 Zbl 1098.20011

[Cartier 2007] P. Cartier, "A primer of Hopf algebras", pp. 537–615 in *Frontiers in number theory, physics, and geometry, II*, edited by P. M. Pierre Cartier, Bernard Julia and P. Vanhove, Springer, Berlin, 2007. MR 2008b:16059 Zbl 1184.16031

[Diaconis et al. 2014] P. Diaconis, C. Y. A. Pang, and A. Ram, "Hopf algebras and Markov chains: two examples and a theory", *J. Algebraic Combin.* **39**:3 (2014), 527–585. MR 3183482 Zbl 1291.05220

[Flajolet and Sedgewick 2009] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2009. MR 2010h:05005 Zbl 1165.05001

[Gerstenhaber and Schack 1991] M. Gerstenhaber and S. D. Schack, "The shuffle bialgebra and the cohomology of commutative algebras", *J. Pure Appl. Algebra* **70**:3 (1991), 263–272. MR 92e:13008 Zbl 0728.13003

[Kang and Kim 1996] S.-J. Kang and M.-H. Kim, "Free Lie algebras, generalized Witt formula, and the denominator identity", *J. Algebra* **183**:2 (1996), 560–594. MR 97e:17042 Zbl 0862.17001

[Kashina et al. 2002] Y. Kashina, G. Mason, and S. Montgomery, "Computing the Frobenius–Schur indicator for abelian extensions of Hopf algebras", *J. Algebra* **251**:2 (2002), 888–913. MR 2003f:16061 Zbl 1012.16040

[Kashina et al. 2006] Y. Kashina, Y. Sommerhäuser, and Y. Zhu, "On higher Frobenius–Schur indicators", pp. viii+65 Mem. Amer. Math. Soc. **855**, Amer. Math. Soc., Providence, RI, 2006. MR 2007k:16071 Zbl 1163.16029

[Kashina et al. 2012] Y. Kashina, S. Montgomery, and S.-H. Ng, "On the trace of the antipode and higher indicators", *Israel J. Math.* **188** (2012), 57–89. MR 2897723 Zbl 1260.16027

[Kassel 1995] C. Kassel, *Quantum groups*, Graduate Texts in Mathematics **155**, Springer, New York, 1995. MR 96e:17041 Zbl 0808.17003

[Linchenko and Montgomery 2000] V. Linchenko and S. Montgomery, "A Frobenius–Schur theorem for Hopf algebras", *Algebr. Represent. Theory* **3**:4 (2000), 347–355. MR 2001k:16073 Zbl 0971.16018

[Loday 1992] J.-L. Loday, *Cyclic homology*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **301**, Springer, Berlin, 1992. MR 94a:19004 Zbl 0780.18009

[Loday and Ronco 2006] J.-L. Loday and M. Ronco, "On the structure of cofree Hopf algebras", *J. Reine Angew. Math.* **592** (2006), 123–155. MR 2007b:16084 Zbl 1096.16019

[Lothaire 1997] M. Lothaire, *Combinatorics on words*, Cambridge University Press, 1997. MR 98g:68134 Zbl 0874.20040

[Macdonald 1995] I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed., Oxford University Press, New York, 1995. MR 96h:05207 Zbl 0824.05059

[Malvenuto and Reutenauer 1995] C. Malvenuto and C. Reutenauer, "Duality between quasi-symmetric functions and the Solomon descent algebra", *J. Algebra* **177**:3 (1995), 967–982. MR 97d:05277 Zbl 0838.05100

[Milnor and Moore 1965] J. W. Milnor and J. C. Moore, "On the structure of Hopf algebras", *Ann. of Math.* (2) **81** (1965), 211–264. MR 30 #4259 Zbl 0163.28202

[Montgomery 1993] S. Montgomery, *Hopf algebras and their actions on rings*, CBMS Regional Conference Series in Mathematics **82**, Amer. Math. Soc., Providence, RI, 1993. MR 94i:16019 Zbl 0793.16029

[Ng and Schauenburg 2008] S.-H. Ng and P. Schauenburg, "Central invariants and higher indicators for semisimple quasi-Hopf algebras", *Trans. Amer. Math. Soc.* **360**:4 (2008), 1839–1860. MR 2009d:16065 Zbl 1141.16028

[Novelli et al. 2013] J.-C. Novelli, F. Patras, and J.-Y. Thibon, "Natural endomorphisms of quasi-shuffle Hopf algebras", *Bull. Soc. Math. France* **141** (2013), 107–130. MR 3031674 Zbl 1266.05175

[OEIS] OEIS, "The On–Line Encyclopedia of Integer Sequences", http://oeis.org.

[Pang 2014] A. Pang, *Hopf algebras and Markov chains*, Ph.D. thesis, Stanford University, 2014, https://stacks.stanford.edu/file/druid:vy459jk2393/thesis_submit-augmented.pdf. Zbl 1291.05220

[Patras 1993] F. Patras, "La décomposition en poids des algèbres de Hopf", *Ann. Inst. Fourier* (*Grenoble*) **43**:4 (1993), 1067–1087. MR 95d:16054 Zbl 0795.16028

[Patras and Schocker 2006] F. Patras and M. Schocker, "Twisted descent algebras and the Solomon–Tits algebra", *Adv. Math.* **199**:1 (2006), 151–184. MR 2006k:16086 Zbl 1154.16029

[Quillen 1969] D. Quillen, "Rational homotopy theory", *Ann. of Math.* (2) **90** (1969), 205–295. MR 41 #2678 Zbl 0191.53702

[Radford 2012] D. E. Radford, *Hopf algebras*, Series on Knots and Everything **49**, World Scientific, Hackensack, NJ, 2012. MR 2894855 Zbl 1266.16036

[Reutenauer 1993] C. Reutenauer, *Free Lie algebras*, London Mathematical Society Monographs. New Series **7**, The Clarendon Press, Oxford University Press, New York, 1993. MR 94j:17002 Zbl 0798.17001

[Sage and Vega 2012] D. S. Sage and M. D. Vega, "Twisted Frobenius–Schur indicators for Hopf algebras", *J. Algebra* **354** (2012), 136–147. MR 2012k:16087 Zbl 1280.16027

[Schmitt 1994] W. R. Schmitt, "Incidence Hopf algebras", *J. Pure Appl. Algebra* **96**:3 (1994), 299–330. MR 95m:16033 Zbl 0808.05101

[Serre 1977] J.-P. Serre, *Linear representations of finite groups*, Graduate Texts in Mathematics **42**, Springer, New York, 1977. MR 56 #8675 Zbl 0355.20006

[Shimizu 2012] K. Shimizu, "Some computations of Frobenius–Schur indicators of the regular representations of Hopf algebras", *Algebr. Represent. Theory* **15**:2 (2012), 325–357. MR 2892511 Zbl 1260.16028

[Sloane and Plouffe 1995] N. J. A. Sloane and S. Plouffe, *The encyclopedia of integer sequences*, Academic Press, San Diego, CA, 1995. MR 96a:11001 Zbl 0845.11001

[Stanley 2012] R. P. Stanley, *Enumerative combinatorics*, vol. 1, 2nd ed., Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, 2012. MR 2868112 Zbl 1288.00026

[Stembridge 1997] J. R. Stembridge, "Enriched P-partitions", *Trans. Amer. Math. Soc.* **349**:2 (1997), 763–788. MR 97f:06006 Zbl 0863.06005

[Sweedler 1969] M. E. Sweedler, *Hopf algebras*, W. A. Benjamin, New York, 1969. MR 40 #5705 Zbl 0194.32901

[Taft 1971] E. J. Taft, "The order of the antipode of finite-dimensional Hopf algebra", *Proc. Nat. Acad. Sci. U.S.A.* **68** (1971), 2631–2633. MR 44 #4075 Zbl 0222.16012

[Takeuchi 1971] M. Takeuchi, "Free Hopf algebras generated by coalgebras", *J. Math. Soc. Japan* **23** (1971), 561–582. MR 45 #1958 Zbl 0217.05902

[Wilf 2006] H. S. Wilf, *generatingfunctionology*, 3rd ed., A K Peters, Wellesley, MA, 2006. MR 2006i:05014 Zbl 1092.05001

Communicated by Susan Montgomery
Received 2014-03-27 Revised 2014-10-23 Accepted 2014-12-01

maguiar@math.cornell.edu *Department of Mathematics, Cornell University, Ithaca, NY 14853, United States*

lauve@math.luc.edu *Department of Mathematics and Statistics, Loyola University Chicago, Chicago, IL 60660, United States*

ALGEBRA AND NUMBER THEORY 9:3 (2015)
dx.doi.org/10.2140/ant.2015.9.585

Secant spaces and syzygies of special line bundles on curves

Marian Aprodu and Edoardo Sernesi

On a special line bundle L on a projective curve C we introduce a geometric condition called (Δ_q). When $L = K_C$, this condition implies $\operatorname{gon}(C) \geq q + 2$. For an arbitrary special L, we show that (Δ_3) implies that L has the well-known property (M_3), generalising a similar result proved by Voisin in the case $L = K_C$.

1. Introduction

In this paper we introduce some new geometric methods in the study of the Koszul cohomology groups of a projective curve with coefficients in an invertible sheaf. The basic set-up is as follows.

Let C be a smooth complex projective curve of genus g, and L a very ample line bundle of degree d on C with $h^0(C, L) = r + 1$. Consider a coherent sheaf $\mathcal{F}$ on C and let $V = H^0(C, L)$; one has natural complexes of vector spaces

$$\bigwedge^{p+1} V \otimes H^0(\mathcal{F} \otimes L^{q-1}) \longrightarrow \bigwedge^{p} V \otimes H^0(\mathcal{F} \otimes L^{q}) \longrightarrow \bigwedge^{p-1} V \otimes H^0(\mathcal{F} \otimes L^{q+1}),$$

whose cohomology $K_{p,q}(C, \mathcal{F}; L)$ is called the (p, q) *(mixed) Koszul cohomology group* of C with respect to $\mathcal{F}$ and L. These vector spaces give information about the minimal resolution of the graded module

$$\gamma(C, \mathcal{F}; L) = \bigoplus_k H^0(\mathcal{F} \otimes L^k)$$

over the symmetric (polynomial) algebra $R = S^* V$ in a well-known way (see [Aprodu and Nagel 2010]). The most important cases are obtained for $\mathcal{F} = \mathcal{O}_C$; the corresponding graded R-module $\bigoplus_k H^0(L^k)$ is denoted by $\gamma(C; L)$ and its Koszul

Aprodu thanks the University of Trento and the Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre for hospitality during the preparation of this work. Sernesi is grateful to the IMAR and to the University of Trento for hospitality during the preparation of this work. The authors have been partly supported by a RIP of CIRM-Trento. Aprodu was partly supported by the grant PN-II-ID-PCE-2012-4-0156. Sernesi was partly supported by the project MIUR-PRIN 2010/11 *Geometria delle varietà algebriche.*
MSC2010: primary 14N05; secondary 14N25, 14M12.
Keywords: projective curves, Brill–Noether theory, syzygies, secant loci.

cohomology groups by $K_{p,q}(C; L)$. The choice $L = K_C$ is of central importance, and its study is at the origin of several results and conjectures on this subject. The guiding notions are the so-called properties (N_p).

Definition 1.1. The line bundle L *has property* (N_0) if and only if the natural restriction map $\rho : R \longrightarrow \gamma(C; L)$ is surjective, i.e., L is normally generated. For $p \geq 1$, we say that the bundle L *satisfies the property* (N_p) if and only if it is normally generated and $K_{i,j}(C; L) = 0$ for all $j \neq 1$ and all $1 \leq i \leq p$.

Roughly speaking, (N_p) holds if and only if the minimal resolution of $\gamma(C; L)$ behaves nicely up to the p-th step. These notions have provided an excellent motivation on these problems in two important cases, namely in the case $L = K_C$ and in the case $\deg(L) \gg 0$. As an example, we recall the following:

Theorem 1.2 [Green and Lazarsfeld 1985]. *If* $\deg(L) \geq 2g+1+p$, *then* L *has property* (N_p). *If* $\deg(L) \geq 2g + p$, *then* L *has property* (N_p) *unless* C *is hyperelliptic or* L *embeds* C *in* $\mathbb{P}^{g+p}$ *with a* $(p+2)$*-secant* p*-plane.*

Property (N_p) for special line bundles is also highly interesting; the study of possible divisorial cases in the moduli space of pairs (C, L), for special line bundles L with $h^1(C, L) > 1$ and which fail property (N_p), has revealed a whole class of counterexamples for the slope conjecture [Farkas 2009]. However, the relations between the properties (N_p) and the geometry of the projective model $\varphi_L(C)$ when L is a special line bundle different from K_C, especially if $h^1(C, L) > 1$, remain somewhat mysterious. Already (N_0) and (N_1) have escaped a systematic classification for obvious reasons: normal generation and ideal generation of special projective curves behave essentially wildly and it is therefore very difficult to get even a conjectural picture of how the resolution of $\gamma(C; L)$ might look like (see [Aprodu and Nagel 2010, Section 4.4] for a short discussion).

A possible solution comes from the study of other properties of $\gamma(C; L)$, called (M_q), which were introduced in [Green and Lazarsfeld 1986] for $q \geq 1$. We shall work with a slightly weaker condition than there, in the spirit of [Ehbauer 1994].

Definition 5.3. The line bundle L *has property* (M_q) if $K_{n,1}(C; L) = 0$ for all $n \geq r - q$.

These are properties enjoyed by the *tail* of the resolution of $\gamma(C; L)$; i.e., property (M_q) holds for L if the resolution of $\gamma(C; L)$ has a nice behaviour at the last q steps. Another, perhaps more suggestive, point of view consists of considering the resolution of the module $\gamma(C, K_C; L)$. Since it is dual to $\gamma(C; L)$, properties (M_q) for L correspond to nice behaviour of the *head* of the resolution of $\gamma(C, K_C; L)$. In a landmark paper, Petri [1925] had already focused his attention on the module

$\gamma(C, K_C; L)$ *when L is special*. Arbarello and Sernesi [1978] showed that Petri's analysis contains a proof of (M_1) for all L on a nonrational curve C and a characterisation of the validity of (M_2) when L is special. Note that when $L = K_C$ the self-duality of the resolution of $\gamma(C; K_C)$ implies that property (M_q) is equivalent to property (N_{q-1}), so the result discussed in [Arbarello and Sernesi 1978] generalises Petri's celebrated analysis of the ideal of the canonical model of a nonhyperelliptic curve (see [Saint-Donat 1973]).

The present paper is devoted to the study of (M_3) for a special L. This property has been already studied and characterised for $L = K_C$ by Schreyer [1991], by Voisin [1988] and when $\deg(L) \gg 0$ by Ehbauer [1994]. The main issue in considering the case of any special line bundle, not considered by them, is to find natural geometric conditions on C and L. We introduce the following definition:

Definition 2.3. Assume that $r \geq 4$, and let $2 \leq q \leq 1 + r/2$. We say that a reduced effective divisor $D = x_1 + \cdots + x_{r-q+2}$ on C satisfies condition (Δ_q) with respect to L if the following conditions are satisfied:

(a) $h^0(L(-D)) = q$.

(b) $L(-D)$ is basepoint-free.

(c) $h^0(L(-D + x_i)) = h^0(L(-D))$ for all $i = 1, \ldots, r - q + 2$.

In the case $L = K_C$, a divisor D satisfies condition (Δ_q) if it defines a primitive g^1_{g-q+1}. In general, D defines an $(r-q)$-plane in $\mathbb{P}^r$ which is precisely $(r-q+2)$-secant to $\varphi_L(C) \subset \mathbb{P}^r$. This condition has appeared in [Green and Lazarsfeld 1985] in the case $q = 2$ and in [Voisin 1988], where it is called (H_1), in the case $q = 3$. In both cases they have proved to be the key for (M_2) to hold for K_C (equivalent to Petri's theorem) and for (M_3) to hold for K_C, respectively. More precisely, a divisor $D = x_1 + \cdots + x_{g-1}$ satisfying condition (Δ_2) for K_C defines a primitive g^1_{g-1}, and the existence of such a D can be seen to be equivalent to C being not exceptional, i.e., to $\mathrm{Cliff}(C) \geq 2$: this is how Green and Lazarsfeld arrive at Petri's theorem involving (Δ_2) and using the Mumford–Martens theorem. On the other hand, via an elaborate analysis, Voisin showed for $g \geq 11$ that (Δ_3) plus $\mathrm{Cliff}(C) \geq 3$ imply that a general projection in $\mathbb{P}^5$ of the canonical model of C satisfies (M_3). It is interesting to note that this is achieved by excluding in particular that the projected curve lies in certain surfaces that are intersection of quadrics in $\mathbb{P}^5$. Here one cannot but observe the analogy with the way Ehbauer [1994] proved (M_3) for L such $\deg(L) \gg 0$: while his method is different from Voisin's, he is led to consider the same list of surfaces.

Our main result involves condition (Δ_3) plus a transversality condition as the key hypothesis. Specifically, we prove the following:

Theorem 5.4. *Assume $g \geq 14$, $r \geq 5$, that L is very ample and special of degree $\geq r+13$, that each component of the locus of $(r-1)$-secant $(r-3)$-planes has the expected dimension $r-4$, and that the general such $(r-3)$-plane in each component satisfies (Δ_3) with respect to L. Then L satisfies (M_3) unless* $\mathrm{Cliff}(C) \leq 2$.

The relation between condition (Δ_3) and the vanishing of the $K_{i,1}(C; L)$ for all $i \geq r-3$ is roughly the following: Nonzero elements of the $K_{i,1}(C; L)$ can be seen to correspond to certain subvarieties containing the curve $\varphi_L(C) \subset \mathbb{P}^r$ and defined by quadrics. On the other hand, the existence of divisors satisfying (Δ_3) plays the role of a generality condition which prevents the curve from being contained in such a variety. This simple contradiction works quite efficiently once the curve is projected in $\mathbb{P}^5$, and that's how we prove the theorem. Note that the condition $\mathrm{Cliff}(C) \geq 3$ cannot be removed, as easy examples show.

For higher q we have a similar contradiction. But the verification that (M_q) holds once hypotheses similar to those of the theorem are satisfied becomes much more involved as $q \geq 4$, and would require a classification of certain classes of varieties that is not yet available.

It is interesting to note that in Theorem 1.2 the existence of secant spaces is related to the exceptions to the validity of (N_p); hence, it is not satisfied in general. On the other hand, in Theorem 5.4 the existence of secant spaces, implied by condition (Δ_3), is satisfied in general.

A final note in the case $L = K_C$. The condition $\mathrm{Cliff}(C) \geq 2$ already implies the existence of divisors satisfying (Δ_2). Similarly, the use of condition (Δ_3) made by Voisin [1988] plays a role in the proof, but is not required for the validity of (M_3): all that is required is that $\mathrm{Cliff}(C) \geq 3$; in fact the main difficulty in that work consists of proving that $\mathrm{Cliff}(C) \geq 3$ implies the existence of D satisfying (Δ_3). This suggests, more generally, that $\mathrm{Cliff}(C) \geq q$ might imply the existence of divisors D satisfying (Δ_q) with respect to K_C.

The paper is organised as follows. In Section 2 we introduce the main condition (Δ_q) and study its general properties. In Section 3 we specialise to the case of canonical curves. In Section 4 we relate condition (Δ_q) to the geometry of the curve in $\mathbb{P}^r$, and in Section 5 we recall the definition of syzygy schemes and prove Theorem 5.4.

2. The condition (Δ_q)

2A. ***Secant loci.*** For any $n \geq 1$, we denote by C_n the n-th symmetric product of C and by $\Xi_n \subset C \times C_n$ the universal divisor. Let

$$
\begin{tikzcd}
C & C \times C_n \arrow[l, "\pi"'] \arrow[d, "\pi_n"] \\
 & C_n
\end{tikzcd}
$$

be the projections. For any globally generated line bundle L on C, the sheaf on C_n

$$E_L := \pi_{n*}(\pi^* L \otimes \mathcal{O}_{\Xi_n})$$

is locally free of rank n and is called the *secant bundle* of L. We have a homomorphism of locally free sheaves on C_n

$$\begin{tikzcd} \pi_{n*}\pi^* L \arrow[r, "e_{L,n}"] \arrow[d, equal] & E_L \\ H^0(L) \otimes \mathcal{O}_{C_n} & \end{tikzcd}$$

Note that $e_{L,n}$ is generically surjective if $n \leq r$.

We will denote by $V_n^k(L) \subset C_n$ the closed subscheme defined by the condition

$$\operatorname{rank}(e_{L,n}) \leq k.$$

Standard facts about determinantal subschemes (see, for example, [Arbarello et al. 1985]) imply that if nonempty, then $V_n^k(L)$ has dimension $\geq n-(r+1-k)(n-k)$, which is the *expected dimension*.

Of special interest are the cases $k = n-1$. The scheme $V_n^{n-1}(L)$ is supported on the set of $D \in C_n$ which do not impose independent conditions on L, and its expected dimension is $2n-r-2$. If $n = r$, we can prove the following:

Lemma 2.1. *If $r \geq 4$ then $V_r^{r-1}(L)$ is nonempty and of pure dimension $r-2$.*

Proof. Let Σ be a nonempty component of $V_r^{r-1}(L)$ with $\operatorname{codim}(\Sigma) \leq 1$, i.e., with $\dim(\Sigma) \geq r-1$. Consider the morphisms

$$\begin{tikzcd} C_{r-1} \times C \arrow[r, "\sigma"] \arrow[d, "\pi_{r-1}"] & C_r \\ C_{r-1} & \end{tikzcd}$$

Then $\pi_{r-1}(\sigma^{-1}(\Sigma)) = C_{r-1}$. This implies that if $x_1, \dots, x_{r-1} \in C$ are general points then the pencil $|L(-x_1 - \cdots - x_{r-1})|$ has basepoints, which is impossible. Therefore $V_r^{r-1}(L)$ has pure dimension $r-2$.

For the same reason, if $A = x_1 + \cdots + x_{r-2}$ is a general effective divisor of degree $r-2$, then $L(-A)$ is basepoint-free and not composed with an involution. The plane curve $\Gamma := \varphi_{L(-A)}(C) \subset \mathbb{P}^2$ is singular and birational to C. Letting $x_{r-1}, x_r \in C$ be such that $\varphi_{L(-A)}(x_{r-1}) = \varphi_{L(-A)}(x_r)$ is a singular point of Γ, the divisor $x_1 + \cdots + x_{r-2} + x_{r-1} + x_r$ belongs to $V_r^{r-1}(L)$, which shows nonemptiness. $\square$

Let us record the following useful fact, which is a direct generalisation of [Arbarello et al. 1985, Lemma 1.7, p. 163]:

Lemma 2.2. *Assume that $q \geq 2$, $r-q+2 \geq 4$ and $V^{r-q+1}_{r-q+2}(L) \neq \varnothing$. Then no irreducible component of $V^{r-q+1}_{r-q+2}(L)$ is contained in $V^{r-q}_{r-q+2}(L)$.*

Proof. Let $D = x_1 + \cdots + x_{r-q+2}$ be a general element in a component of $V^{r-q+1}_{r-q+2}(L)$. Assume by contradiction that $D \in V^{r-q}_{r-q+2}(L)$. Then $\dim\langle D\rangle \leq r-q-1$. We may assume that $\langle D\rangle = \langle x_1 + \cdots + x_{r-q+1}\rangle$. Then for a general $x \in C$ we have $\dim\langle x_1 + \cdots + x_{r-q+1} + x\rangle \leq r-q$ and therefore $x_1 + \cdots + x_{r-q+1} + x \in V^{r-q+1}_{r-q+2}(L)$. To conclude, note that $x_1 + \cdots + x_{r-q+1} + x$, D belong to the same component of $V^{r-q+1}_{r-q+2}(L)$ and $\dim\langle D\rangle < \dim\langle x_1 + \cdots + x_{r-q+1} + x\rangle$, contradicting the generality of D. $\square$

A consequence of Lemma 2.2 is that the locally closed subscheme $S_{r-q+2}(L) \subset C_{r-q+2}$ defined as

$$S_{r-q+2}(L) := V^{r-q+1}_{r-q+2}(L) \setminus V^{r-q}_{r-q+2}(L)$$

is dense in any irreducible component of $V^{r-q+1}_{r-q+2}(L)$. In particular, any property which is satisfied by general divisors in any irreducible component of $V^{r-q+1}_{r-q+2}(L)$ is also valid for $S_{r-q+2}(L)$. Note that the expected dimension is $r-2q+2$ in this case. For the particular case $q=2$, Lemma 2.1 shows that the dimension of $S_r(L)$ coincides with the expected dimension $r-2$.

2B. *Condition* ($\boldsymbol{\Delta_q}$). We introduce our basic condition:

Definition 2.3. Assume that $r \geq 4$, and let $2 \leq q \leq 1 + r/2$. We say that a reduced effective divisor $D = x_1 + \cdots + x_{r-q+2}$ on C satisfies condition (Δ_q) with respect to L if the following conditions are satisfied:

(a) $h^0(L(-D)) = q$.

(b) $L(-D)$ is basepoint-free.

(c) $h^0(L(-D + x_i)) = h^0(L(-D))$ for all $i = 1, \dots, r-q+2$.

In terms of projective geometry, the conditions defining (Δ_q) can be rephrased as follows:

(a) The linear span $\langle D\rangle \subset \mathbb{P}^r$ is an $(r-q)$-plane.

(b) $\langle D\rangle \cap C = \mathrm{Supp}(D)$.

(c) $x_1, \dots, x_{r-q+2}$ are in linearly general position in $\langle D\rangle$ (but not in $\mathbb{P}^r$ of course); i.e., $\langle D - x_i\rangle = \langle D\rangle$ for all i.

In terms of symmetric products, the conditions defining (Δ_q) correspond to the following:

(a) $D \in S_{r-q+2}(L)$.

(b) $\{D\} + C \subset S_{r-q+3}(L)$.

(c) $D \notin \mathrm{Im}\{V^{r-q}_{r-q+1}(L) \times C \to C_{r-q+2}\}$.

Note that, from Lemma 2.2, a general point in any irreducible component of $V^{r-q+1}_{r-q+2}(L)$ satisfies condition (a). Clearly, divisors $D = x_1 + \cdots + x_{r-q+2}$ as in Definition 2.3 fill an open subset of $S_{r-q+2}(L)$.

Proposition 2.4. *Assume that L is special and embeds C with a $(r-q+2)$-secant $(r-q)$-plane $\langle D\rangle \subset \mathbb{P}^r$. Then $h^0(\mathcal{O}_C(D)) \le 2$.*

Proof. Assume that $L = K_C(-B)$, and set $r_B := h^0(\mathcal{O}_C(B)) - 1 = h^1(L) - 1$. From the Riemann–Roch theorem applied to L, we obtain $\deg(B) = r_B - r + g - 1$, and hence $\deg(B + D) = g - q + r_B + 1$. From Riemann–Roch applied to $L(-D)$, we obtain $h^0(\mathcal{O}_C(B + D)) = r_B + 2$. Since the addition map of divisors $|B| \times |D| \to |B + D|$ is finite on its image, it follows that $\dim |D| \le 1$. □

Remark 2.5. (i) A divisor D satisfies (Δ_q) with respect to K_C if and only if $|D|$ is a primitive g^1_{g-q+1}. In particular, (Δ_2) is equivalent to $|D|$ being a primitive g^1_{g-1} on C, and such a D does not exist if and only if C is trigonal or a nonsingular plane quintic (see [Green and Lazarsfeld 1985]). Note that hyperelliptic curves are excluded automatically by our assumptions if $L = K_C$. We shall treat the canonical case in a separate section.

(ii) If L is nonspecial of degree $d = g + r \ge 2g$, then there is no divisor $D \in C_{r-g+1}$ satisfying condition (Δ_{g+1}) with respect to L. In fact this would imply that $L(-D)$ is basepoint-free of degree $(g+r)-(r-g+1) = 2g-1$ and dimension $r-(r-g) = g$, and this is impossible. If $g = 1$, this means that no $D \in C_r$ satisfies (Δ_2) with respect to L: in fact, $C \subset \mathbb{P}^r$ has degree $r + 1$ and any r distinct points of C are independent.

Terminology. Assume L to be special and very ample, $h^0(L) = r + 1$, and let $2 \le q \le r - 1$. It is convenient to introduce the following:

- We say that *condition* (Δ_q) *holds on a component* V of $V^{r-q+1}_{r-q+2}(L)$ if the general element $D \in V$ satisfies (Δ_q) with respect to L. We say that (Δ_q) *holds on* C *with respect to* L if it holds on every component of $V^{r-q+1}_{r-q+2}(L)$.
- We say that (Δ_q) holds on C with respect to L *in the strong sense* if it holds, and moreover all components of $V^{r-q+1}_{r-q+2}(L)$ have dimension equal to the expected dimension $r - 2q + 2$. A necessary condition for this to happen is that $r \ge 2q - 2$.
- When we say "$\dim(Z) = d$", we mean that each irreducible component of Z has dimension d.

Most of our results are proved only under the assumption that (Δ_q) holds in the strong sense.

Proposition 2.6. *Assume that* $\dim(V^{r-q+1}_{r-q+2}(L)) = r-2q+2$. *Then* (Δ_q) *holds on* C *with respect to* L *in the strong sense if and only if the following conditions are satisfied*:

(1) $\dim(V^{r-q+1}_{r-q+3}(L)) \le r-2q+1$.

(2) $\dim(V^{r-q}_{r-q+1}(L)) = r-2q$.

Proof. Note that the expected dimension of the locus $V^{r-q+1}_{r-q+3}(L)$ is $r-3q+3 \le r-2q+1$.

The proof relies on the observation that *any* map defined by addition of divisors is finite on its image. Assume $\dim(V^{r-q+1}_{r-q+3}(L)) \le r-2q+1$ and $\dim(V^{r-q}_{r-q+1}(L)) = r-2q$. Let $D \in S_{r-q+2}(L)$ be a general element in an irreducible component. Then by definition $h^0(L(-D)) = q$, hence condition (a) from Definition 2.3 is satisfied. We prove that $L(-D)$ has no basepoints, i.e., condition (b). Suppose that x is a basepoint of $L(-D)$; then $D+x$ is in $V^{r-q+1}_{r-q+3}(L)$ and depends on $r-2q+2$ parameters, contradicting the assumption on $\dim(V^{r-q+1}_{r-q+3}(L))$. We have seen that condition (c) is equivalent to $D \notin \operatorname{Im}\{V^{r-q}_{r-q+1}(L) \times C \to C_{r-q+2}\}$. By the dimensionality assumptions, the image of the addition map cannot fill a dense set of a component of $S_{r-q+2}(L)$.

Conversely, assume that (Δ_q) hold on C with respect to L in the strong sense. Suppose that $V^{r-q}_{r-q+1}(L)$ has a component Z with $\dim(Z) \ge r-2q+1$. Then by the dimensionality hypothesis, the image of the set $Z+C$ inside $V^{r-q+1}_{r-q+2}(L)$ must fill a component, and all its points violate (Δ_q). If there is a component Y of $V^{r-q+1}_{r-q+3}(L)$ having dimension $\ge r-2q+2$, then a general element $D' \in Y$ can be written as $D' = D+x$, where, again by the dimensionality assumption, D must fill a component of $V^{r-q+1}_{r-q+2}(L)$. From the definition, D fails property (b) of (Δ_q), a contradiction. □

Remark 2.7. Recall that in the case $q=2$ the dimension of the locus $V^{r-1}_r(L)$ equals the expected dimension $r-2$ (Lemma 2.1) but it can be reducible: when $L = K_C$ and $g \ge 6$ this happens precisely when C is either trigonal or bielliptic (see [Teixidor i Bigas 1984]). In the trigonal case $V^{g-2}_{g-1}(K_C)$ has two components, and in both of them (Δ_2) does not hold. In the bielliptic case (Δ_2) holds in one component but not in the other. A characterisation of the pairs (C, L) for which $V^{r-1}_r(L)$ is reducible is unknown to us when L is arbitrary.

Lemma 2.8. *Assume* $r \ge 5$ *and* $2 \le q \le (r+1)/2$. *Assume that* (Δ_q) *holds on* C *with respect to* L *in the strong sense. Then, for every general* $x \in C$, (Δ_q) *holds on* C *with respect to* $L(-x)$ *in the strong sense.*

Proof. As noted before, it suffices to prove the same statement for the locally closed subschemes S_{r-q+1}. Let $x \in C$ be a point such that, for each irreducible component of $S_{r-q+2}(L)$, it is not in the support of all divisors of that component and it is in the support of some divisor in it that satisfies (Δ_q) with respect to L. We have a diagram of spaces and maps

$$\begin{tikzcd} C_{r-q+1} \times \{x\} \arrow[r, "\phi"] & C_{r-q+2} \\ S_{r-q+1}(L(-x)) \arrow[u] \arrow[r] & S_{r-q+2}(L) \arrow[u] \end{tikzcd}$$

where all the maps are inclusions. Let $\Sigma \subset S_{r-q+1}(L(-x))$ be an irreducible component. Assume that $\dim(\Sigma) \geq r - 2q + 2$. Then $\phi(\Sigma)$ is a component of $\overline{S_{r-q+2}(L)}$ and all divisors in $\phi(\Sigma)$ contain x in their support. This contradicts our assumptions. The second possibility is that $\dim(\Sigma) = r - 2q + 1$ and that all divisors $D \in \Sigma$ do not satisfy (Δ_q) with respect to $L(-x)$. Then $\phi(\Sigma) \subset \overline{S_{r-q+2}(L)}$ and all $D + x \in \phi(\Sigma)$ do not satisfy (Δ_q) with respect to L. Since this condition is satisfied for a general choice of $x \in C$, we deduce that there is a component of $\overline{S_{r-q+2}(L)}$ with no elements satisfying (Δ_q) with respect to L, a contradiction. $\square$

3. The case $L = K_C$

In this case the notation specialises as follows:

- $V^{g-q}_{g-q+1}(K_C) = C^1_{g-q+1}$.
- $S_{g-q+1}(K_C) = C^1_{g-q+1} \setminus C^2_{g-q+1}$.
- The expected dimension of $V^{g-q}_{g-q+1}(K_C)$ is $g - 2q + 1$.
- A divisor $D \in C^1_{g-q+1}$ satisfies (Δ_q) with respect to K_C for some $q \geq 2$ if and only if it defines a *primitive* g^1_{g-q+1}, i.e., it is complete, basepoint-free and the residual is also basepoint-free.

For brevity, when in this section we say that a *condition* (Δ_q) *is satisfied*, we assume implicitly "with respect to K_C".

The condition (Δ_q) is well defined in the range $2 \leq q \leq g-1$. When $[(g-1)/2] < q \leq g-1$, the existence of a $D \in C_{g-q+1}$ satisfying (Δ_q) is equivalent to the existence of a primitive g^1_{g-q+1} with $g - q + 1 < (g+3)/2$, and therefore C becomes more and more special as q grows, because its gonality decreases. On the other hand, when $2 \leq q \leq [(g-1)/2]$, the condition that there exists D satisfying (Δ_q) should imply that $\mathrm{Cliff}(C) \geq q$ (this is true for $q = 2, 3$, see the Remark 3.4 below). In this range, if this implication is true then the existence of a $D \in C_{g-q+1}$ satisfying (Δ_q) implies that C is more and more general as q grows. We are able to clarify this, assuming only that C has Clifford dimension 1.

Proposition 3.1. *Assume $g \geq 2q+1$ and $q \geq 2$. Consider the following conditions*:

(i) *The condition (Δ_q) holds on C in the strong sense.*

(ii) *$C \subset \mathbb{P}^{g-1}$ is not contained in a q-dimensional variety of minimal degree $g-q$.*

(iii) *For all $1 \leq e \leq q$, there does not exist a $\overline{D} \in C_{e+1}$ satisfying (Δ_{g-e}).*

(iv) $\mathrm{gon}(C) \geq q+2$.

We have (i) $\Longrightarrow$ (ii) $\Longleftrightarrow$ (iii) $\Longleftrightarrow$ (iv).

Proof. (iv) $\Longleftrightarrow$ (iii). $\mathrm{gon}(C) < q+2$ if and only if there exists a primitive g^1_{e+1} for some $1 \leq e \leq q$, and this is equivalent to the existence of $\overline{D} \in C_{e+1}$ satisfying (Δ_{g-e}).

(ii) $\Longleftrightarrow$ (iii). The existence of a primitive g^1_{e+1} for some $1 \leq e \leq q$ is equivalent to the existence of $A \in W^1_{q+1} \setminus W^2_{q+1}$, possibly with basepoints. The union of the linear spans $\langle E \rangle$ for $E \in |A|$ is a q-dimensional variety of minimal degree.

(i) $\Longrightarrow$ (iii). If there exists $\overline{D} \in C^1_{e+1}$ satisfying (Δ_{g-e}) for some $1 \leq e \leq q$, then the locus

$$W := \{\overline{D} + x_1 + \cdots + x_{g-(q+e)} : \overline{D} \in C^1_{e+1} \text{ satisfying } (\Delta_{g-e}),\, x_i \in C\} \subset C_{g-q+1}$$

consists of divisors not satisfying (Δ_q) and has dimension

$$\dim(W) \geq g-(q+e)+1 \geq g-2q+1.$$

Therefore $\overline{W}$ is a component of C^1_{g-q+1}, contradicting (i). □

Remark 3.2. The proof of the implication (i) $\Longrightarrow$ (iii) fails if $g = 2q$. In fact, a general curve C of genus $g = 2q$ has a primitive g^1_{q+1} and (Δ_q) holds on C in the strong sense. In this case $V^q_{q+1}(K_C) = C^1_{q+1}$ is reducible in several components of dimension 1: their number is given by Castelnuovo's formula [Arbarello et al. 1985, p. 211].

Remark 3.3. The implication (ii) $\Longrightarrow$ (i) does not hold. In fact, if C is a bielliptic curve then $\mathrm{gon}(C) = 4$. On the other hand, C^1_{g-1} has two components [Teixidor i Bigas 1984], both having dimension $g-3$, equal to the expected dimension, but (Δ_2) holds only on one of them. Therefore, in this case the implication holds only in a weak sense.

Remark 3.4. If (Δ_2) holds then $\mathrm{Cliff}(C) \geq 2$. This has been proved in [Green and Lazarsfeld 1985] using Mumford–Martens. Note that they only assumed that (Δ_2) holds on *some* component of C^1_{g-1}. The implication (Δ_3) holds $\Longrightarrow \mathrm{gon}(C) \geq 5$ has been considered in [Voisin 1988]. In both cases $q = 2, 3$, the converse implication

$$\mathrm{Cliff}(C) \geq q \implies (\Delta_q) \text{ holds on some component of } C^1_{g-q+1}$$

has also been proved.

Remark 3.5. Assume g is odd. On a general curve C of Clifford dimension 1 there is a $D \in C_{(g+3)/2}$ satisfying $(\Delta_{(g-1)/2})$. The reason is that C has gonality $(g+3)/2$, and a pencil computing its gonality is necessarily primitive. Therefore a divisor D in the pencil satisfies $(\Delta_{(g-1)/2})$.

In the case $L = K_C$, Proposition 2.6 implies:

Proposition 3.6. *Let C be a curve of genus $g \geq 2q+2$ such that the dimension of the locus $W^1_{g-q}(C)$ equals the expected dimension $g-2q-2$ and* $\dim(W^2_{g-q+2}(C)) \leq g-2q-2$. *Then (Δ_q) holds on C in the strong sense.*

Proof. Since $\dim(W^1_{g-q}(C)) = g-2q-2$, we obtain $\dim(V^{g-q-1}_{g-q}(K_C)) = g-2q-1$, which is (2) of Proposition 2.6 in this case. From "excess linear series" it follows that the dimension of $W^1_{g-q+1}(C)$ also equals the expected dimension $g-2q$, and hence $\dim(V^{g-q}_{g-q+1}(K_C)) = g-2q+1$. Finally, $\dim(W^2_{g-q+2}(C)) \leq g-2q-2$ implies that $\dim(V^{g-q}_{g-q+2}(K_C)) \leq g-2q$, as $V^{g-q}_{g-q+2}(K_C) = C^2_{g-q+2}$. Hence all the conditions required in Proposition 2.6 are satisfied. □

Remark 3.7. If the curve C is of gonality $(q+1)$ or less, then the hypotheses of Proposition 3.6 are not satisfied. Indeed, if A is a g^1_{q+1}, then $W^1_{g-q}(C)$ contains the variety $\{A\} + W_{g-2q-1}(C)$, which is of dimension $g-2q-1$.

If the curve C is instead of gonality $(q+2)$, then the hypothesis that

$$\dim(W^1_{g-q}(C)) = \rho(g, 1, g-q) = g-2q-2$$

coincides with the *linear growth condition* on the dimension of Brill–Noether loci, from [Aprodu 2005]. It was proved there that this condition implies Green's conjecture, i.e., condition (M_q).

Remark 3.8. If $q = 2$, and C is neither trigonal, bielliptic nor plane quintic, the hypotheses of Proposition 3.6 are satisfied. Indeed, if one of the two fails, then we obtain a contradiction with the Mumford–Martens dimension theorem. Likewise, for $q = 3$ the failure of the hypotheses contradicts Keem's dimension theorem [Voisin 1988, Proposition II.0].

Corollary 3.9. *Assume that $g \geq 2q+2$, $q \geq 2$, and that* $\dim(C^2_{g-q+2}) \leq g-2q$. *If (Δ_{q+1}) holds on C in the strong sense then (Δ_q) also holds on C in the strong sense.*

Applying Proposition 3.6 and Lemma 2.8, we obtain the following existence result:

Corollary 3.10. *For a general triple (C, L, D), with L special and $D \in V^{r-q+1}_{r-q+2}(L)$, the condition (Δ_q) is satisfied.*

The meaning of generality in the statement is that L is a general projection of the canonical bundle, and hence the speciality index equals 1. More precise existence results are proved by Coppens and Martens, and by Farkas; see [Farkas 2008, Theorem 0.5] and the references therein.

4. Condition (Δ_q) and geometry

Proposition 4.1. *Assume that $r \geq \max\{4, 2q-1\}$. Suppose that L is special and condition (Δ_q) holds on C with respect to L in the strong sense. Then $\varphi_L(C) \subset \mathbb{P}^r$ is not contained in a q-dimensional variety of minimal degree $(r-q+1)$ unless $r = 2q-1$ and C has a basepoint-free g^1_{q+1}.*

Proof. Assume that $r \geq 2q$. We note that C has no g^1_{q+1}. Indeed, if we have a g^1_{q+1}, then $A + C_{r-2q+1}$, with $A \in |g^1_{q+1}|$, fill up a component of $V^{r-q+1}_{r-q+2}(L)$, and any element of this locus fails condition (c) of the definition of (Δ_q).

Assume by contradiction that $\varphi_L(C) \subset X$, a q-dimensional variety of minimal degree $r-q+1$. Then X is ruled by a one-dimensional family of $(q-1)$-planes. Let Λ be a general such $(q-1)$-plane, and let $E = \Lambda \cap \varphi_L(C)$ and $n = \deg(\Lambda \cap \varphi_L(C))$. Then $n \geq q+2$ by what we have just shown. Decompose $E = A + B$ with $\deg(A) = q+1$. Let $D = A + y_1 + \cdots + y_{r-2q+1}$ with the y_i general points of C. Then $D \in V^{r-q+1}_{r-q+2}(L)$, but it does not satisfy (Δ_q). On the other hand, the divisor D depends on $1 + (r-2q+1) = r-2q+2$ parameters. Therefore it is a general point of a component of $V^{r-q+1}_{r-q+2}(L)$, a contradiction.

In the case $r = 2q-1$, the only possibility for C to be on a variety of minimal degree is that C have a basepoint-free g^1_{q+1}, and, in this case, $S_{q+1}(L)$ will have a rational component. The case when X is a cone over the Veronese surface can be treated similarly, by general projection to $\mathbb{P}^{2q-1}$ using Lemma 2.8. □

Note that if C is contained in an e-dimensional variety of minimal degree $(r-e+1)$ with $e \leq q$, then it is contained also in a q-dimensional variety of minimal degree $(r-q+1)$ [Harris 1981].

As we will see, the validity of property (M_3) is tightly connected with properties of surfaces of low degree in $\mathbb{P}^5$. As an illustration of the geometric content of Definition 2.3, we study surfaces of degree $n \leq 6$.

Proposition 4.2. *Assume that $r = 5$, L is special and (Δ_3) holds on C with respect to L in the strong sense. Then $\varphi_L(C) \subset \mathbb{P}^5$ is not contained in a nonsingular surface of degree ≤ 6 unless it has a g^1_4.*

Proof. Assume that $C \subset S$, a nonsingular surface of degree $n \leq 6$. Consider the case $n = 6$. The possibilities for a nonsingular surface of degree 6 in $\mathbb{P}^5$ are described in [Ionescu 1984], and are the following: (i) an elliptic scroll with sectional genus $g = 1$ and $e = 0$; (ii) a Castelnuovo surface with sectional genus $g = 2$ defined by the embedding in $\mathbb{P}^5$ of the blow-up $X = \mathrm{Bl}_{p_1,\dots,p_6,q}(\mathbb{P}^2)$ of $\mathbb{P}^2$ at seven general points via the very ample linear system $|\mathscr{L}| = |4H - E_1 - E_2 - \cdots - E_6 - 2A|$ (with the obvious notation) corresponding to the system of plane quartics passing simply through $p_1, \dots, p_6$ and doubly through q.

In case (i), let $\ell \subset S$ be a general line of the ruling, and let $k = \deg(\mathcal{O}_C(\ell))$. Then $k \geq 2$, and if $k = 2$ then C is bielliptic, so it has a g^1_4. If $k \geq 3$, then adding a general $p \in C$ to a subdivisor of degree 3 of $\mathcal{O}_C(\ell)$ we obtain an element of $S_4(L)$ which does not satisfy (Δ_3) and which depends on two parameters, a contradiction.

In case (ii), the system $|H - A|$ is a pencil of conics on the surface S. The divisors $D \in |\mathcal{O}_C(H - A)|$ have degree say $m \geq 3$ and $\dim |D| \geq 1$. If $m \leq 4$ then C has a g^1_4. Otherwise the divisors D contain subdivisors of degree 4 contradicting the other conditions.

If $n = 5$, then S is a Del Pezzo surface. Let $|\gamma|$ be a pencil of conics on S and let $N = \mathcal{O}_C(\gamma)$. Then N gives a g^1_4 or contradicts (Δ_3), depending on whether $\deg(N) \leq 4$ or $\deg(N) \geq 5$.

If $n = 4$, the conclusion follows from Proposition 4.1. □

5. Condition (Δ_3) and Koszul cohomology

In this section, we briefly recall the relation between Koszul cohomology and vector bundles, as well as the definition of syzygy schemes.

Consider X a smooth projective variety, and let L be a globally generated line bundle on X. We let

$$\varphi_L : X \to \mathbb{P}(H^0(L)^\vee) \cong \mathbb{P}^r, \quad r + 1 = h^0(L)$$

be the morphism defined by L.

We have an exact sequence

$$0 \longrightarrow M_L \longrightarrow H^0(L) \otimes \mathcal{O}_X \longrightarrow L \longrightarrow 0, \tag{1}$$

where $M_L = \varphi^*(\Omega_{\mathbb{P}^r}(1))$ is locally free of rank r. If $r = 1$, i.e., if $|L|$ is a basepoint-free pencil, then $M_L = L^{-1}$. Taking the n-th exterior power $(1 \leq n \leq r)$ we obtain the exact sequence

$$0 \longrightarrow \bigwedge^n M_L \longrightarrow \bigwedge^n H^0(L) \otimes \mathcal{O}_X \longrightarrow \bigwedge^{n-1} M_L \otimes L \longrightarrow 0. \tag{2}$$

For any coherent sheaf $\mathcal{F}$ on X, twisting the sequence above with $\mathcal{F}$, with powers of L and taking global sections, we obtain isomorphisms

$$K_{n,m}(X, \mathcal{F}; L) \cong \operatorname{Coker}\bigl\{\bigwedge^{n+1} H^0(L) \otimes H^0(\mathcal{F} \otimes L^{m-1}) \to H^0\bigl(\bigwedge^n M_L \otimes \mathcal{F} \otimes L^m\bigr)\bigr\}.$$

The syzygy schemes were introduced and studied in [Green 1984; Ehbauer 1994]. The idea behind the definition of syzygy schemes is that one reason for which a linearly normal curve C in $\mathbb{P}^r$ has some nonvanishing $K_{n,1}$ is that C lies on a variety of special type. The varieties under question are cut out by quadrics; more precisely, by the quadrics involved in syzygies.

The general set-up is the following. Let C be a smooth curve, L a globally generated (preferably very ample) line bundle on C and set $V = H^0(L)$. Start with the short exact sequence of sheaves on the projective space

$$0 \longrightarrow \mathcal{I}_C \longrightarrow \mathcal{O}_{\mathbb{P}^r} \longrightarrow \mathcal{O}_C \longrightarrow 0.$$

Note that for any n and m we have $K_{n,m}(\mathbb{P}^r, \mathcal{O}_C; \mathcal{O}_{\mathbb{P}^r}(1)) \cong K_{n,m}(C; L)$.

Taking Koszul cohomology with respect to $\mathcal{O}_{\mathbb{P}^r}(1)$, and using the vanishing of Koszul cohomology on the projective space, we obtain isomorphisms

$$K_{n,m}(C; L) \cong K_{n-1,m+1}(\mathbb{P}^r, \mathcal{I}_C; \mathcal{O}_{\mathbb{P}^r}(1)),$$

for any n and m except for the cases $(n, m) = (0, 0)$ or $(n, m) = (1, -1)$. On the other hand, from the general description of mixed Koszul cohomology, we know that

$$K_{n-1,m+1}(\mathbb{P}^r, \mathcal{I}_C; \mathcal{O}_{\mathbb{P}^r}(1)) \cong \mathrm{Coker}\left\{\bigwedge^n V \otimes H^0(\mathcal{I}_C(m)) \to H^0(\Omega^{n-1}_{\mathbb{P}^r}(n+m) \otimes \mathcal{I}_C)\right\}.$$

Observe that for the case $m = 1$ we have $H^0(\mathcal{I}_C(m)) = 0$, and hence we obtain an isomorphism

$$K_{n,1}(C; L) \cong H^0(\Omega^{n-1}_{\mathbb{P}^r}(n+1) \otimes \mathcal{I}_C);$$

in particular, any nonzero Koszul cohomology class $\alpha \in K_{n,1}(C; L)$ corresponds to a section in $H^0(\Omega^{n-1}_{\mathbb{P}^r}(n+1))$ vanishing along C. The zero-scheme of this section is called the *syzygy scheme* associated to α, and is denoted by $\mathrm{Syz}(\alpha)$. Note that a syzygy scheme is cut out by quadrics, as the sheaf $\Omega^{n-1}_{\mathbb{P}^r}(n+1)$ is a subsheaf of $\bigwedge^{n-1} V \otimes \mathcal{O}_{\mathbb{P}^r}(2)$. The scheme-theoretic intersection of all the syzygy schemes is denoted by $\mathrm{Syz}_n(C)$. It contains C and is cut out by quadrics as well.

We record next two remarkable classification results concerning syzygy schemes, due to Green and Ehbauer.

Theorem 5.1 (Green's $K_{p,1}$). *If $K_{r-1,1}(C, L) \neq 0$, then C is a rational normal curve and* $\mathrm{Syz}_{r-1}(C) = C$. *If C is of degree $\geq r + 2$ and $K_{r-2,1}(C, L) \neq 0$, then* $\mathrm{Syz}_{r-2}(C)$ *is a surface of minimal degree* $(r - 1)$.

Theorem 5.2 (Ehbauer). *If C has degree $\geq r + 13$ and $K_{r-3,1}(C, L) \neq 0$, then* $\mathrm{Syz}_{r-3}(C)$ *is either a surface of minimal degree* $(r - 1)$, *a surface of degree r or a threefold of minimal degree* $(r - 2)$.

We recall the following:

Definition 5.3. The line bundle L *has property* (M_q) if $K_{n,1}(C; L) = 0$ for all $n \geq r - q$.

We prove:

Theorem 5.4. *Assume $g \geq 14$, $r \geq 5$, L is very ample and special of degree $\geq r+13$, and (Δ_3) holds on C with respect to L in the strong sense. Then L satisfies (M_3) unless* $\mathrm{gon}(C) \leq 4$.

Proof. Applying Ehbauer's characterisation of syzygy schemes, if L fails property (M_3), then C lies either on a surface of minimal degree, on a threefold of minimal degree or on a surface of degree r. The first two cases are excluded by Proposition 4.1. Projecting generically to $\mathbb{P}^5$ and applying Lemma 2.8 and Proposition 4.2, we see that C cannot lie on a smooth surface of degree 5. If it lies on a singular surface of degree 5 in $\mathbb{P}^5$, then, projecting from a singular point, the curve in $\mathbb{P}^4$ lies on a surface of minimal degree. In particular, since the curve is of gonality ≥ 5, the image of C in $\mathbb{P}^4$ has a k-secant line for some $k \geq 5$, and hence the image of C in $\mathbb{P}^5$ has a one-dimensional family of k-secant 2-planes with $k \geq 5$, which contradicts the assumptions. $\square$

Remark 5.5. The same argument together with Green's $K_{p,1}$-theorem gives a similar statement for the weaker property (M_2).

References

[Aprodu 2005] M. Aprodu, "Remarks on syzygies of d-gonal curves", *Math. Res. Lett.* **12**:2-3 (2005), 387–400. MR 2006d:14028 Zbl 1084.14032

[Aprodu and Nagel 2010] M. Aprodu and J. Nagel, *Koszul cohomology and algebraic geometry*, University Lecture Series **52**, American Mathematical Society, Providence, RI, 2010. MR 2011f:14051 Zbl 1189.14001

[Arbarello and Sernesi 1978] E. Arbarello and E. Sernesi, "Petri's approach to the study of the ideal associated to a special divisor", *Invent. Math.* **49**:2 (1978), 99–119. MR 80c:14020 Zbl 0399.14019

[Arbarello et al. 1985] E. Arbarello, M. Cornalba, P. A. Griffiths, and J. Harris, *Geometry of algebraic curves, I*, Grundlehren der Mathematischen Wissenschaften **267**, Springer, New York, 1985. MR 86h:14019 Zbl 0559.14017

[Teixidor i Bigas 1984] M. Teixidor i Bigas, "For which Jacobi varieties is Sing Θ reducible?", *J. Reine Angew. Math.* **354** (1984), 141–149. MR 86c:14025 Zbl 0542.14020

[Ehbauer 1994] S. Ehbauer, "Syzygies of points in projective space and applications", pp. 145–170 in *Zero-dimensional schemes* (Ravello, 1992), edited by F. Orecchia and L. Chiantini, de Gruyter, Berlin, 1994. MR 95h:13012 Zbl 0840.14032

[Farkas 2008] G. Farkas, "Higher ramification and varieties of secant divisors on the generic curve", *J. Lond. Math. Soc.* (2) **78**:2 (2008), 418–440. MR 2010f:14029 Zbl 1155.14024

[Farkas 2009] G. Farkas, "Koszul divisors on moduli spaces of curves", *Amer. J. Math.* **131**:3 (2009), 819–867. MR 2010f:14030 Zbl 1176.14006

[Green 1984] M. L. Green, "Koszul cohomology and the geometry of projective varieties", *J. Differential Geom.* **19**:1 (1984), 125–171. MR 85e:14022 Zbl 0559.14008

[Green and Lazarsfeld 1985] M. Green and R. Lazarsfeld, "A simple proof of Petri's theorem on canonical curves", pp. 129–142 in *Geometry today* (Rome, 1984), edited by E. Arbarello et al., Progr. Math. **60**, Birkhäuser, Boston, 1985. MR 88h:14038 Zbl 0577.14018

[Green and Lazarsfeld 1986] M. Green and R. Lazarsfeld, "On the projective normality of complete linear series on an algebraic curve", *Invent. Math.* **83**:1 (1986), 73–90. MR 87g:14022 Zbl 0594.14010

[Harris 1981] J. Harris, "A bound on the geometric genus of projective varieties", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **8**:1 (1981), 35–68. MR 82h:14010 Zbl 0467.14005

[Ionescu 1984] P. Ionescu, "Embedded projective varieties of small invariants", pp. 142–186 in *Algebraic geometry* (Bucharest, 1982), edited by L. Bădescu and D. Popescu, Lecture Notes in Math. **1056**, Springer, Berlin, 1984. MR 85m:14024 Zbl 0542.14024

[Petri 1925] K. Petri, "Über Spezialkurven, I", *Math. Ann.* **93**:1 (1925), 182–209. MR 1512232 Zbl 51.0510.01

[Saint-Donat 1973] B. Saint-Donat, "On Petri's analysis of the linear system of quadrics through a canonical curve", *Math. Ann.* **206** (1973), 157–175. MR 49 #2752 Zbl 0315.14010

[Schreyer 1991] F.-O. Schreyer, "A standard basis approach to syzygies of canonical curves", *J. Reine Angew. Math.* **421** (1991), 83–123. MR 92j:14040 Zbl 0729.14021

[Voisin 1988] C. Voisin, "Courbes tétragonales et cohomologie de Koszul", *J. Reine Angew. Math.* **387** (1988), 111–121. MR 89e:14036 Zbl 0652.14012

Communicated by David Eisenbud
Received 2014-04-25 Revised 2015-01-27 Accepted 2015-03-02

marian.aprodu@imar.ro *Simion Stoilow Institute of Mathematics of the Romanian Academy, P.O. Box 1-764, 014700 Bucharest, Romania*

Faculty of Mathematics and Computer Science, University of Bucharest, 14 Academiei Street, 010014 Bucharest, Romania

sernesi@mat.uniroma3.it *Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, Largo San Leonardo Murialdo, I-00146 Roma, Italy*

ALGEBRA AND NUMBER THEORY 9:3 (2015)
dx.doi.org/10.2140/ant.2015.9.601

Complex group algebras of the double covers of the symmetric and alternating groups

Christine Bessenrodt, Hung Ngoc Nguyen,
Jørn B. Olsson and Hung P. Tong-Viet

We prove that the double covers of the alternating and symmetric groups are determined by their complex group algebras. To be more precise, let $n \geq 5$ be an integer, G a finite group, and let $\widehat{\mathsf{A}}_n$ and $\widehat{\mathsf{S}}_n^{\pm}$ denote the double covers of A_n and S_n, respectively. We prove that $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{A}}_n$ if and only if $G \cong \widehat{\mathsf{A}}_n$, and $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{S}}_n^{+} \cong \mathbb{C}\widehat{\mathsf{S}}_n^{-}$ if and only if $G \cong \widehat{\mathsf{S}}_n^{+}$ or $\widehat{\mathsf{S}}_n^{-}$. This in particular completes the proof of a conjecture proposed by the second and fourth authors that every finite quasisimple group is determined uniquely up to isomorphism by the structure of its complex group algebra. The known results on prime power degrees and relatively small degrees of irreducible (linear and projective) representations of the symmetric and alternating groups together with the classification of finite simple groups play an essential role in the proofs.

1. Introduction

The complex group algebra of a finite group G, denoted by $\mathbb{C}G$, is the set of formal sums $\left\{\sum_{g \in G} a_g g \mid a_g \in \mathbb{C}\right\}$, equipped with natural rules for addition, multiplication, and scalar multiplication. Wedderburn's theorem implies that $\mathbb{C}G$ is isomorphic to the direct sum of matrix algebras over $\mathbb{C}$ whose dimensions are exactly the degrees of the (nonisomorphic) irreducible complex representations of G. Therefore, the study of complex group algebras and the relation to their base groups is important in group representation theory.

In an attempt to understand the connection between the structure of a finite group and its complex group algebra, in Question 2 of the landmark paper [Brauer 1963]

Nguyen was partially supported by the NSA Young Investigator grant H98230-14-1-0293 and a Faculty Scholarship Award from Buchtel College of Arts and Sciences at The University of Akron. Tong-Viet's work is based on the research supported in part by the National Research Foundation of South Africa (grant 93408).

MSC2010: primary 20C30; secondary 20C15, 20C33.

Keywords: symmetric groups, alternating groups, complex group algebras, Schur covers, double covers, irreducible representations, character degrees.

it was asked: *when do nonisomorphic groups have isomorphic complex group algebras*? Since this question might be too general to be solved completely, it is more feasible to study more explicit questions/problems whose solutions will provide a partial answer to Brauer's question. For instance, if two finite groups have isomorphic complex group algebras and one of them is solvable, is it true that the other is also solvable? Or, if two finite groups have isomorphic complex group algebras and one of them has a normal Sylow p-subgroup, can we conclude the same for the other group? We refer the reader to [Brauer 1963] or Section 9 of the survey paper [Navarro 2010] for more discussions on complex group algebras.

A natural problem that arises from Brauer's question is the following: given a finite group G, determine all finite groups (up to isomorphism) with complex group algebras isomorphic to that of G. This problem is easy for abelian groups but difficult for solvable groups in general. If G is any finite abelian group of order n, then $\mathbb{C}G$ is isomorphic to a direct sum of n copies of $\mathbb{C}$, so the complex group algebras of any two abelian groups are isomorphic if and only if the two groups have the same order. For solvable groups, the probability that two groups have isomorphic complex group algebra is often fairly "high". For instance, it was pointed out in [Huppert 2000] that among 2328 groups of order 2^7, there are only 30 different complex group algebra structures. In contrast to solvable groups, simple groups or more generally quasisimple groups seem to have a stronger connection to their complex group algebras. In [Nguyen and Tong-Viet 2014], two of the four current authors have conjectured that every finite quasisimple group is determined uniquely up to isomorphism by its complex group algebra, and proved it for all quasisimple groups except the nontrivial perfect central covers of the alternating groups.

Let A_n denote the alternating group of degree n. (Throughout the paper we always assume that $n \geq 5$, unless otherwise stated.) The *Schur covers* (or *covering groups*) of the alternating and symmetric groups were first studied and classified in [Schur 1911] in connection with their projective representations. It is known that A_n has one isomorphism class of Schur covers, which is indeed the double cover $\widehat{\mathsf{A}}_n$ except when n is 6 or 7, where triple and 6-fold covers also exist. We are able to prove that every double cover of an alternating group of degree at least 5 is determined uniquely by its complex group algebra.

Theorem A. *Let $n \geq 5$. Let G be a finite group and $\widehat{\mathsf{A}}_n$ the double cover of A_n. Then $G \cong \widehat{\mathsf{A}}_n$ if and only if $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{A}}_n$.*

We prove a similar result for the triple and 6-fold (perfect central) covers of A_6 and A_7, and therefore complete the proof of the aforementioned conjecture.

Theorem B. *Let G be a finite group and H a quasisimple group. Then $G \cong H$ if and only if $\mathbb{C}G \cong \mathbb{C}H$.*

Proof. This is a consequence of Theorem A, Theorem 6.2, and [Nguyen and Tong-Viet 2014, Corollary 1.4]. ☐

The symmetric group S_n has two isomorphism classes of Schur double covers, denoted by $\hat{\mathsf{S}}_n^-$ and $\hat{\mathsf{S}}_n^+$. It turns out that these two covers are isoclinic, and therefore their complex group algebras $\mathbb{C}\hat{\mathsf{S}}_n^+$ and $\mathbb{C}\hat{\mathsf{S}}_n^-$ are isomorphic [Morris 1962]. Our next result solves the above problem for the double covers of the symmetric groups.

Theorem C. *Let $n \geq 5$. Let G be a finite group and $\hat{\mathsf{S}}_n^\pm$ the double covers of S_n. Then $\mathbb{C}G \cong \mathbb{C}\hat{\mathsf{S}}_n^+$ (or equivalently $\mathbb{C}G \cong \mathbb{C}\hat{\mathsf{S}}_n^-$) if and only if $G \cong \hat{\mathsf{S}}_n^+$ or $G \cong \hat{\mathsf{S}}_n^-$.*

Let $\mathrm{Irr}(G)$ denote the set of all irreducible representations (or characters) of a group G over the complex field. As mentioned above, two finite groups have isomorphic complex group algebras if and only if they have the same set of degrees (counting multiplicities) of irreducible characters. Therefore, the proofs of our main results as expected depend heavily on the representation theories of the symmetric and alternating groups, their double covers, and quasisimple groups in general. In particular, we make use of known results on relatively small degrees and prime power degrees of the irreducible characters of $\hat{\mathsf{A}}_n$ and $\hat{\mathsf{S}}_n^\pm$.

The remainder of the paper is organized as follows. In the next section, we give a brief overview of the representation theory of the symmetric and alternating groups and their double covers, and then collect some results on prime power character degrees of these groups. The results on minimal degrees are then presented in Section 3. In Section 4, we establish some useful lemmas that will be needed later in the proofs of the main results. The proof of Theorem A is carried out in Section 5 and exceptional covers of A_6 and A_7 are treated in Section 6.

The last four sections are devoted to the proof of Theorem C. Let G be a finite group such that $\mathbb{C}G \cong \mathbb{C}\hat{\mathsf{S}}_n^\pm$. We will show that $G' = G''$, and therefore there exists a normal subgroup M of G such that $M \subseteq G'$ and the chief factor G'/M is isomorphic to a direct product of k copies of a nonabelian simple group S. To prove that G is isomorphic to one of $\hat{\mathsf{S}}_n^\pm$, one of the key steps is to show that this chief factor is isomorphic to A_n. We will do this by using the classification of finite simple groups to eliminate almost all possibilities for k and S. As we will see, it turns out that the case of simple groups of Lie type in even characteristic is most difficult.

In Section 7, we prove a nonexistence result for particular character degrees of $\hat{\mathsf{S}}_n^\pm$, and we apply it in Section 8 to show that S cannot be a simple group of Lie type in even characteristic. We then eliminate other possibilities for S in Section 9, and complete the proof of Theorem C in Section 10.

Notation. Since $\mathbb{C}\hat{\mathsf{S}}_n^+ \cong \mathbb{C}\hat{\mathsf{S}}_n^-$, when working with character degrees of $\hat{\mathsf{S}}_n^\pm$, it suffices to consider just one of the two covers. For the sake of convenience, we will write $\hat{\mathsf{S}}_n$ to denote either one of the two double covers of S_n. If X and Y are two multisets, we write $X \subseteq Y$, and say that X is a submultiset of Y, if the multiplicity of any element in X does not exceed that of the same element in Y. For a finite group G, the number of conjugacy classes of G is denoted by $k(G)$. We write $\mathrm{Irr}_{2'}(G)$ to mean the set of all irreducible characters of G of odd degree. The set and the multiset of character degrees of G are denoted respectively by $\mathrm{cd}(G)$ and $\mathrm{cd}^*(G)$. Finally, we denote by $d_i(G)$ the i-th smallest nontrivial character degree of G. Other notation is standard or will be defined when needed.

2. Prime power degrees of $\widehat{\mathsf{S}}_n$ and $\widehat{\mathsf{A}}_n$

In this section, we collect some results on irreducible characters of prime power degree of $\widehat{\mathsf{S}}_n$ and $\widehat{\mathsf{A}}_n$. The irreducible characters of the double covers of the symmetric and alternating groups are divided into two kinds: faithful characters, which are also known as spin characters, and nonfaithful characters, which can be viewed as ordinary characters of S_n or A_n.

2A. ***Characters of S_n and A_n.*** To begin with, let us recall some notation and terminology of partitions and Young diagrams in connection with representation theory of the symmetric and alternating groups. A partition λ of n is a finite sequence of natural numbers $(\lambda_1, \lambda_2, \dots, \lambda_m)$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ and $\lambda_1 + \lambda_2 + \cdots + \lambda_m = n$. If $\lambda_1 > \lambda_2 > \cdots > \lambda_m$, we say that λ is a *bar partition* of n (also called a *strict partition* of n). The Young diagram associated to λ is an array of n nodes with λ_i nodes on the i-th row. At each node (i, j), we define the *hook length* $h(i, j)$ to be the number of nodes to the right and below the node (i, j), including the node (i, j).

It is well known that the irreducible characters of S_n are in one-to-one correspondence with partitions of n. The degree of the character χ_λ corresponding to λ is given by the *hook-length formula* of Frame, Robinson and Thrall [Frame et al. 1954]:

$$f_\lambda := \chi_\lambda(1) = \frac{n!}{\prod_{i,j} h(i, j)}.$$

Two partitions of n whose Young diagrams transform into each other when reflected about the line $y = -x$, with the coordinates of the upper-left node taken to be $(0, 0)$, are called conjugate partitions. The partition conjugate to λ is denoted by $\bar{\lambda}$. If $\lambda = \bar{\lambda}$, we say that λ is self-conjugate. The irreducible characters of A_n can be obtained by restricting those of S_n to A_n. More explicitly, $\chi_\lambda \downarrow_{\mathsf{A}_n} = \chi_{\bar{\lambda}} \downarrow_{\mathsf{A}_n}$ is irreducible if λ is not self-conjugate. Otherwise, $\chi_\lambda \downarrow_{\mathsf{A}_n}$ is the sum of two different irreducible characters of A_n of the same degree. In short, the degrees of irreducible characters of A_n are labeled by partitions of n and are given by

$$\tilde{f}_\lambda = \begin{cases} f_\lambda & \text{if } \lambda \neq \bar{\lambda}, \\ f_\lambda/2 & \text{if } \lambda = \bar{\lambda}. \end{cases}$$

Irreducible representations of prime power degree of the symmetric and alternating groups were classified by A. Balog, C. Bessenrodt, J. B. Olsson and K. Ono [Balog et al. 2001]. This result is critical in eliminating simple groups other than A_n involved in the structure of finite groups whose complex group algebras are isomorphic to $\mathbb{C}\widehat{\mathsf{A}}_n$ or $\mathbb{C}\widehat{\mathsf{S}}_n$.

Lemma 2.1 [Balog et al. 2001, Theorem 2.4]. *Let $n \geq 5$. An irreducible character $\chi_\lambda \in \mathrm{Irr}(\mathsf{S}_n)$ corresponding to a partition λ of n has prime power degree $f_\lambda = p^r > 1$ if and only if one of the following occurs*:

(1) $n = p^r + 1$, $\lambda = (n-1, 1)$ *or* $(2, 1^{n-2})$, *and* $f_\lambda = n - 1$.

(2) $n = 5$, $\lambda = (2^2, 1)$ *or* $(3, 2)$, *and* $f_\lambda = 5$.

(3) $n = 6$, $\lambda = (4, 2)$ *or* $(2^2, 1^2)$, *and* $f_\lambda = 9$; $\lambda = (3^2)$ *or* (2^3), *and* $f_\lambda = 5$; $\lambda = (3, 2, 1)$ *and* $f_\lambda = 16$.

(4) $n = 8$, $\lambda = (5, 2, 1)$ *or* $(3, 2, 1^3)$, *and* $f_\lambda = 64$.

(5) $n = 9$, $\lambda = (7, 2)$ *or* $(2^2, 1^5)$, *and* $f_\lambda = 27$.

Lemma 2.2 [Balog et al. 2001, Theorem 5.1]. *Let $n \geq 5$. An irreducible character degree $\tilde{f}_\lambda$ of A_n corresponding to a partition λ of n is a prime power $p^r > 1$ if and only if one of the following occurs*:

(1) $n = p^r + 1$, $\lambda = (n-1, 1)$ *or* $(2, 1^{n-2})$, *and* $\tilde{f}_\lambda = n - 1$.

(2) $n = 5$, $\lambda = (2^2, 1)$ *or* $(3, 2)$, *and* $\tilde{f}_\lambda = 5$; $\lambda = (3, 1^2)$ *and* $\tilde{f}_\lambda = 3$.

(3) $n = 6$, $\lambda = (4, 2)$ *or* $(2^2, 1^2)$ *and* $\tilde{f}_\lambda = 9$; $\lambda = (3^2)$ *or* (2^3) *and* $\tilde{f}_\lambda = 5$; $\lambda = (3, 2, 1)$ *and* $\tilde{f}_\lambda = 8$.

(4) $n = 8$, $\lambda = (5, 2, 1)$ *or* $(3, 2, 1^3)$, *and* $\tilde{f}_\lambda = 64$.

(5) $n = 9$, $\lambda = (7, 2)$ *or* $(2^2, 1^5)$, *and* $\tilde{f}_\lambda = 27$.

2B. ***Spin characters of S_n and A_n.*** We now recall the spin representation theory of the symmetric and alternating groups, due to Schur [Hoffman and Humphreys 1992; Morris 1962; Schur 1911; Wagner 1977]. To each bar partition $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ (i.e., $\mu_1 > \mu_2 > \cdots > \mu_m$) of n, there corresponds one or two irreducible characters (depending on whether $n - m$ is even or odd, respectively) of $\widehat{\mathsf{S}}_n$, of degree

$$g_\mu = 2^{\lfloor (n-m)/2 \rfloor} \bar{g}_\mu,$$

where $\bar{g}_\mu$ denotes the number of shifted standard tableaux of shape μ. This number can be computed by an analogue of the hook-length formula, the *bar formula* [Hoffman and Humphreys 1992, Proposition 10.6]. The length $b(i, j)$ of the (i, j)-bar of μ is the length of the $(i, j+1)$-hook in the shift-symmetric diagram of μ

(obtained by reflecting the shifted diagram of μ along the diagonal and pasting it onto μ; see [Macdonald 1995, p. 14] for details). Then

$$\bar{g}_\mu = \frac{n!}{\prod_{i,j} b(i,j)}.$$

The spin character degree may also be computed by the formula

$$g_\mu = 2^{\lfloor (n-m)/2 \rfloor} \frac{n!}{\mu_1!\mu_2!\cdots\mu_m!} \prod_{i<j} \frac{\mu_i - \mu_j}{\mu_i + \mu_j}.$$

Again, one can get faithful irreducible characters of $\widehat{\mathsf{A}}_n$ by restricting those of $\widehat{\mathsf{S}}_n^{\pm}$ to $\widehat{\mathsf{A}}_n$ in the following way. If $n-m$ is odd, then the restrictions of the two characters of $\widehat{\mathsf{S}}_n^{\pm}$ labeled by μ to $\widehat{\mathsf{A}}_n$ are the same and irreducible. Otherwise, the restriction of the one character labeled by μ is the sum of two irreducible characters of the same degree $g_\mu/2$. Let $\tilde{g}_\mu$ be the degree of the irreducible spin character(s) of $\widehat{\mathsf{A}}_n$ labeled by the bar partition μ; we then have

$$\tilde{g}_\mu = \begin{cases} g_\mu & \text{if } n-m \text{ is odd,} \\ g_\mu/2 & \text{if } n-m \text{ is even.} \end{cases}$$

The classification of spin representations of prime power degree of the symmetric and alternating groups has been done by the first and third authors of the current paper in [Bessenrodt and Olsson 2002].

Lemma 2.3 [Bessenrodt and Olsson 2002, Theorem 4.2]. *Let $n \geq 5$, and let μ be a bar partition of n. The spin irreducible character degree g_μ of $\widehat{\mathsf{S}}_n$ corresponding to μ is a prime power if and only if one of the following occurs*:

(1) $\mu = (n)$ *and* $g_\mu = 2^{\lfloor (n-1)/2 \rfloor}$.

(2) $n = 2^r + 2$ *for some* $r \in \mathbb{N}$, $\mu = (n-1, 1)$, *and* $g_\mu = 2^{2^{r-1}+r}$.

(3) $n = 5$, $\mu = (3, 2)$, *and* $g_\mu = 4$.

(4) $n = 6$, $\mu = (3, 2, 1)$, *and* $g_\mu = 4$.

(5) $n = 8$, $\mu = (5, 2, 1)$, *and* $g_\mu = 64$.

Lemma 2.4 [Bessenrodt and Olsson 2002, Theorem 4.3]. *Let $n \geq 5$, and let μ be a bar partition of n. The spin irreducible character degree $\tilde{g}_\mu$ of $\widehat{\mathsf{A}}_n$ corresponding to μ is a prime power if and only if one of the following occurs*:

(1) $\mu = (n)$ *and* $\tilde{g}_\mu = 2^{\lfloor (n-2)/2 \rfloor}$.

(2) $n = 2^r + 2$ *for some* $r \in \mathbb{N}$, $\mu = (n-1, 1)$, *and* $\tilde{g}_\mu = 2^{2^{r-1}+r-1}$.

(3) $n = 5$, $\mu = (3, 2)$, *and* $\tilde{g}_\mu = 4$.

(4) $n = 6$, $\mu = (3, 2, 1)$, *and* $\tilde{g}_\mu = 4$.

(5) $n = 8$, $\mu = (5, 2, 1)$, *and* $\tilde{g}_\mu = 64$.

3. Low degrees of $\widehat{\mathsf{S}}_n$ and $\widehat{\mathsf{A}}_n$

We present in this section some results on minimal degrees of both ordinary and spin characters of the symmetric and alternating groups. We start with ordinary characters:

Lemma 3.1 [Rasala 1977]. *The following hold*:

(1) $d_1(\mathsf{S}_n) = n-1$ *if* $n \geq 5$.

(2) $d_2(\mathsf{S}_n) = n(n-3)/2$ *if* $n \geq 9$.

(3) $d_3(\mathsf{S}_n) = (n-1)(n-2)/2$ *if* $n \geq 9$.

(4) $d_4(\mathsf{S}_n) = n(n-1)(n-5)/6$ *if* $n \geq 13$.

(5) $d_5(\mathsf{S}_n) = (n-1)(n-2)(n-3)/6$ *if* $n \geq 13$.

(6) $d_6(\mathsf{S}_n) = n(n-2)(n-4)/3$ *if* $n \geq 15$.

(7) $d_7(\mathsf{S}_n) = n(n-1)(n-2)(n-7)/24$ *if* $n \geq 15$.

Lemma 3.2 [Tong-Viet 2011]. *If* $n \geq 15$, *then* $d_i(\mathsf{A}_n) = d_i(\mathsf{S}_n)$ *for* $1 \leq i \leq 4$, *and*, *if* $n \geq 22$, *then* $d_i(\mathsf{A}_n) = d_i(\mathsf{S}_n)$ *for* $1 \leq i \leq 7$.

The minimal degrees of spin irreducible representations of $\widehat{\mathsf{A}}_n$ and $\widehat{\mathsf{S}}_n$ were obtained in [Kleshchev and Tiep 2004; 2012]. These minimal degrees are indeed the degrees of the *basic spin* and *second basic spin* representations. Let $\mathfrak{d}_1(\widehat{\mathsf{A}}_n)$ and $\mathfrak{d}_1(\widehat{\mathsf{S}}_n)$ denote the smallest degrees of irreducible spin characters of $\widehat{\mathsf{A}}_n$ and $\widehat{\mathsf{S}}_n$, respectively.

Lemma 3.3 [Kleshchev and Tiep 2004, Theorem A]. *Let* $n \geq 8$. *The smallest degrees of the irreducible spin characters of* $\widehat{\mathsf{A}}_n$ *and* $\widehat{\mathsf{S}}_n$ *are* $\mathfrak{d}_1(\widehat{\mathsf{A}}_n) = 2^{\lfloor (n-2)/2 \rfloor}$ *and* $\mathfrak{d}_1(\widehat{\mathsf{S}}_n) = 2^{\lfloor (n-1)/2 \rfloor}$ *respectively*, *and there is no degree between* $\mathfrak{d}_1(\widehat{\mathsf{A}}_n)$ *and* $2\mathfrak{d}_1(\widehat{\mathsf{A}}_n)$ *in the alternating case or* $\mathfrak{d}_1(\widehat{\mathsf{S}}_n)$ *and* $2\mathfrak{d}_1(\widehat{\mathsf{S}}_n)$ *in the symmetric one.*

Using the above results, we easily deduce the following:

Lemma 3.4. *The following hold*:

(1) *If* $n \geq 31$, *then* $d_i(\widehat{\mathsf{S}}_n) = d_i(\mathsf{S}_n)$ *for* $1 \leq i \leq 7$.

(2) *If* $n \geq 34$, *then* $d_i(\widehat{\mathsf{A}}_n) = d_i(\mathsf{A}_n)$ *for* $1 \leq i \leq 7$.

Proof. We observe that $\mathfrak{d}_1(\widehat{\mathsf{S}}_n) = 2^{\lfloor (n-1)/2 \rfloor} > n(n-1)(n-2)(n-7)/24 = d_7(\mathsf{S}_n)$ if $n \geq 31$. Therefore part (1) follows by Lemmas 3.1 and 3.3. Similarly, we have $\mathfrak{d}_1(\widehat{\mathsf{A}}_n) = 2^{\lfloor (n-2)/2 \rfloor} > n(n-1)(n-2)(n-7)/24 = d_7(\mathsf{A}_n)$ if $n \geq 34$, and thus part (2) follows. $\square$

Lemma 3.5. *Let* G *be either* $\widehat{\mathsf{A}}_n$ *or* $\widehat{\mathsf{S}}_n$.

(1) *If* $n \geq 8$, *then* $d_1(G) = n-1$.

(2) *If* $n \geq 10$, *then* $d_2(G) = \min\{n(n-3)/2, \mathfrak{d}_1(G)\}$. *Furthermore*, *if* $n \geq 12$, *then* $d_2(G) > 2n$.

(3) *If* $n \geq 16$, *then* $d_3(G) = (n-1)(n-2)/2$ *and* $d_4(G) = \min\{n(n-1)(n-5)/6, \mathfrak{d}_1(G)\}$.

(4) *If* $n \geq 28$, *then* $d_4(G) = n(n-1)(n-5)/6$, $d_5(G) = (n-1)(n-2)(n-3)/6$, $d_6(G) = n(n-2)(n-4)/3$, *and* $d_7(G) = \min\{n(n-1)(n-2)(n-7)/24, \mathfrak{d}_1(G)\}$.

Proof. When $n \geq 8$, we see that $\mathfrak{d}_1(\widehat{\mathsf{S}}_n) \geq \mathfrak{d}_1(\widehat{\mathsf{A}}_n) = 2^{\lfloor (n-2)/2 \rfloor} > n-1$ and $d_1(\mathsf{S}_n) = d_1(\mathsf{A}_n) = n-1$, which implies that $d_1(\widehat{\mathsf{A}}_n) = d_1(\widehat{\mathsf{S}}_n) = n-1$, and part (1) follows. When $n \geq 10$, we observe that $d_2(\mathsf{A}_n) = d_2(\mathsf{S}_n) = n(n-3)/2$, and part (2) then follows from part (1).

Suppose that $n \geq 16$. It is easy to check that $2^{\lfloor (n-2)/2 \rfloor} > (n-1)(n-2)/2 = d_3(\mathsf{A}_n) = d_3(\mathsf{S}_n)$. It follows that $d_3(\widehat{\mathsf{A}}_n) = d_3(\widehat{\mathsf{S}}_n) = (n-1)(n-2)/2$ and $d_4(G) = \min\{n(n-1)(n-5)/6, \mathfrak{d}_1(G)\}$, as claimed.

Finally suppose that $n \geq 28$. We check that $2^{\lfloor (n-2)/2 \rfloor} > n(n-2)(n-4)/3 = d_6(\mathsf{A}_n) = d_6(\mathsf{S}_n)$, and part (4) then follows. □

4. Some useful lemmas

We begin with an easy observation:

Lemma 4.1. *There always exists a prime number p with $n < p \leq d_2(G)$ for $G = \widehat{\mathsf{A}}_n$ or $\widehat{\mathsf{S}}_n$, provided that $n \geq 9$.*

Proof. It is routine to check the statement for $9 \leq n \leq 11$ by using [Conway et al. 1985]. So we can assume that $n \geq 12$. By Lemma 3.5, it suffices to prove that there exists a prime between n and $2n$. However, this is the well-known Bertrand–Chebyshev theorem [Harborth and Kemnitz 1981]. □

The next three lemmas are critical in the proof of Theorem A.

Lemma 4.2. *Let S be a simple group of Lie type of rank l defined over a field of q elements, with q even. Then*

$$|\mathrm{Irr}_{2'}(S)| \geq \begin{cases} q^l/(l+1, q-1) & \text{if } S \text{ is of type } A, \\ q^l/(l+1, q+1) & \text{if } S \text{ is of type } {}^2A, \\ q^l/3 & \text{if } S \text{ is of type } E_6, {}^2E_6, \\ q^l & \text{otherwise.} \end{cases}$$

Proof. Let S_{sc} be the finite Lie-type group of simply connected type corresponding to S. By [Brunat 2010, Corollary 3.6], S_{sc} has q^l semisimple conjugacy classes. To each semisimple class s of S_{sc}, Lusztig's classification of complex characters of finite groups of Lie type says that there corresponds a semisimple character of the dual group, say S^*_{sc}, of S_{sc} of degree

$$|S_{\mathrm{sc}}|_{2'}/|\boldsymbol{C}_{S_{\mathrm{sc}}}(s)|_{2'}.$$

This means that the dual S_{sc}^* of S_{sc} has at least q^l irreducible characters of odd degree.

If S is of type A, we have $S_{\mathrm{sc}}^* = \mathrm{PGL}_{l+1}(q) = S.(l+1, q-1)$, and the lemma follows for linear groups. A similar argument works for unitary groups and E_6 as well as ${}^2E_6$. If S is not of these types, we will have $S = S_{\mathrm{sc}} = S_{\mathrm{sc}}^*$ and the lemma also follows. $\square$

Lemma 4.3. *Let $n \geq 5$, and let $n = 2^{k_1} + 2^{k_2} + \cdots + 2^{k_t}$ be the binary expansion of n, with $k_1 > k_2 > \cdots > k_t \geq 0$. Then*

$$|\mathrm{Irr}_{2'}(\widehat{\mathsf{A}}_n)| \leq |\mathrm{Irr}_{2'}(\widehat{\mathsf{S}}_n)| = 2^{k_1+k_2+\cdots+k_t}.$$

Proof. If $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ is a bar partition of n, then the 2-part of the spin character degree g_μ of $\widehat{\mathsf{S}}_n$ labeled by μ is at least

$$2^{\lfloor (n-m)/2 \rfloor},$$

which is at least 2 as $n \geq 5$ and $n - m \geq 3$. We note that if $n - m = 3$ then $\tilde{g}_\mu = g_\mu$. Therefore, the 2-part of the degree $\tilde{g}_\mu$ of $\widehat{\mathsf{A}}_n$ is at least 2 as well. In particular, we see that every spin character degree of $\widehat{\mathsf{A}}_n$ as well as $\widehat{\mathsf{S}}_n$ is even. It follows that

$$|\mathrm{Irr}_{2'}(\widehat{\mathsf{A}}_n)| = |\mathrm{Irr}_{2'}(\mathsf{A}_n)| \quad \text{and} \quad |\mathrm{Irr}_{2'}(\widehat{\mathsf{S}}_n)| = |\mathrm{Irr}_{2'}(\mathsf{S}_n)|.$$

As mentioned in [McKay 1972], the number of odd degree irreducible characters of A_n does not exceed that of S_n. Now the lemma follows from the formula for the number of odd degree characters of S_n given in [Macdonald 1971, Corollary 1.3]. $\square$

Lemma 4.4. *Let $n = 2^{k_1} + 2^{k_2} + \cdots + 2^{k_t}$ be the binary expansion of n, with $k_1 > k_2 > \cdots > k_t \geq 0$.*

(1) *If $k_1 + k_2 + \cdots + k_t \geq \sqrt{(n-3)/2}$, then $n < 2^{15}$.*

(2) *if $k_1 + k_2 + \cdots + k_t \geq \sqrt{n-3} - 3$, then $n < 2^{13}$.*

(3) *if $k_1 + k_2 + \cdots + k_t \geq (n-3)/18$, then $n < 2^{10}$.*

(4) *if $k_1 + k_2 + \cdots + k_t \geq (n-3)/30$, then $n < 2^{11}$.*

Proof. We only give here a proof for part (2). The other statements are proved similarly. As $n = 2^{k_1} + 2^{k_2} + \cdots + 2^{k_t}$, we get $k_1 = \lfloor \log_2 n \rfloor$ and hence

$$k_1 + k_2 + \cdots + k_t \geq \lfloor \log_2 n \rfloor (\lfloor \log_2 n \rfloor + 1)/2.$$

However, it is easy to check that $\sqrt{(n-3)} - 3 > \lfloor \log_2 n \rfloor (\lfloor \log_2 n \rfloor + 1)/2$ if $n \geq 2^{14}$. For $2^{13} \leq n < 2^{14}$, the statement follows by direct computation. $\square$

The following lemma is probably known, but we include a short proof for the reader's convenience. It will be needed in the proof of Theorem C.

Lemma 4.5. *Let $k(\widehat{\mathsf{A}}_n)$ and $k(\widehat{\mathsf{S}}_n)$ denote the number of conjugacy classes of $\widehat{\mathsf{A}}_n$ and $\widehat{\mathsf{S}}_n$, respectively. Then*

$$k(\widehat{\mathsf{S}}_n) < 2k(\widehat{\mathsf{A}}_n).$$

Proof. Let $\mathrm{Irr}_{\mathrm{faithful}}(G)$ and $\mathrm{Irr}_{\mathrm{nonfaithful}}(G)$ denote the sets of faithful irreducible characters and nonfaithful irreducible characters, respectively, of a group G. Let a and b be the numbers of self-conjugate partitions and of pairs of nonself-conjugate partitions, respectively, of n. Also, let c and d be the numbers of bar partitions of n with $n-m$ even and odd, respectively. We have

$$|\mathrm{Irr}_{\mathrm{nonfaithful}}(\widehat{\mathsf{S}}_n)| = k(\mathsf{S}_n) = a+2b, \quad |\mathrm{Irr}_{\mathrm{faithful}}(\widehat{\mathsf{S}}_n)| = c+2d,$$
$$|\mathrm{Irr}_{\mathrm{nonfaithful}}(\widehat{\mathsf{A}}_n)| = k(\mathsf{A}_n) = 2a+b, \quad |\mathrm{Irr}_{\mathrm{faithful}}(\widehat{\mathsf{A}}_n)| = 2c+d.$$

Therefore,

$$k(\widehat{\mathsf{S}}_n) = a+2b+c+2d \quad \text{and} \quad k(\widehat{\mathsf{A}}_n) = 2a+b+2c+d,$$

and the lemma follows. □

5. Complex group algebra of $\widehat{\mathsf{A}}_n$ — Theorem A

The aim of this section is to prove Theorem A. Let G be a finite group such that $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{A}}_n$. Then G has exactly one linear character, which is the trivial one, so that G is perfect. Let M be a maximal normal subgroup of G. We then have that G/M is nonabelian simple, and moreover

$$\mathrm{cd}^*(G/M) \subseteq \mathrm{cd}^*(G) = \mathrm{cd}^*(\widehat{\mathsf{A}}_n).$$

To prove the theorem, it is clear that we first have to show $G/M \cong \mathsf{A}_n$. We will work towards this aim.

Proposition 5.1. *Let S be a nonabelian simple group such that $\mathrm{cd}^*(S) \subseteq \mathrm{cd}^*(\widehat{\mathsf{A}}_n)$. Then S is isomorphic to A_n or to a simple group of Lie type in even characteristic.*

Proof. We will eliminate other possibilities for S by using the classification of finite simple groups. If $5 \le n \le 9$, then the set of prime divisors of S is contained in that of $\widehat{\mathsf{A}}_n$, which in turn is contained in $\{2,3,5,7\}$; hence by using [Huppert and Lempken 2000, Theorem III] and [Conway et al. 1985], the result follows easily. From now on we assume that $n \ge 10$.

(i) *Alternating groups*: Suppose that $S = \mathsf{A}_m$ with $5 \le m \ne n$. Since $\mathrm{cd}(A_m) \subseteq \mathrm{cd}(\widehat{\mathsf{A}}_n)$, we get $d_1(\mathsf{A}_m) \ge d_1(\widehat{\mathsf{A}}_n)$. As $d_1(\widehat{\mathsf{A}}_n) = n-1 \ge 9$ by Lemma 3.5, it follows that $d_1(\mathsf{A}_m) \ge 9$. Thus $m \ge 10$, and so $m-1 = d_1(\mathsf{A}_m) \ge d_1(\widehat{\mathsf{A}}_n) = n-1$. In particular, we have $m > n$ as $m \ne n$. It follows that $|S| > 2|\mathsf{A}_n|$ and this violates the hypothesis that $\mathrm{cd}^*(S) \subseteq \mathrm{cd}^*(\widehat{\mathsf{A}}_n)$.

(ii) *Simple groups of Lie type in odd characteristic*: Suppose that $S = G(p^k)$, a simple group of Lie type defined over a field of p^k elements with p odd. Since $|S|_p$ is the degree of the Steinberg character of S, we have $|S|_p \in \mathrm{cd}(\widehat{\mathsf{A}}_n)$. As $|S|_p$ is an odd prime power, Lemma 2.4 implies that $|S|_p$ must be the degree of a nonfaithful character of $\widehat{\mathsf{A}}_n$. In other words, $|S|_p \in \mathrm{cd}(\mathsf{A}_n)$. Using Lemma 2.2, we deduce that $|S|_p = n-1$. Hence, $|S|_p = d_1(\widehat{\mathsf{A}}_n)$ is the smallest nontrivial degree of $\widehat{\mathsf{A}}_n$ by Lemma 3.5. However, by [Tong-Viet 2012, Lemma 8] we have $d_1(S) < |S|_p = d_1(\widehat{\mathsf{A}}_n)$, which is impossible as $\mathrm{cd}(S) \subseteq \mathrm{cd}(\widehat{\mathsf{A}}_n)$.

(iii) *Sporadic simple groups and the Tits group*: Using GAP (version 4.4.12), we can assume that $n \geq 14$. To eliminate these cases, observe that $n \geq \max\{p(S), 14\}$, where $p(S)$ is the largest prime divisor of $|S|$, and that $d_i(S) \geq d_i(\widehat{\mathsf{A}}_n)$ for all $i \geq 1$. With this lower bound on n, we find the lower bounds for $d_i(\widehat{\mathsf{A}}_n)$ with $1 \leq i \leq 7$ using Lemmas 3.4 and 3.5. Choose $i \in \{2, 3, \dots, 7\}$ such that $d_i(\widehat{\mathsf{A}}_n) > d_j(S)$ for some $j \geq 1$ such that $|i-j|$ is minimal. If $j \geq i$, then we obtain a contradiction. If $j < i$, then $d_j(S) \in \{d_k(\widehat{\mathsf{A}}_n)\}_{k=j}^{i-1}$. Solving these equations for n, we then obtain that either these equations have no solution, or that, for each solution of n, we can find some $k \geq 1$ with $d_k(\widehat{\mathsf{A}}_n) > d_k(S)$. As an example, assume that $S = \mathrm{O'N}$. Then $n \geq 31$ since $p(S) = 31$. We have $d_7(\widehat{\mathsf{A}}_n) = n(n-1)(n-2)(n-7)/24 \geq 26970$. As $d_4(S) = 26752 < d_7(\widehat{\mathsf{A}}_n)$, it follows that $d_4(S) \in \{d_4(\widehat{\mathsf{A}}_n), d_5(\widehat{\mathsf{A}}_n), d_6(\widehat{\mathsf{A}}_n)\}$. However, one can check that these equations have no integer solutions. □

Proposition 5.2. *Let S be a nonabelian simple group such that $|S| \mid n!$ and $\mathrm{cd}^*(S) \subseteq \mathrm{cd}^*(\widehat{\mathsf{A}}_n)$. Then $S \cong \mathsf{A}_n$.*

Proof. In light of Proposition 5.1 and its proof, it remains to assume that $n \geq 9$ and prove that S cannot be a simple group of Lie type in even characteristic. Assume to the contrary that $S = G_l(2^k)$, a simple group of Lie type of rank l defined over a field of $q = 2^k$ elements. As above, we then have $|S|_2 \in \mathrm{cd}(\widehat{\mathsf{A}}_n)$. By Lemmas 2.2 and 2.4, we have that $|S|_2 = n-1$, $|S|_2 = 2^{\lfloor (n-2)/2 \rfloor}$, or $|S|_2 = 2^{n/2+\log_2(n-2)-2}$ when $n = 2^r + 2$. Since the case $|S|_2 = n-1$ can be eliminated as in the proof of the previous proposition, we can assume further that

$$|S|_2 = 2^{\lfloor (n-2)/2 \rfloor} \quad \text{or} \quad |S|_2 = 2^{n/2+\log_2(n-2)-2} \quad \text{when } n = 2^r + 2.$$

Recalling the hypothesis that $\mathrm{cd}^*(S) \subseteq \mathrm{cd}^*(\widehat{\mathsf{A}}_n)$, we have

$$|\mathrm{Irr}_{2'}(S)| \leq |\mathrm{Irr}_{2'}(\widehat{\mathsf{A}}_n)|. \tag{5-1}$$

(i) $S = B_l(2^k) \cong C_l(2^k)$, $D_l(2^k)$, *or* ${}^2D_l(2^k)$. Then $|S|_2 = 2^{kl^2}$ or $2^{kl(l-1)}$. In particular,

$$|S|_2 \leq 2^{kl^2}.$$

As $|S|_2 \geq 2^{\lfloor (n-2)/2 \rfloor}$, it follows that

$$kl^2 \geq \lfloor (n-2)/2 \rfloor \geq (n-3)/2.$$

Therefore

$$kl \geq \sqrt{(n-3)/2}.$$

Using Lemma 4.2, we then obtain

$$|\mathrm{Irr}_{2'}(S)| \geq q^l = 2^{kl} \geq 2^{\sqrt{(n-3)/2}}.$$

Now Lemma 4.3 and the inequality (5-1) imply

$$k_1 + k_2 + \cdots + k_t \geq \sqrt{(n-3)/2},$$

where $n = 2^{k_1} + 2^{k_2} + \cdots + 2^{k_t}$ is the binary expansion of n. Invoking Lemma 4.4(1), we obtain that $n < 2^{15}$.

(ii) $S = A_l(2^k)$ *or* ${}^2A_l(2^k)$. Arguing as above, we have

$$k_1 + k_2 + \cdots + k_t \geq \sqrt{n-3} - 3,$$

which forces $n < 2^{13}$ by Lemma 4.4(2).

(iii) *S is a simple group of exceptional Lie type*. Using Lemma 4.4(3,4), we deduce that $n < 2^{11}$.

For each of the above cases, a computer program has checked that either $|S|$ does not divide $n!$ or S has an irreducible character degree not belonging to $\mathrm{cd}(\widehat{\mathsf{A}}_n)$ for "small" n. This contradiction completes the proof. Let us describe the example where $S = \mathrm{Sp}_{2l}(2^k) \cong \Omega_{2l+1}(2^k)$. Then we have

$$kl^2 = \lfloor (n-2)/2 \rfloor \quad \text{or} \quad kl^2 = n/2 + \log_2(n-2) - 2 \quad \text{when } n = 2^r + 2.$$

Moreover, the condition $|S| \mid n!$ is equivalent to

$$2^{kl^2} \prod_{i=1}^{l} (2^{2ki} - 1) \mid n!\,.$$

By computer calculations, we can determine all triples (k, l, n) with $n < 2^{15}$ satisfying the above conditions. It turns out that, for each such triple, n is at most 170 and one of the three smallest character degrees of S is not a character degree of $\widehat{\mathsf{A}}_n$. The low-degree characters of simple groups of Lie type can be found in [Tiep and Zalesskii 1996; Lübeck 2001; Nguyen 2010]. □

We are now ready to prove the first main result:

Proof of Theorem A. Recall the hypothesis that G is a finite group such that $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{A}}_n$. Therefore $\mathrm{cd}^*(G) = \mathrm{cd}^*(\widehat{\mathsf{A}}_n)$. In particular, we have $|G| = |\widehat{\mathsf{A}}_n|$ and $G = G'$ since $\widehat{\mathsf{A}}_n$ has only one linear character. Let M be a maximal normal subgroup of G. Then G/M is a nonabelian simple group, say S. It follows that $\mathrm{cd}^*(S) = \mathrm{cd}^*(G/M) \subseteq \mathrm{cd}^*(G)$ and hence

$$\mathrm{cd}^*(S) \subseteq \mathrm{cd}^*(\widehat{\mathsf{A}}_n).$$

We also have

$$|S| \mid |G| = |\widehat{\mathsf{A}}_n| = n!\,.$$

Applying Propositions 5.1 and 5.2, we deduce that $S \cong \mathsf{A}_n$.

We have shown that $G/M \cong \mathsf{A}_n$. Since $|G| = |\widehat{\mathsf{A}}_n| = 2|\mathsf{A}_n|$, we obtain $|M| = 2$. In particular, M is central in G and therefore $M \subseteq \boldsymbol{Z}(G) \cap G'$. Thus $G \cong \widehat{\mathsf{A}}_n$, as desired. □

6. Triple and 6-fold covers of A_6 and A_7

In this section, we aim to prove that every perfect central cover of A_6 or A_7 is uniquely determined up to isomorphism by the structure of its complex group algebra. To do that, we need the following result from [Nguyen and Tong-Viet 2014, Lemma 2.5]. Here and in what follows, we write $\mathrm{Mult}(S)$ and $\mathrm{Schur}(S)$ to denote the Schur multiplier and the Schur covering group (or the Schur cover for short), respectively, of a simple group S.

Lemma 6.1. *Let S be a nonabelian simple group different from an alternating group of degree greater than* 13. *Assume that S is different from* $\mathrm{PSL}_3(4)$ *and* $\mathrm{PSU}_4(3)$. *Let G be a perfect group and $M \lhd G$ such that $G/M \cong S$, $|M| \leq |\mathrm{Mult}(S)|$, and* $\mathrm{cd}(G) \subseteq \mathrm{cd}(\mathrm{Schur}(S))$. *Then G is uniquely determined up to isomorphism by S and the order of G.*

Now we prove the main result of this section:

Theorem 6.2. *Let G be a finite group and H a perfect central cover of* A_6 *or A_7. Then $G \cong H$ if and only if $\mathbb{C}G \cong \mathbb{C}H$.*

Proof. First, as $\mathsf{A}_6 \cong \mathrm{PSL}_2(9)$, every perfect central cover of A_6 can be viewed as a quasisimple classical group, a case which is already studied in [Nguyen 2013, Theorem 1.1]. So it remains to consider the perfect central covers of A_7. Let H be one of those, and assume that G is a finite group such that $\mathbb{C}G \cong \mathbb{C}H$.

As before, we see that G is perfect and, if M is a normal maximal subgroup of G, we have that G/M is nonabelian simple and $\mathrm{cd}^*(G/M) \subseteq \mathrm{cd}^*(H)$. In particular,

$$\mathrm{cd}^*(G/M) \subseteq \mathrm{cd}^*(\mathrm{Schur}(\mathsf{A}_7)),$$

where Schur(A_7) is the Schur cover (or the 6-fold cover) of A_7. It follows that $|G/M| \leq 6|\mathsf{A}_7| = 7560$. Inspecting [Conway et al. 1985], we come up with

$$G/M \cong \mathrm{PSL}_2(q) \text{ with } 5 \leq q \leq 23, \text{ or } \mathrm{PSL}_3(3), \text{ or } \mathrm{PSU}_3(3), \text{ or } \mathsf{A}_7.$$

Since each of the possibilities for G/M except A_7 does not satisfy the inclusion $\mathrm{cd}^*(G/M) \subseteq \mathrm{cd}^*(\mathrm{Schur}(\mathsf{A}_7))$, we deduce that $G/M \cong \mathsf{A}_7$.

On the other hand, as $\mathbb{C}G \cong \mathbb{C}H$, we have $|G| = |H|$. It follows that $|M| = |\boldsymbol{Z}(H)| \leq 6$. Using Lemma 6.1, we conclude that $G \cong H$. □

7. Excluding critical character degrees of $\widehat{\mathsf{S}}_n$

In this section, we prove a nonexistence result for special character degrees of $\widehat{\mathsf{S}}_n$ which will be applied in the next section. Indeed, with the following proposition, we prove a little more, as only the case of even n will be needed (in fact, the proof shows that also versions with slightly modified 2-powers can be obtained).

Proposition 7.1. *Let $n \in \mathbb{N}$. If $2^{[\frac{n-2}{2}]}(n-1)$ is a character degree of $\widehat{\mathsf{S}}_n$, then $n \leq 8$ and the degree is an ordinary degree f_λ for $\lambda \in \{(2), (2,1), (4,2^2)\}$ (or their conjugates), or the spin degree g_μ for $\mu = (4,2)$.*

The strategy for the proof is inspired by the methods used in [Balog et al. 2001; Bessenrodt and Olsson 2002] to classify the irreducible characters of prime power degrees. A main ingredient is a number-theoretic result which is a variation of [Balog et al. 2001, Theorem 3.1].

First, we define $M(n)$ to be the set of pairs of finite sequences of integers $s_1 < s_2 < \cdots < s_r \leq n$, $t_1 < t_2 < \cdots < t_r \leq n$, with all numbers different from $n-1$, that satisfy

(i) $s_i < t_i$ for all i,

(ii) s_1 and t_1 are primes $> n/2$, and

(iii) for $1 \leq i \leq r-1$, s_{i+1} and t_{i+1} contain prime factors exceeding $2n - s_i - t_i$ and not dividing $n-1$.

We then set $t(n) := \max\{t_r \mid ((s_i)_{i=1,\dots,r}, (t_i)_{i=1,\dots,r}) \in M(n)\}$, and $t(n) = 0$ when $M(n) = \varnothing$. Note that, for all $n \geq 15$, there are at least two primes $p, q \neq n-1$ with $n/2 < p < q \leq n$ (e.g., use [Harborth and Kemnitz 1981]); hence for all $n \geq 15$ the set $M(n)$ is not empty.

Theorem 7.2. *Let $n \in \mathbb{N}$. Then $n - t(n) \leq 225$.*

For $15 \leq n \leq 10^9$, we have the tighter bounds

$$n - t(n) \begin{cases} = 7 & \text{for } n \in \{30, 54\}, \\ = 5 & \text{for } n \in \{18, 24, 28, 52, 102, 128, 224\}, \\ \leq 4 & \text{otherwise}. \end{cases}$$

Proof. For $n > 3.9 \cdot 10^8$, the proof follows the lines of the arguments for [Balog et al. 2001, Theorem 3.1], noticing that in the construction given there the numbers in the sequences are below $n-1$ and that the chosen prime factors do not divide $n-1$; this then gives $n - t(n) \le 225$.

A computer calculation (with Maple) shows that, for all $n \le 10^9$, we have the claimed bounds and values for $n - t(n)$. □

For any partition λ of n, we denote by $l(\lambda)$ the length of λ, and we let $l_1(\lambda)$ be the multiplicity of 1 in λ. We put $h_i = h(i,1)$ for $1 \le i \le l(\lambda)$; these are the *first-column hook lengths* of λ. We set $\mathrm{fch}(\lambda) = \{h_1, \dots, h_{l(\lambda)}\}$.

First we want to show that Proposition 7.1 holds for ordinary characters. Via computer calculations, the claim is easily checked up to $n = 44$, and in particular, we find the stated exceptions for $n < 9$. Thus we have to show that

$$f_\lambda = 2^{[\frac{n-2}{2}]}(n-1) \tag{$*$}$$

cannot hold for $n \ge 9$; if necessary, we may even assume that $n > 44$.

To employ Theorem 7.2, we need some preparation, which is similar to corresponding results in [Balog et al. 2001].

Proposition 7.3. *If q is a prime with $n - l_1(\lambda) \le q \le n$ and $q \nmid f_\lambda$, then*

$$q,\ 2q,\ \dots,\ \left[\frac{n}{q}\right]q \in \mathrm{fch}(\lambda).$$

Proof. Put $w = [n/q]$, $n = wq + r$, $0 \le r < q$. By assumption, we have $(w-1)q \le (w-1)q + r = n - q \le l_1 := l_1(\lambda)$. Thus $q, 2q, \dots, (w-1)q \in \mathrm{fch}(\lambda)$. If $wq \le l_1$, then we are done. Assume that $l_1 < wq$. At most w hooks in λ are of lengths divisible by q (see, e.g., [Olsson 1993, Proposition (3.6)]). If there are only the above $w-1$ hooks in the first column of length divisible by q, then $q \mid f_\lambda$ since $\prod_{i=1}^{w}(iq) \mid n!$, a contradiction. Let $h(i,j)$ be the additional hook length divisible by q. Since $\lambda \ne (1^n)$, $l_1 \le h_2$. If $h_2 > l_1$, then $h(i,j) + h(2,1) > q + l_1 \ge n$. By [Balog et al. 2001, Corollary 2.8] we get $j = 1$. If $h_2 = l_1$, then $\lambda = (n - l_1, 1^{l_1})$, and since $l_1 < wq$ there has to be a hook of length divisible by q in the first row. Since $n - l_1 \le q$ we must have $h_1 = wq$. □

In analogy to [Balog et al. 2001, Corollary 2.10], we deduce:

Corollary 7.4. *Let $1 \le i < j \le l(\lambda)$. If $h \le n$ has a prime divisor q satisfying $2n - h_i - h_j < q$ and $q \nmid f_\lambda$, then $h \in \mathrm{fch}(\lambda)$.*

We now combine these results with Theorem 7.2, similarly to [Balog et al. 2001]; as stated earlier we may assume that $n \ge 15$, and hence there are least two primes $p, q \ne n-1$ with $n/2 < p < q \le n$. Assuming ($*$) for λ, the hook formula implies that there have to be hooks of length p and q in λ. As argued in [Balog et al. 2001], we then have $p, q \in \mathrm{fch}(\lambda)$ or $p, q \in \mathrm{fch}(\bar{\lambda})$; without loss of generality, we may assume $p, q \in \mathrm{fch}(\lambda)$. Then the assumption $(*)$ forces *any* prime between $n/2$ and n, except $n-1$ if this is prime, to be in $\mathrm{fch}(\lambda)$. This gives an indication towards the connection with the sequences belonging to the pairs in $M(n)$.

Indeed, we have the following proposition, which is proved similarly to the corresponding result in [Balog et al. 2001].

Proposition 7.5. *Let $n \geq 15$. Let $((s_i)_{i=1,\ldots,r}, (t_i)_{i=1,\ldots,r}) \in M(n)$. Let λ be a partition of n such that $(*)$ holds. Then $\{s_1, \ldots, s_r, t_1, \ldots, t_r\} \subset \mathrm{fch}(\lambda)$ or $\{s_1, \ldots, s_r, t_1, \ldots, t_r\} \subset \mathrm{fch}(\bar{\lambda})$.*

In particular, $n-h_1 \leq 225$, and we have tighter bounds for $n-h_1$ when $n \leq 10^9$ as given in Theorem 7.2.

Now we can embark on the first part of the proof of Proposition 7.1, showing the nonexistence of ordinary irreducible characters of the critical degree for $n \geq 9$. As remarked before, we may assume $n \geq 44$.

First part of the proof of Proposition 7.1. Set $m = [(n-2)/2]$, and assume that the partition λ of n satisfies

$$f_\lambda = 2^m(n-1). \tag{$*$}$$

Let $c = n - h_1$. By [Balog et al. 2001, Proposition 4.1] we have the following bound for the 2-part of the degree:

$$(f_\lambda)_2 \leq n^2 \cdot ((2c+2)!)_2.$$

By Proposition 7.5, we have $c \leq 225$, and hence $((2c+2)!)_2 \leq (452!)_2 = 2^{448}$. Thus

$$2^m \leq 2^{448} n^2 \leq 2^{448}(2m+3)^2.$$

A short computation gives $m \leq 467$, and hence $n \leq 937$. By Proposition 7.5, $c \leq 5$, unless $n = 54$, where we only get $c \leq 7$. But for $n = 54$ we can argue as follows: As λ satisfies $(*)$, without loss of generality $43, 47 \in \mathrm{fch}(\lambda)$. Then $l_1(\lambda) \geq 35$ (by [Balog et al. 2001, Proposition 2.6]), and hence $17, 34 \in \mathrm{fch}(\lambda)$; since $54 > 3 \cdot 17$, by the hook formula there has to be one more hook of length divisible by 17 in λ. As $c \leq 7$, this is in the first row or column; if it is not in the first column, we get a contradiction considering this hook and the one of length 43. Thus $51 \in \mathrm{fch}(\lambda)$, and hence $c \leq 3$.

Hence for all $n \leq 937$ we have $((2c+2)!)_2 \leq (12!)_2 = 2^{10}$, and

$$2^m \leq 2^{10} n^2 \leq 2^{10}(2m+3)^2.$$

This implies $m \leq 20$, and hence $n \leq 43$, where the assertion was checked directly. □

Next we deal with the spin characters of $\widehat{\mathsf{S}}_n$. Recall that for a bar partition λ of n, $\bar{g}_\lambda$ is the number of shifted standard tableaux of shape λ, and the spin character degree associated to λ is $g_\lambda = 2^{\lfloor\frac{n-l(\lambda)}{2}\rfloor}\bar{g}_\lambda$. Hence the condition on the spin degree translates into the condition (†) on $\bar{g}_\lambda$ given below.

Proposition 7.6. *Let λ be a bar partition of n. Then*

$$\bar{g}_\lambda = \begin{cases} 2^{\lceil\frac{l(\lambda)-2}{2}\rceil}(n-1) & \textit{for } n \textit{ even},\\ 2^{\lceil\frac{l(\lambda)-3}{2}\rceil}(n-1) & \textit{for } n \textit{ odd}, \end{cases} \tag{†}$$

only if $n \le 6$ and $\lambda = (2)$, $\lambda = (4,2)$.

We note that, for $n \le 34$, the assertion is easily checked by computer calculation (using John Stembridge's Maple package QF), so we may assume that $n > 34$ when needed.

We set $b_i = b(1,i)$ for the *first row bar lengths* of λ, and $\mathrm{frb}(\lambda)$ for the set of first row bar lengths of λ (see [Olsson 1993] for details on the combinatorics of bars).

In analogy to the case of ordinary characters where we have modified the results in [Balog et al. 2001], we adapt the results in [Bessenrodt and Olsson 2002] for the case under consideration now. Similarly to Proposition 7.3 before, we have a version of [Bessenrodt and Olsson 2002, Proposition 2.5] where, instead of the prime power condition for $\bar{g}_\lambda$, the condition $q \nmid \bar{g}_\lambda$ is assumed for the prime q under consideration. For the corresponding variant of [Bessenrodt and Olsson 2002, Lemma 2.6] that says that any prime q with $n/2 < q \le n$ and $q \nmid \bar{g}_\lambda$ is a first row bar length of λ, we need two primes $p_1, p_2 \ne n-1$ with $p_1 + p_2 - n > n/2$. For $n \ge 33$, $n \ne 42$, we always find two primes $p_1, p_2 \ne n-1$ such that $\frac{3}{4}n < p_1 < p_2 \le n$. But for $n = 42$, the primes $p_1 = 31$ and $p_2 = 37$ are big enough to have $p_1 + p_2 - n > n/2$.

The largest bar length of λ is $b_1 = b(1,1) = \lambda_1 + \lambda_2$. As before, the preparatory results just described together with our arithmetical Theorem 7.2 show that $n - b_1$ is small for a bar partition λ satisfying (†). More precisely, we obtain:

Proposition 7.7. *Let $n \ge 15$. Let $((s_1,\dots,s_r),(t_1,\dots,t_r)) \in M(n)$. Assume that λ is a bar partition of n that satisfies* (†). *Then $s_1,\dots,s_r,t_1,\dots,t_r \in \mathrm{frb}(\lambda)$.*

In particular, if λ satisfies (†), *then $n - b_1 \le 225$, and we have tighter bounds for $n - b_1$ when $n \le 10^9$ as given in Theorem 7.2.*

Now we can get into the second part of the proof of Proposition 7.1, showing the nonexistence of spin irreducible characters of the critical degree for $n \ge 7$.

Second part of the proof of Proposition 7.1. We have already seen that it suffices to prove Proposition 7.6, and that we may assume $n \geq 15$. Set

$$r = \begin{cases} \left\lceil \dfrac{l(\lambda)-2}{2} \right\rceil & \text{for } n \text{ even,} \\ \left\lceil \dfrac{l(\lambda)-3}{2} \right\rceil & \text{for } n \text{ odd,} \end{cases}$$

and assume that λ is a bar partition of n that satisfies (†).

Let $c = n - b_1$. As seen above, we have $c \leq 225$, and hence $l(\lambda) \leq 23$. Thus $r \leq 11$ in any case, and hence $\bar{g}_\lambda \leq 2^{11}(n-1)$.

Now, by [Bessenrodt and Olsson 2002, Proposition 2.2] we know that $\bar{g}_\lambda \geq \frac{1}{2}(n-1)(n-4)$ unless we have one of the following situations: $\lambda = (n)$and $\bar{g}_\lambda = 1$, or $\lambda = (n-1, 1)$ and $\bar{g}_\lambda = n-2$. None of these exceptional cases is relevant here, and thus we obtain $n - 4 \leq 2^{12}$. But for $n \leq 4100$ we already know that $c \leq 7$. Then $l(\lambda) \leq 6$ and $r \leq 2$, and hence

$$\tfrac{1}{2}(n-1)(n-4) \leq \bar{g}_\lambda \leq 4(n-1).$$

But then $n - 4 \leq 8$, a contradiction. Thus we have now completed the proof of Proposition 7.1. □

8. Eliminating simple groups of Lie type in even characteristic

Let G be a finite group whose complex group algebra is isomorphic to that of $\widehat{\mathsf{S}}_n$. In order to show that G is isomorphic to one of the two double covers of S_n, one has to eliminate the involvement of all nonabelian simple groups other than A_n in the structure of G. The most difficult case turns out to be the simple groups of Lie type in even characteristic.

For the purpose of the next lemma, let $\mathscr{C}$ be the set consisting of the following simple groups:

$$\{{}^2F_4(2)', \mathrm{PSL}_4(2), \mathrm{PSL}_3(4), \mathrm{PSU}_4(2), \mathrm{PSU}_6(2), \mathrm{P}\Omega_8^+(2), \mathrm{PSp}_6(2), {}^2B_2(8), G_2(4), {}^2E_6(2)\}.$$

Lemma 8.1. *If S is a simple group of Lie type in characteristic 2 such that $|S|_2 \geq 2^4$ and $S \notin \mathscr{C}$, then $|S|_2 < 2^{(e(S)-1)/2}$, where $e(S)$ is the smallest nontrivial degree of an irreducible projective representation of S.*

Proof. Assume that S is defined over a finite field of size $q = 2^f$. If $|S|_2 = q^{N(S)}$, then the inequality in the lemma is equivalent to

$$e(S) > 2N(S)f + 1. \tag{8-1}$$

The values of $e(S)$ are available in [Tiep and Zalesskii 1996, Table II] for classical groups and in [Lübeck 2001] for exceptional groups. The arguments for simple classical groups are quite similar. So let us consider the linear groups. Assume that $S = \mathrm{PSL}_m(q)$ with $m \geq 2$. We have $N(S) = m(m-1)/2$. First we assume that $m = 2$. As $|S|_2 = q \geq 2^4$, we deduce that $e(S) = q-1$ and $f \geq 4$. In this case, we obtain $e(S)-1 = 2^f - 2$ and $2N(S)f + 1 = 2f + 1$. As $f \geq 4$, we see that (8-1) holds. Next we assume that $m \geq 3$ and $S \neq \mathrm{PSL}_3(4), \mathrm{PSL}_4(2)$. As $|S|_2 \geq 2^4$, we deduce that $S \neq \mathrm{PSL}_3(2)$. We have $e(S) = (q^m - q)/(q-1)$. Now (8-1) is equivalent to

$$\frac{q^m - q}{q-1} > m(m-1)f + 1.$$

It is routine to check that this inequality holds for any $m \geq 3$ and $q \geq 2$.

The arguments for exceptional Lie type groups are also similar. For instance, if $S = {}^2B_2(2^{2m+1})$ with $m \geq 1$, then $|S|_2 = 2^{2(2m+1)}$ and $e(S) = 2^m(2^{2m+1}-1)$. The inequality can now be easily checked. □

Proposition 8.2. *Let G be a finite group and S a simple group of Lie type in characteristic 2. Suppose that $M \subseteq G'$ is a normal subgroup of G such that $G'/M \cong S$ and $|G : G'| = 2$. Then $\mathrm{cd}^*(G) \neq \mathrm{cd}^*(\hat{\mathsf{S}}_n)$ for every integer $n \geq 10$.*

Proof. By way of contradiction, assume that $\mathrm{cd}^*(G) = \mathrm{cd}^*(\hat{\mathsf{S}}_n)$ for some $n \geq 10$. Let St_S be the Steinberg character of $G'/M \cong S$. As St_S extends to G/M and $|G/M : G'/M| = 2$, by Gallagher's theorem (see [Isaacs 1994, Corollary 6.17] for instance) G/M has two irreducible characters of degree $\mathrm{St}_S(1) = |S|_2$. As $n \geq 10$, Lemma 3.5(1) yields that $d_1(\hat{\mathsf{S}}_n) = n-1$.

We claim that $|S|_2 > d_1(\hat{\mathsf{S}}_n) = n-1$. Suppose for a contradiction that $|S|_2 = d_1(\hat{\mathsf{S}}_n)$. If $G/M \cong S \times C_2$, then $\mathrm{cd}(G/M) = \mathrm{cd}(S) \subseteq \mathrm{cd}(\hat{\mathsf{S}}_n)$, which implies that $d_1(S) \geq d_1(\hat{\mathsf{S}}_n) = |S|_2$. However, this is impossible since S always has a nontrivial character degree smaller than $|S|_2$ (see [Tong-Viet 2012, Lemma 8] for instance). Now assume that G/M is almost simple with socle S. If $S \not\cong \mathrm{PSL}_2(q)$ with $q \geq 4$, then $d_1(G/M) < |S|_2 = d_1(\hat{\mathsf{S}}_n)$ by [Tong-Viet 2011, Lemma 2.4], which leads to a contradiction as before since $\mathrm{cd}(G/N) \subseteq \mathrm{cd}(\hat{\mathsf{S}}_n)$. Therefore, assume that $S = \mathrm{PSL}_2(q)$ with $q = 2^f \geq 4$. Then $q = |S|_2 = n-1 \geq 9$. If $q \equiv -1 \pmod 3$, then $d_1(G) = q-1 < q = d_1(\hat{\mathsf{S}}_n)$ by [Tong-Viet 2011, Lemma 2.5], which is impossible. Hence, $q \equiv 1 \pmod 3$ and $q+1 \in \mathrm{cd}(G/M) \subseteq \mathrm{cd}(\hat{\mathsf{S}}_n)$. It follows that $q+1 = (n-1)+1 = n \in \mathrm{cd}(\hat{\mathsf{S}}_n)$. If $n \geq 12$, then $d_2(\hat{\mathsf{S}}_n) > 2n > n > d_1(\hat{\mathsf{S}}_n)$, hence n is not a degree of $\hat{\mathsf{S}}_n$. Thus $10 \leq n \leq 11$. However, we see that $n-1$ is not a power of 2 in either case. The claim is proved.

Assume that $n = 2k + 1 \geq 11$ is odd. By Lemmas 2.1 and 2.3, $|S|_2 = 2^k$ is the degree of the basic spin character of $\widehat{\mathsf{S}}_n$. However, by [Wales 1979, Table I] such a degree has multiplicity 1, which contradicts the fact proved above that G has at least two irreducible characters of degree $|S|_2$.

Assume $n = 2k \geq 10$ is even. By Lemmas 2.1 and 2.3 and [Wales 1979, Table I], $\widehat{\mathsf{S}}_n$ always has the character degree 2^{k-1} with multiplicity 2, and, if $n = 2^r + 2$, then it has the character degree $2^{k-1}(n-2) = 2^{2^{r-1}+r}$ with multiplicity 1. These are in fact the only nontrivial 2-power character degrees of $\widehat{\mathsf{S}}_n$. As in the previous case, by comparing the multiplicity, we see that $|S|_2 \neq 2^{2^{r-1}+r}$. Thus $|S|_2 = 2^{k-1}$ is the degree of the basic spin character of $\widehat{\mathsf{S}}_n$ with multiplicity 2. Notice that $k \geq 5$ and hence $|S|_2 = 2^{k-1} \geq 2^4$.

Now let $\psi \in \mathrm{Irr}(G)$ with $\psi(1) = n-1$. As $|G : G'| = 2$ and $\psi(1)$ is odd, we deduce that $\phi = \psi \downarrow_{G'} \in \mathrm{Irr}(G')$ and $\phi(1) = n-1$. Let $\theta \in \mathrm{Irr}(M)$ be an irreducible constituent of $\phi \downarrow_M$. Then $\phi \downarrow_M = e(\theta_1 + \cdots + \theta_t)$, where $t = |G' : I_{G'}(\theta)|$, and each θ_i is conjugate to $\theta \in \mathrm{Irr}(M)$. If θ is not G'-invariant, then $\phi(1) = et\theta(1) = n - 1 \geq \min(S)$, where $\min(S)$ is the smallest nontrivial index of a maximal subgroup of S. We see that $\min(S) > d_1(S) \geq e(S)$, where $e(S)$ is the minimal degree of a projective irreducible representation of S, and so $n-1 \geq e(S)$. If θ is G'-invariant and $\phi \downarrow_M = e\theta$ with $e > 1$, then e is the degree of a projective irreducible representation of S. It follows that $n-1 \geq e \geq e(S)$. In both cases, we always have

$$k - 1 = \frac{n-2}{2} \geq \frac{e(S)-1}{2}.$$

Therefore,

$$|S|_2 = 2^{k-1} \geq 2^{(e(S)-1)/2}.$$

By Lemma 8.1, we deduce that $S \in \mathscr{C}$. Solving the equation $|S|_2 = 2^{(n-2)/2}$, we get the degree n. However, by using [Conway et al. 1985] and Lemma 3.5, we can check that $\mathrm{cd}(G/M) \nsubseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$ in any of these cases. For example, assume that $S \cong {}^2E_6(2)$. Then $|S|_2 = 2^{36} = 2^{(n-2)/2}$, so $n = 74$. By Lemma 3.5, we have $d_1(\widehat{\mathsf{S}}_n) = n - 1 = 73$ and $d_2(\widehat{\mathsf{S}}_n) = n(n-3)/2 = 2627$. Using [Conway et al. 1985], we know that $\mathrm{cd}(G/M)$ contains the degree 1938. Clearly, $d_1(\widehat{\mathsf{S}}_n) < 1938 < d_2(\widehat{\mathsf{S}}_n)$, so $1938 \notin \mathrm{cd}(\widehat{\mathsf{S}}_n)$, hence $\mathrm{cd}(G/M) \nsubseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$, a contradiction.

Finally we assume that $et = 1$. Then θ extends to $\phi \in \mathrm{Irr}(G')$ and to $\psi \in \mathrm{Irr}(G)$. Hence $\phi \downarrow_M = \theta$ and so, by Gallagher's Theorem, we have $\psi\tau \in \mathrm{Irr}(G)$ for every $\tau \in \mathrm{Irr}(G/M)$. In particular,

$$2^{k-1}(n-1) = \psi(1)|S|_2 \in \mathrm{cd}(G) = \mathrm{cd}(\widehat{\mathsf{S}}_n),$$

which is impossible by Proposition 7.1. $\square$

9. Eliminating simple groups other than A_n

We continue to eliminate the involvement of simple groups other than A_n in the structure of G with $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{S}}_n$.

Proposition 9.1. *Let G be an almost-simple group with nonabelian simple socle S. Suppose that $\mathrm{cd}^*(G) \subseteq \mathrm{cd}^*(\widehat{\mathsf{S}}_n)$ for some $n \geq 10$. Then $S \cong \mathsf{A}_n$ or S is isomorphic to a simple group of Lie type in characteristic 2.*

Proof. We make use of the classification of finite simple groups.

(i) *S is a sporadic simple group or the Tits group.* Using GAP (version 4.4.12), we can assume that $n \geq 19$. By Lemma 3.5(2), we have $d_2(\widehat{\mathsf{S}}_n) = n(n-3)/2 \geq 152$. Since $d_2(G) \geq d_2(\widehat{\mathsf{S}}_n) \geq 152$, using [Conway et al. 1985], we only need to consider the following simple groups:

$$J_3, \mathrm{Suz}, \mathrm{McL}, \mathrm{Ru}, \mathrm{He}, \mathrm{Co}_1, \mathrm{Co}_2, \mathrm{Co}_3, \mathrm{Fi}_{22}, \mathrm{O'N}, \mathrm{HN}, \mathrm{Ly}, \mathrm{Th}, \mathrm{Fi}_{23}, J_4, \mathrm{Fi}'_{24}, B, M.$$

To eliminate these groups, we first observe that $n \geq p(S)$, the largest prime divisor of $|S|$, and $d_i(G) \geq d_i(\widehat{\mathsf{S}}_n)$ for all $i \geq 1$. Now with the lower bound $n \geq \max\{19, p(S)\}$, we can find the lower bounds for $d_i(\widehat{\mathsf{S}}_n)$ with $1 \leq i \leq 7$ using Lemmas 3.4 and 3.5. Choose $i \in \{2, 3, \dots, 7\}$ such that $d_i(\widehat{\mathsf{S}}_n) > d_j(G)$ for some $j \geq 1$ such that $|i-j|$ is minimal. If $i \geq j$, then we obtain a contradiction. Otherwise, $d_j(G) \in \{d_k(\widehat{\mathsf{S}}_n)\}_{k=j}^{i-1}$. Solving these equations for n, we then obtain that either these equations have no solution, or that, for each solution of n, we can find some $k \geq 1$ with $d_k(\widehat{\mathsf{S}}_n) > d_k(G)$.

For an example of such a demonstration, assume that $S = \mathrm{O'N}$. In this case, we have $|\mathrm{Out}(S)| = 2$, so $G = S$ or $G = S.2$. Since $p(S) = 31$, we have $n \geq 31$. Assume first that $G = S = \mathrm{O'N}$. Then $d_4(\mathrm{O'N}) = 26752$ and, since $n \geq 31$, by Lemma 3.4 $d_7(\widehat{\mathsf{S}}_n) \geq 26970 > d_4(\mathrm{O'N})$. It follows that $d_4(\mathrm{O'N}) \in \{d_i(\widehat{\mathsf{S}}_n)\}_{i=4}^{6}$. However, we can check that these equations are impossible. Now assume $G = \mathrm{O'N}.2$. Then $d_2(G) = 26752 < 26970 \leq d_7(\widehat{\mathsf{S}}_n)$ so that $d_2(G) \in \{d_i(\widehat{\mathsf{S}}_n)\}_{i=2}^{6}$. As above, these equations cannot hold for any $n \geq 31$. Thus $\mathrm{cd}(G) \not\subseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$.

For another example, let $S = M$. Since $|\mathrm{Out}(S)| = 1$, we have $G = S$ so that $p(S) = 71 \in \pi(\widehat{\mathsf{S}}_n)$ and hence $n \geq 71$. As $d_1(M) = 196883 < 914480 \leq d_7(\widehat{\mathsf{S}}_n)$, we deduce that $d_1(M) \in \{d_i(\widehat{\mathsf{S}}_n) \mid i = 1, \dots, 6\}$. Solving these equations, we obtain $n = 196884$. But then $d_2(\widehat{\mathsf{S}}_n) > 21296876 = d_2(M)$. Thus $\mathrm{cd}(M) \not\subseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$.

(ii) *$S = \mathsf{A}_m$ with $m \geq 7$.* Note that we consider $\mathsf{A}_5 \cong \mathrm{PSL}_2(5)$ and $\mathsf{A}_6 \cong \mathrm{PSL}_2(9)$ as groups of Lie type. Let $\lambda = (m-1, 1)$, a partition of m. Since $m \geq 7$, λ is not self-conjugate, hence the irreducible character χ_λ of S_m is still irreducible upon restriction to A_m. Note that $\mathrm{Aut}(\mathsf{A}_m) = \mathsf{S}_m$ as $m \geq 7$. Then $G \in \{\mathsf{A}_m, \mathsf{S}_m\}$ and G has an irreducible character of degree $m-1$. Since $\mathrm{cd}(G) \subseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$, we have $m - 1 \geq d_1(\widehat{\mathsf{S}}_n) = n - 1$, so $m \geq n$. If $m = n$ then we are done. On the other hand,

if $m > n$ then
$$|G| \geq |S| = |\mathsf{A}_m| > 4|\mathsf{A}_n| = |\widehat{\mathsf{S}}_n|,$$
and this violates the hypothesis $\mathrm{cd}^*(G) \subseteq \mathrm{cd}^*(\widehat{\mathsf{S}}_n)$.

(iii) *S is a simple group of Lie type in odd characteristic.* Suppose that $S = G(p^k)$, a simple group of Lie type defined over a field of p^k elements with p odd. Let St_S be the Steinberg character of S. Then, as St extends to G and $\mathrm{St}_S(1) = |S|_p$, we have $|S|_p \in \mathrm{cd}(\widehat{\mathsf{S}}_n)$. Using Lemma 2.3, which says that all possible prime power degrees of spin characters of S_n are even, we deduce that $|S|_p \in \mathrm{cd}(\mathsf{S}_n)$. By Lemma 2.1, we then obtain that $|S|_p = n-1$ since $n \geq 10$. By Lemma 3.5, $n-1 = d_1(\widehat{\mathsf{S}}_n)$ is the smallest nontrivial degree of $\widehat{\mathsf{S}}_n$. Assume first that $S \neq \mathrm{PSL}_2(q)$. Then $d_1(G) < |S|_p = d_1(\widehat{\mathsf{S}}_n)$ by [Tong-Viet 2011, Lemma 2.4], which is a contradiction as $\mathrm{cd}(G) \subseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$. Now it remains to consider the case $S = \mathrm{PSL}_2(q)$. We have $q = n-1 \geq 9$. If G has a character degree which is smaller than $|S|_p = q$, then we obtain a contradiction as before. So, by [Tong-Viet 2011, Lemma 2.5], we have $p \neq 3$ and $q \equiv 1 \pmod 3$ or $p = 3$ and $q \equiv 1 \pmod 4$. In both cases, G has an irreducible character of degree $q+1 = n = d_1(\widehat{\mathsf{S}}_n)+1$. If $n \geq 12$, then $d_2(\widehat{\mathsf{S}}_n) \geq 2n > n > d_1(\widehat{\mathsf{S}}_n)$ by Lemma 3.5, so that n is not a character degree of $\widehat{\mathsf{S}}_n$. Assume that $10 \leq n \leq 11$. Then $n = 10$ and $q = 9$. However, using [Conway et al. 1985], we can check that $\mathrm{cd}(G) \nsubseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$ for every almost-simple group G with socle $\mathrm{PSL}_2(9) \cong \mathsf{A}_6$. □

Combining Propositions 9.1 and 8.2, we obtain the following results, which will be crucial in the proof of Theorem A:

Proposition 9.2. *Let G be a finite group and let $M \subseteq G'$ be a normal subgroup of G such that G/M is an almost-simple group with socle $S \neq \mathsf{A}_n$, where $|G : G'| = 2$ and $G'/M \cong S$. Then $\mathrm{cd}^*(G) \neq \mathrm{cd}^*(\widehat{\mathsf{S}}_n)$.*

Proof. If $n \geq 10$, then the result follows from Propositions 9.1 and 8.2. It remains to assume that $5 \leq n \leq 9$ and suppose by contradiction that $\mathrm{cd}^*(G) = \mathrm{cd}^*(\widehat{\mathsf{S}}_n)$. Then $|G| = 2n!$ and so $|S| \mid 2n!$, hence $\pi(S) \subseteq \pi(\widehat{\mathsf{S}}_n) \subseteq \{2,3,5,7\}$. By [Huppert and Lempken 2000, Theorem III], one of the following holds:

(1) If $\pi(S) = \{2,3,5\}$, then $S \cong \mathsf{A}_5$, A_6 or $\mathrm{PSp}_4(3)$.

(2) If $\pi(S) = \{2,3,7\}$, then $S \cong \mathrm{PSL}_2(7)$, $\mathrm{PSL}_2(8)$ or $\mathrm{PSU}_3(3)$.

(3) If $\pi(S) = \{2,3,5,7\}$, then $S \cong \mathsf{A}_k$ with $7 \leq k \leq 10$, J_2, $\mathrm{PSL}_2(49)$, $\mathrm{PSL}_3(4)$, $\mathrm{PSU}_3(5)$, $\mathrm{PSU}_4(3)$, $\mathrm{PSp}_4(7)$, $\mathrm{PSp}_6(2)$ or $\mathrm{P}\Omega_8^+(2)$.

Now it is routine to check that $\mathrm{cd}(G/M) \nsubseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$ unless $S \cong \mathsf{A}_n$, where G/M is almost simple with socle S. □

Proposition 9.3. *Let G be a finite group and let $M \subseteq G'$ be a normal subgroup of G. Suppose that* $\mathrm{cd}^*(G) = \mathrm{cd}^*(\widehat{\mathsf{S}}_n)$ *and $G/M \cong G'/M \times C_2 \cong S^k \times C_2$ for some positive integer k and some nonabelian simple group S. Then $k = 1$ and $S \cong \mathsf{A}_n$.*

Proof. Since $\mathrm{cd}^*(S) \subseteq \mathrm{cd}^*(S^k)$, the hypotheses imply that

$$\mathrm{cd}^*(S) \subseteq \mathrm{cd}^*(\widehat{\mathsf{S}}_n).$$

Assume first that $5 \leq n \leq 9$. Since $|S^k| = |S|^k$ divides $|\widehat{\mathsf{S}}_n| = 2n!$, we deduce that $\pi(S) \subseteq \pi(\widehat{\mathsf{S}}_n)$ and, in particular, $\pi(S) \subseteq \{2, 3, 5, 7\}$. The possibilities for S are listed in the proof of Proposition 9.2 above. Observe that $|S|$ is always divisible by a prime r with $r \geq 5$. Hence, $r^k \mid |\widehat{\mathsf{S}}_n|$, which implies that $k = 1$ as $|\widehat{\mathsf{S}}_n|$ divides $2 \cdot 9!$. Now the fact that $S \cong \mathsf{A}_n$ follows easily.

From now on we can assume that $n \geq 10$. Using Proposition 9.1, we obtain $S = \mathsf{A}_n$ or S is a simple group of Lie type in even characteristic. It suffices to show that $k = 1$, and then the result follows from Proposition 8.2.

Assume that the latter case holds. Then S is a simple group of Lie type in characteristic 2. By Lemmas 2.1 and 2.3, $\widehat{\mathsf{S}}_n$ has at most two distinct nontrivial 2-power character degrees, which are $n-1$ and $2^{\lfloor (n-1)/2 \rfloor}$, or $2^{\lfloor (n-1)/2 \rfloor}$ and $2^{2^{r-1}+r}$ with $n = 2^r + 2$. By way of contradiction, assume that $k \geq 2$. If $k \geq 3$, then $G/M \cong S^k \times C_2$ has irreducible characters of degrees

$$|S|_2^k > |S|_2^{k-1} > |S|_2^{k-2} > 1.$$

Obviously, this is impossible as $\mathrm{cd}(G/M) \subseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$. Therefore, $k = 2$. In this case, G/M has character degree $|S|_2^2$ with multiplicity at least 2 and $|S|_2$ with multiplicity at least 4. It follows that either $2^{\lfloor (n-1)/2 \rfloor} = |S|_2^2$ and $n - 1 = |S|_2$, or $2^{2^{r-1}+r} = |S|_2^2$ and $2^{\lfloor (n-1)/2 \rfloor} = |S|_2$. However, both cases are impossible by comparing the multiplicity.

It remains to eliminate the case $S \cong \mathsf{A}_n$ and $k \geq 2$. By comparing the orders, we see that

$$2(n!/2)^k |M| = 2n!.$$

After simplifying, we obtain

$$|M|(n!)^{k-1} = 2^k.$$

Since $n \geq 10$, we see that, if $k \geq 2$, then the left side is divisible by 5, while the right side is not. We conclude that $k = 1$, and the proof is now complete. $\square$

10. Completion of the proof of Theorem C

We need one more result before proving Theorem C.

Proposition 10.1. *Let G be a finite group and let S be a nonabelian simple group. Suppose that $|G : G'| = 2$ and $G' \cong S^2$ is the unique minimal normal subgroup of G. Then* $\mathrm{cd}(G) \not\subseteq \mathrm{cd}(\hat{\mathsf{S}}_n)$.

Proof. Assume, to the contrary, that $\mathrm{cd}(G) \subseteq \mathrm{cd}(\hat{\mathsf{S}}_n)$. Let $\alpha \in \mathrm{Irr}(S)$ with $\alpha(1) > 1$ and put $\theta = \alpha \otimes 1 \in \mathrm{Irr}(G')$. Observe that θ is not G-invariant, so that $I_G(\theta) = G'$; hence $\theta^G \in \mathrm{Irr}(G)$ and so $\theta^G(1) = 2\alpha(1) \in \mathrm{cd}(\hat{\mathsf{S}}_n)$. On the other hand, if $\varphi = \alpha \otimes \alpha \in \mathrm{Irr}(G')$, then φ is G-invariant and, since G/G' is cyclic, we deduce that φ extends to $\psi \in \mathrm{Irr}(G)$, so $\psi(1) = \alpha(1)^2 \in \mathrm{cd}(\hat{\mathsf{S}}_n)$. Thus, we conclude that

$$\text{if } a \in \mathrm{cd}(S)\backslash\{1\} \quad \text{then } 2a, a^2 \in \mathrm{cd}(\hat{\mathsf{S}}_n). \tag{10-1}$$

Let r be an odd prime divisor of $|S|$. The Ito–Michler theorem then implies that r divides some character degree, say a, of S. Since $a^2 \in \mathrm{cd}(\hat{\mathsf{S}}_n)$ by (10-1), we have $r^2 \mid 2n!$ and hence $n \geq 2r$ as $r > 2$. Thus, we have shown that

$$\text{if } r \in \pi(S) - \{2\} \quad \text{then } r^2 \mid 2n! \quad \text{and} \quad n \geq 2r. \tag{10-2}$$

Using the classification of finite simple groups, we consider the following cases:

(i) $S = \mathsf{A}_m$, *with* $m \geq 7$. As $7 \in \pi(S)$, it follows from (10-2) that $n \geq 14$. Since $m-1 \in \mathrm{cd}(S)$, both $2(m-1)$ and $(m-1)^2$ are in $\mathrm{cd}(\hat{\mathsf{S}}_n)$ by (10-1). As $m \geq 7$, we also have that $m(m-3)/2, (m-1)(m-2)/2 \in \mathrm{cd}(S)$ and so $m(m-3), (m-1)(m-2) \in \mathrm{cd}(\hat{\mathsf{S}}_n)$. We claim that $m < n$. Suppose for a contradiction that $m \geq n$. As $n \geq 14$, by Lemma 4.1, there exists a prime r such that $n/2 < r \leq n$. Hence, the r-part of $2n!$ is just r. However, as $r \leq n \leq m$, r divides $|\mathsf{A}_m|$ and so $r^2 \mid 2n!$ by (10-2), a contradiction. Thus $m < n$ as claimed.

Since $m \geq 7$, we obtain that

$$1 < 2(m-1) < m(m-3) < (m-1)(m-2) < (m-1)^2. \tag{10-3}$$

By Lemma 3.5(1), we have $d_1(\hat{\mathsf{S}}_n) = n-1$, so $2(m-1) \geq n-1$, and thus $n \leq 2m-1$. As $n \geq 14$, we deduce that $m \geq 8$.

Assume first that $m \in \{8, 9, 10\}$. Then $(m-1)^2 \in \mathrm{cd}(\hat{\mathsf{S}}_n)$ is a prime power. As $3^3 \neq (m-1)^2 > d_1(\hat{\mathsf{S}}_n)$, Lemmas 2.1 and 2.3 yield that $(m-1)^2$ is a power of 2, and thus $m = 9$. Since $\{2^3, 3^3\} \subseteq \mathrm{cd}(\mathsf{A}_9)$, we have $\{2^4, 2^6, 3^6\} \subseteq \mathrm{cd}(\hat{\mathsf{S}}_n)$ by (10-2). As $n \geq 14$, we have $2^{\lfloor (n-1)/2 \rfloor} > 2^4$ and $n(n-3)/2 > 2^4$, so $d_2(\hat{\mathsf{S}}_n) > 2^4$ by Lemma 3.5(2). This forces $2^4 = d_1(\hat{\mathsf{S}}_n) = n-1$ or, equivalently, $n = 17$. But then Lemmas 2.1 and 2.3 yield $3^6 = d_1(\hat{\mathsf{S}}_n)$, which is impossible.

Assume next that $m \geq 11$. Then $n \geq 22$ by (10-2). By Lemma 3.5(2), we have $d_2(\hat{\mathsf{S}}_n) > 2n > 2m$. In particular, $2(m-1) < d_2(\hat{\mathsf{S}}_n)$. By (10-1), we have $2(m-1) \in \mathrm{cd}(\hat{\mathsf{S}}_n)$, hence $2(m-1) = d_1(\hat{\mathsf{S}}_n) = n-1$, which implies that $n = 2m-1$. By Lemma 3.5(3), we have $d_3(\hat{\mathsf{S}}_n) = (n-1)(n-2)/2$ and thus by (10-3) we obtain that

$$(m-1)(m-2) \geq d_3(\hat{\mathsf{S}}_n) = (n-1)(n-2)/2 = (m-1)(2m-3).$$

After simplifying, we have $m-2 \geq 2m-3$ or, equivalently, $m \leq 1$, a contradiction.

(ii) *S is a finite simple group of Lie type in characteristic p, with $S \neq {}^2F_4(2)'$.* As $|S|$ is always divisible by an odd prime $r \geq 5$, we have $n \geq 2r \geq 10$ by (10-2). Let St_S denote the Steinberg character of S. We can check that $\mathrm{St}_S(1) = |S|_p \geq 4$. Since $\mathrm{St}_S(1) \in \mathrm{cd}(S)$, $2\,\mathrm{St}_S(1)$ and $\mathrm{St}_S(1)^2$ are character degrees of $\widehat{\mathsf{S}}_n$ by (10-1). As $2\,\mathrm{St}_S(1) < \mathrm{St}_S(1)^2$, we have $\mathrm{St}(1)^2 > d_1(\widehat{\mathsf{S}}_n) = n-1$. Since $n \geq 10$, Lemmas 2.1 and 2.3 yield that $\mathrm{St}_S(1)^2$ is a 2-power. Hence, $2\,\mathrm{St}(1)$ is also a 2-power. By [Tong-Viet 2012, Lemma 8], there exists a nontrivial character degree x of S such that $1 < x < \mathrm{St}_S(1)$. It follows that $2x < 2\,\mathrm{St}_S(1)$ is also a character degree of $\widehat{\mathsf{S}}_n$. Therefore, $2\,\mathrm{St}_S(1) > d_1(\widehat{\mathsf{S}}_n) = n-1$. Hence, $\widehat{\mathsf{S}}_n$ has two distinct nontrivial 2-power character degrees, neither of which is $n-1$. It follows that $n = 2^r + 2 \geq 10$, and furthermore

$$2\,\mathrm{St}_S(1) = 2^{\lfloor (n-1)/2 \rfloor} \quad \text{and} \quad \mathrm{St}_S(1)^2 = 2^{2^{r-1}+r}$$

by Lemma 2.3. We write $\mathrm{St}_S(1) = 2^N$. Then $2^{r-1} + r = 2N$ and $2^{r-1} = N+1$ since $\lfloor (n-1)/2 \rfloor = 2^{r-1}$. Solving these equations, we have $r = N-1$ and $2^{r-1} = r+2$. As $n \geq 10$, we deduce that $r \geq 3$. In this case, it is easy to check that the equation $2^{r-1} = r+2$ has no integer solution.

(iii) *S is a sporadic simple group or the Tits group.* Since the arguments are fairly similar, we consider just the case $S = J_3$ as an example. Recall that $p(S)$ is the largest prime divisor of $|S|$. By (10-2), we have $n \geq 2p(S)$. Since $n \geq 2p(S) \geq 22$, we have $d_2(\widehat{\mathsf{S}}_n) = n(n-3)/2 \geq p(S)(2p(S)-3)$ by applying Lemma 3.5(3). For $i = 1, 2$, we have $2d_i(S) \in \mathrm{cd}(G) \subseteq \mathrm{cd}(\widehat{\mathsf{S}}_n)$ with $1 < 2d_1(S) < 2d_2(S)$. For each possibility for S, we can check using [Conway et al. 1985] that $p(S)(2p(S)-3) > 2d_2(S)$; hence $d_2(\widehat{\mathsf{S}}_n) > 2d_2(S) > 2d_1(S)$, which is a contradiction. □

We are now ready to prove the main result, Theorem C, which we restate below for the reader's convenience:

Theorem C. *Let $n \geq 5$. Let G be a finite group and $\widehat{\mathsf{S}}_n^{\pm}$ the double covers of S_n. Then $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{S}}_n^{+}$ (or equivalently $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{S}}_n^{-}$) if and only if $G \cong \widehat{\mathsf{S}}_n^{+}$ or $G \cong \widehat{\mathsf{S}}_n^{-}$.*

Proof. By the hypothesis that $\mathbb{C}G \cong \mathbb{C}\widehat{\mathsf{S}}_n$, we have $|G| = 2n!$, and, as $\widehat{\mathsf{S}}_n$ has two linear characters, we also have $|G : G'| = 2$.

First we claim that $G' = G''$. Assume not. Then $H := G/G'$ is a group whose commutator subgroup H' is nontrivial abelian of index 2. Now the induction of any nonprincipal (linear) character of H' to H must be irreducible and 2-dimensional. This is not possible since $\widehat{\mathsf{S}}_n$ with $n \geq 5$ does not have any irreducible character of degree 2. Thus $G' = G''$.

As $G' = G''$ and G' is nontrivial, one can choose a normal subgroup M of G such that $M < G'$ and

$$G'/M \cong S^k,$$

where S is a nonabelian simple group and S^k is a chief factor of G. Let

$$C/M := \boldsymbol{C}_{G/M}(G'/M).$$

(A) First we consider the case $C = M$. Then G'/M is the unique minimal normal subgroup of G/M. Therefore G/M permutes the direct factors of G'/M (which is isomorphic to S^k). It then follows that $k \leq 2$ as $|G : G'| = 2$. Invoking Proposition 10.1, we deduce that $k = 1$ and thus $G'/M \cong S$. Therefore, G'/M is the socle of G/M. As $\mathrm{cd}^*(G/M) \subseteq \mathrm{cd}^*(\widehat{\mathsf{S}}_n)$, Proposition 9.2 then implies that $G'/M \cong \mathsf{A}_n$. Thus $G/M \cong \mathsf{S}_n$ and also $|M| = 2$. In particular, M is central in G and therefore $M \subseteq \boldsymbol{Z}(G) \cap G'$. We conclude that G is one of the two double covers of S_n, as desired.

(B) It remains to consider the case $C > M$. Since $C/M \lhd G/M$ and $\boldsymbol{Z}(G'/M) = 1$, it follows that G' is a proper subgroup of $G'C$. As $|G : G'| = 2$, we then deduce that $G = G'C$ and hence

$$G/M = G'/M \times C/M, \quad \text{where } C/M \cong C_2.$$

Applying Proposition 9.3, we obtain that $k = 1$ and S is isomorphic to A_n. In other words, $G'/M \cong \mathsf{A}_n$. So $G/M \cong \mathsf{A}_n \times C_2$. Comparing the orders, we get $|M| = 2$ and so $M \subseteq \boldsymbol{Z}(G)$. As $M \leq G' = G''$, it follows that $M \leq \boldsymbol{Z}(G') \cap G''$, which in turn implies that G' is the double cover of A_n. We have proved that

$$G' \cong \widehat{\mathsf{A}}_n. \tag{10-4}$$

Moreover, as $C/M \cong C_2$ and $|M| = 2$, we have

$$C \cong C_4 \text{ or } C_2 \times C_2. \tag{10-5}$$

Now we claim that G is an (internal) central product of G' and C with amalgamated central subgroup M. To see this, let $x, y \in G'$ and $c \in C$. Then the facts $C/M = \boldsymbol{C}_{G/M}(G'/M)$ and $M \leq \boldsymbol{Z}(G)$ imply

$$[x, y]^c = [x^c, y^c] = [xm_1, ym_2] = [x, y]$$

for some $m_1, m_2 \in M$. Therefore, C centralizes $G' = G''$ and the claim follows. This claim, together with (10-4) and (10-5), yield

$$4k(\widehat{\mathsf{A}}_n) = 4k(G') = k(G' \times C) \leq k(M)k(G) = 2k(G),$$

where the inequality comes from the well-known result that $k(X) \leq k(N)k(X/N)$ for N a normal subgroup of X (see [Nagao 1962] for instance). Since $\mathrm{cd}^*(G) = \mathrm{cd}^*(\widehat{\mathsf{S}}_n)$, we have $k(G) = k(\widehat{\mathsf{S}}_n)$. It follows that $2k(\widehat{\mathsf{A}}_n) \leq k(\widehat{\mathsf{S}}_n)$. This however contradicts Lemma 4.5, and the theorem is now completely proved. □

Acknowledgements

We close this article with many thanks to the referee for the careful reading of the manuscript and several helpful comments.

References

[Balog et al. 2001] A. Balog, C. Bessenrodt, J. B. Olsson, and K. Ono, "Prime power degree representations of the symmetric and alternating groups", *J. London Math. Soc.* (2) **64**:2 (2001), 344–356. MR 2002g:20025 Zbl 1018.20008

[Bessenrodt and Olsson 2002] C. Bessenrodt and J. B. Olsson, "Prime power degree representations of the double covers of the symmetric and alternating groups", *J. London Math. Soc.* (2) **66**:2 (2002), 313–324. MR 2003e:20012 Zbl 1018.20009

[Brauer 1963] R. Brauer, "Representations of finite groups", pp. 133–175 in *Lectures on Modern Mathematics, I*, edited by T. L. Saaty, Wiley, New York, 1963. MR 31 #2314 Zbl 0124.26504

[Brunat 2010] O. Brunat, "Counting p'-characters in finite reductive groups", *J. London Math. Soc.* (2) **81**:3 (2010), 544–562. MR 2012a:20023 Zbl 1251.20017

[Conway et al. 1985] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups*, Oxford University Press, 1985. MR 88g:20025 Zbl 0568.20001

[Frame et al. 1954] J. S. Frame, G. d. B. Robinson, and R. M. Thrall, "The hook graphs of the symmetric groups", *Canadian J. Math.* **6** (1954), 316–324. MR 15,931g Zbl 0055.25404

[Harborth and Kemnitz 1981] H. Harborth and A. Kemnitz, "Calculations for Bertrand's postulate", *Math. Mag.* **54**:1 (1981), 33–34. MR 82c:10001 Zbl 0453.10006

[Hoffman and Humphreys 1992] P. N. Hoffman and J. F. Humphreys, *Projective representations of the symmetric groups*, Clarendon, New York, 1992. MR 94f:20027 Zbl 0777.20005

[Huppert 2000] B. Huppert, "Some simple groups which are determined by the set of their character degrees, I", *Illinois J. Math.* **44**:4 (2000), 828–842. MR 2001k:20009 Zbl 0972.20006

[Huppert and Lempken 2000] B. Huppert and W. Lempken, "Simple groups of order divisible by at most four primes", *Izv. Gomel. Gos. Univ. Im. F. Skoriny* **2000**:3(16) (2000), 64–75. Zbl 1159.20303

[Isaacs 1994] I. M. Isaacs, *Character theory of finite groups*, Dover, New York, 1994. Corrected reprint of the 1976 original. MR 1280461 Zbl 0849.20004

[Kleshchev and Tiep 2004] A. S. Kleshchev and P. H. Tiep, "On restrictions of modular spin representations of symmetric and alternating groups", *Trans. Amer. Math. Soc.* **356**:5 (2004), 1971–1999. MR 2005a:20020 Zbl 1065.20013

[Kleshchev and Tiep 2012] A. S. Kleshchev and P. H. Tiep, "Small-dimensional projective representations of symmetric and alternating groups", *Algebra Number Theory* **6**:8 (2012), 1773–1816. MR 3033527

[Lübeck 2001] F. Lübeck, "Smallest degrees of representations of exceptional groups of Lie type", *Comm. Algebra* **29**:5 (2001), 2147–2169. MR 2002g:20029 Zbl 1004.20003

[Macdonald 1971] I. G. Macdonald, "On the degrees of the irreducible representations of symmetric groups", *Bull. London Math. Soc.* **3** (1971), 189–192. MR 44 #6865 Zbl 0219.20008

[Macdonald 1995] I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed., Clarendon, New York, 1995. MR 96h:05207 Zbl 0824.05059

[McKay 1972] J. McKay, "Irreducible representations of odd degree", *J. Algebra* **20** (1972), 416–418. MR 44 #4111 Zbl 0235.20009

[Morris 1962] A. O. Morris, "The spin representation of the symmetric group", *Proc. London Math. Soc.* (3) **12** (1962), 55–76. MR 25 #133 Zbl 0104.25202

[Nagao 1962] H. Nagao, "On a conjecture of Brauer for p-solvable groups", *J. Math. Osaka City Univ.* **13** (1962), 35–38. MR 27 #2547 Zbl 0115.25501

[Navarro 2010] G. Navarro, "Problems in character theory", pp. 97–125 in *Character theory of finite groups*, Contemp. Math. **524**, Amer. Math. Soc., Providence, RI, 2010. MR 2012a:20014 Zbl 1235.20009

[Nguyen 2010] H. N. Nguyen, "Low-dimensional complex characters of the symplectic and orthogonal groups", *Comm. Algebra* **38**:3 (2010), 1157–1197. MR 2011h:20025 Zbl 1237.20012

[Nguyen 2013] H. N. Nguyen, "Quasisimple classical groups and their complex group algebras", *Israel J. Math.* **195**:2 (2013), 973–998. MR 3096579 Zbl 1294.20004

[Nguyen and Tong-Viet 2014] H. N. Nguyen and H. P. Tong-Viet, "Characterizing finite quasisimple groups by their complex group algebras", *Algebr. Represent. Theory* **17**:1 (2014), 305–320. MR 3160726 Zbl 1303.20001

[Olsson 1993] J. B. Olsson, *Combinatorics and representations of finite groups*, Fachbereich Mathematik der Universität Essen, Heft **20**, 1993. MR 95b:20020 Zbl 0796.05095

[Rasala 1977] R. Rasala, "On the minimal degrees of characters of S_n", *J. Algebra* **45**:1 (1977), 132–181. MR 55 #477 Zbl 0348.20009

[Schur 1911] I. Schur, "Über die Darstellung der symmetrischen und der alternierenden Gruppe durch gebrochene lineare Substitutionen", *J. Reine Angew. Math.* **139** (1911), 155–250.

[Tiep and Zalesskii 1996] P. H. Tiep and A. E. Zalesskii, "Minimal characters of the finite classical groups", *Comm. Algebra* **24**:6 (1996), 2093–2167. MR 97f:20018 Zbl 0901.20031

[Tong-Viet 2011] H. P. Tong-Viet, "Symmetric groups are determined by their character degrees", *J. Algebra* **334** (2011), 275–284. MR 2012c:20020 Zbl 1246.20007

[Tong-Viet 2012] H. P. Tong-Viet, "Alternating and sporadic simple groups are determined by their character degrees", *Algebr. Represent. Theory* **15**:2 (2012), 379–389. MR 2892513 Zbl 1252.20005

[Wagner 1977] A. Wagner, "An observation on the degrees of projective representations of the symmetric and alternating group over an arbitrary field", *Arch. Math.* **29**:6 (1977), 583–589. MR 57 #444 Zbl 0383.20009

[Wales 1979] D. B. Wales, "Some projective representations of S_n", *J. Algebra* **61**:1 (1979), 37–57. MR 81f:20015 Zbl 0433.20010

Communicated by David Benson
Received 2014-05-19 Revised 2015-01-13 Accepted 2015-02-23

bessen@math.uni-hannover.de *Institut für Algebra, Zahlentheorie und Diskrete Mathematik, Fakultät für Mathematik und Physik, Leibniz Universität Hannover, Welfengarten 1, D-30167 Hannover, Germany*

hungnguyen@uakron.edu *Department of Mathematics, The University of Akron, Akron, OH 44325, United States*

olsson@math.ku.dk *Department of Mathematical Sciences, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark*

Hung.Tong-Viet@up.ac.za *Department of Mathematics and Applied Mathematics, University of Pretoria, Private Bag X20, Hatfield, Pretoria 0002, South Africa*

ALGEBRA AND NUMBER THEORY 9:3 (2015)
dx.doi.org/10.2140/ant.2015.9.629

Fano schemes of determinants and permanents

Melody Chan and Nathan Ilten

Let $D^r_{m,n}$ and $P^r_{m,n}$ denote the subschemes of $\mathbb{P}^{mn-1}$ given by the $r \times r$ determinants (respectively the $r \times r$ permanents) of an $m \times n$ matrix of indeterminates. In this paper, we study the geometry of the Fano schemes $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ parametrizing the k-dimensional planes in $\mathbb{P}^{mn-1}$ lying on $D^r_{m,n}$ and $P^r_{m,n}$, respectively. We prove results characterizing which of these Fano schemes are smooth, irreducible, and connected; and we give examples showing that they need not be reduced. We show that $\boldsymbol{F}_1(D^n_{n,n})$ always has the expected dimension, and we describe its components exactly. Finally, we give a detailed study of the Fano schemes of k-planes on the 3×3 determinantal and permanental hypersurfaces.

1.	Introduction	629
2.	Compression spaces and torus fixed points	635
3.	Tangent space calculations	638
4.	Results on compression spaces	652
5.	Smoothness and connectedness	662
6.	Fano schemes of lines	669
7.	Fano schemes for 3×3 matrices	671
8.	Comparisons and further questions	676
	Acknowledgments	678
	References	678

1. Introduction

Fix an algebraically closed field $\mathbb{K}$. For numbers r, m, n with $1 < r \leq m \leq n$, let $D^r_{m,n}$ and $P^r_{m,n}$ denote the subschemes of $\mathbb{P}^{mn-1}_{\mathbb{K}}$ defined by the $r \times r$ determinants and, respectively, the $r \times r$ permanents of an $m \times n$ matrix

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,n} \end{pmatrix}$$

MSC2010: primary 14M12; secondary 14N20, 14C05, 15A15, 14B10.
Keywords: Fano schemes, determinantal varieties, permanent.

filled with mn independent forms $x_{i,j}$. Whenever we are dealing with $P^r_{m,n}$, we will make the standard assumption that char $\mathbb{K} \neq 2$. In this article, we study the Fano schemes $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$. These are subschemes of the Grassmannian $\mathrm{Gr}(k+1, mn)$ parametrizing those k-dimensional planes in $\mathbb{P}^{mn-1}_{\mathbb{K}}$ that are contained in the schemes $D^r_{m,n}$ and $P^r_{m,n}$, respectively.

We have three main reasons for studying these Fano schemes. First, we would like to understand general Hilbert schemes better. In the case of classical Hilbert schemes — those parametrizing subschemes of $\mathbb{P}^n$ — well-understood examples are scarce, but we do know a number of general results; for example, classical Hilbert schemes are always connected [Hartshorne 1966]. In the case of general Hilbert schemes — those parametrizing subschemes of a fixed closed subscheme $X \subset \mathbb{P}^n$ that have some fixed Hilbert polynomial — even less is known. Our Fano schemes $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ provide an interesting family of Hilbert schemes whose study is tractable but whose geometry is still very rich.

Our second reason for studying $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ comes from geometric complexity theory. It is well known that permanents and determinants behave completely differently from the perspective of complexity theory. Indeed, computing the permanent of a square matrix is $\#P$-hard [Valiant 1979], while the determinant is computable in polynomial time. In fact, one of the central conjectures in complexity theory, due to Valiant [1979], posits that the permanent of an $n \times n$ matrix cannot be computed by affine linear substitution from the determinant of square matrix whose size is polynomial in n. Recently, Mulmuley and Sohoni developed an interesting representation-theoretic approach to this conjecture which is being pursued by a number of authors; see, e.g., [Bürgisser et al. 2011] and the references there. This new approach suggests that it is worthwhile to revisit the determinant and permanent from a geometric perspective. More specifically, linear spaces lying in the determinantal and permanental hypersurfaces in $\mathbb{P}^{n^2-1}$ play a particularly important role [Landsberg 2013, §5]. Indeed, the spaces $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ do exhibit a number of interesting differences, as we will see. (Actually, we began our study with the case $r = m = n$, which is the most interesting case for complexity theorists. We then realized that we could extend our techniques to $D^r_{m,n}$ and $P^r_{m,n}$, and that their Fano schemes have some interesting geometric features that cannot be seen just in the hypersurface case.)

Third, we are interested in contributing to the study of permanental ideals, i.e., the ideals defined by the $r \times r$ permanents of an $m \times n$ matrix. These ideals are much less well studied and behave very differently than determinantal ideals. For example, they can have many primary components, including embedded components, and in general their primary decompositions are not known, with a few nice exceptions, e.g., [Kirkup 2008; Laubenbacher and Swanson 2000]. In our paper, we are able to get new information about the linear spaces in these permanental schemes without computing their primary decompositions at all.

We next summarize the main results in this paper, starting with $D^r_{m,n}$ in Section 1A and turning to $P^r_{m,n}$ in Section 1B.

1A. ***The Fano scheme $F_k(D^r_{m,n})$.*** First, let us define two quantities which appear in our main theorems. Fix r, m, and n and, for $0 \le s \le r-1$, let

$$\kappa(s) = mn - (m-s)(s+n-r+1) - 1, \tag{1}$$

$$\text{and} \quad \delta(s) = (s+1+n-r)(r-s-1) + s(m-s). \tag{2}$$

The interpretation of $\kappa(s)$ is that it is the dimension of the linear space in $\mathbb{P}^{mn-1}$ of maps $\mathbb{K}^n \to \mathbb{K}^m$ sending a fixed $(s+n-r+1)$-dimensional subspace V of $\mathbb{K}^n$ into a fixed s-dimensional subspace W of $\mathbb{K}^m$. So, for example, $\kappa(s)+1$ is the dimension of the space of matrices with a zero block as shown:

$$\begin{array}{cc} & \begin{array}{cc} s{+}1{+}n{-}r & r{-}s{-}1 \end{array} \\ \begin{array}{c} m{-}s \\ \\ \\ s \\ \\ \end{array} & \left(\begin{array}{ccccccc} 0 & 0 & 0 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & 0 & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{array}\right) \end{array}. \tag{3}$$

These linear spaces were studied first in [Eisenbud and Harris 1988] and are very important to our analysis; see Definition 2.1. Next, $\delta(s)$ is simply the dimension of the product of Grassmannians $\mathrm{Gr}(s+n-r+1, n) \times \mathrm{Gr}(s, m)$ that parametrizes our choices of V and W.

Our first theorem on Fano schemes of determinants is a natural generalization of [Eisenbud and Harris 1988, Corollary 2.2] and gives the complete picture in the case $k = 1$ of the Fano scheme of lines.

Theorem 1.1. *The Fano scheme $\boldsymbol{F}_1(D^r_{m,n})$ has exactly r irreducible components, of dimensions*

$$\delta(s) + 2(\kappa(s) - 1) \quad \textit{for } 0 \le s \le r-1.$$

In particular, if $m = n$, then all irreducible components of $\boldsymbol{F}_1(D^r_{n,n})$ have dimension $(n-r)(r-2) + 2nr - n - 5$ If $r > 2$, then all components intersect pairwise. Furthermore, if $m = n = r$, then $\boldsymbol{F}_1(D^n_{n,n})$ has the expected dimension and is a reduced local complete intersection.

We prove this theorem in Section 6. We also explicitly calculate the degree of $\boldsymbol{F}_1(D^n_{n,n})$ as a subscheme of $\mathrm{Gr}(2, n^2)$ in its Plücker embedding for n up to 6; see Proposition 6.3.

n	2	3	4	5	6	7	8	9	10
nonempty $\iff k \le$	1	5	11	19	29	41	55	71	89
singular $\iff k \le$	–	3	8	15	24	35	48	63	80
connected $\iff k \le$	–	3	8	13	21	29	40	51	65
or $k =$					24	35	46–48	57, 60–63	72, 73, 76–80

Table 1. Properties of the Fano scheme $\boldsymbol{F}_k(D^n_{n,n})$ for $n \le 10$.

For higher values of k, we do not have a complete characterization of $\boldsymbol{F}_k(D^r_{m,n})$, but we can still say a lot about these schemes. The first thing we should note is that the Fano scheme $\boldsymbol{F}_k(D^r_{m,n})$ is nonempty if and only if $k < (r-1)n$; this follows from a result of Dieudonné [1949] (see also Section 2 for a quick proof). For these values of k, we can say exactly which of the schemes $\boldsymbol{F}_k(D^r_{m,n})$ are smooth and irreducible.

Theorem 1.2. *Let* $1 \le k < (r-1)n$.

(a) *The Fano scheme* $\boldsymbol{F}_k(D^r_{m,n})$ *is smooth if and only if* $k > (r-2)n$.

(b) $\boldsymbol{F}_k(D^r_{m,n})$ *is irreducible if and only if* $m \ne n$ *and* $k > (r-2)n + m - r + 1$.

See Section 5A for the proof.

Our next main result says exactly when $\boldsymbol{F}_k(D^m_{m,n})$ is connected.

Theorem 1.3. *Suppose that* $1 \le k < (m-1)n$. *Then* $\boldsymbol{F}_k(D^m_{m,n})$ *is disconnected if and only if either*

$$m^2 - 2m < k \le \kappa(0)$$

or there exists an integer s *with* $0 < s < m-1$ *such that*

$$\kappa(s) - \min\{m-s-1, n-m+s\} < k \le \kappa(s).$$

See Section 5A for the proof. Table 1 illustrates the results of Theorems 1.2 and 1.3 for $\boldsymbol{F}_k(D^n_{n,n})$ for $n \le 10$.

Theorem 1.3 is very surprising. It says that connectivity of $\boldsymbol{F}_k(D^m_{m,n})$ can actually be highly nonmonotonic as k varies from 1 to $(m-1)n-1$, since each of the $m-2$ values of s above cuts out an interval of values of k for which $\boldsymbol{F}_k(D^m_{m,n})$ is disconnected. The last row of Table 1 gives examples of this behavior. That said, $\boldsymbol{F}_k(D^m_{m,n})$ is always connected if k is sufficiently small, as we show in Corollary 5.5.

As for connectedness of $\boldsymbol{F}_k(D^r_{m,n})$ when $r < m$, we can still give some necessary conditions and sufficient conditions, although they are not strong enough to completely characterize when $\boldsymbol{F}_k(D^r_{m,n})$ is connected. See Theorem 5.3.

In the next section, we will define a family of special irreducible components $\mathfrak{C}_k(s)$ of $\boldsymbol{F}_k(D^r_{m,n})$ that we call *compression components*. These components are very important to our analysis of $\boldsymbol{F}_k(D^r_{m,n})$. Indeed, we will see that if $k = 1$ or if k is sufficiently large then $\boldsymbol{F}_k(D^r_{m,n})$ consists *only* of compression components (Theorem 1.1 and Corollary 5.1, respectively).

For general k, however, other components may appear, as detected in [Eisenbud and Harris 1988, Theorem 1.1]. For example, we will see in Section 7 that a noncompression component already appears in $\boldsymbol{F}_2(D^3_{3,3})$: it is the component $\mathfrak{C}^*$ of 2-planes of matrices $(\mathrm{GL}_3 \times \mathrm{GL}_3)$-equivalent to the space of 3×3 antisymmetric matrices. On the other hand, the compression components still have the desirable property that every component of $\boldsymbol{F}_k(D^r_{m,n})$ must meet one of them (Remark 2.5). This fact makes it very useful to study local neighborhoods of points on these components.

1B. ***The Fano scheme $\boldsymbol{F}_k(P^r_{m,n})$.*** The study of $\boldsymbol{F}_k(P^r_{m,n})$ is slightly more delicate due to the lack of $\mathrm{GL}_m \times \mathrm{GL}_n$ symmetry. Nonetheless, we can prove several results on $\boldsymbol{F}_k(P^r_{m,n})$ that have a similar flavor to those for $\boldsymbol{F}_k(D^r_{m,n})$. (When we discuss $P^r_{m,n}$, we will always assume that $\mathrm{char}(\mathbb{K}) \neq 2$, since otherwise $P^r_{m,n} = D^r_{m,n}$ and our above results apply.)

For example, just as in the determinantal case, $\boldsymbol{F}_k(P^r_{m,n})$ is nonempty if and only if $k < (r-1)n$ (Proposition 2.6). Furthermore, in that range, we can completely characterize when $\boldsymbol{F}_k(P^r_{m,n})$ is smooth:

Theorem 1.4. *Let $1 \leq k < (r-1)n$. The Fano scheme $\boldsymbol{F}_k(P^r_{m,n})$ is smooth if and only if $n = 2$ or $k > (r-2)n + 1$.*

However, if $\boldsymbol{F}_k(P^r_{m,n})$ is nonempty, it is *never* irreducible, as we prove in Proposition 4.6. See Table 2 for a summary of our results applied to the case of the $P^n_{n,n}$ for $n \leq 9$.

We can also give necessary conditions and sufficient conditions for the connectedness of $\boldsymbol{F}_k(P^r_{m,n})$; see Theorem 5.7. In particular, we show that if k is sufficiently small, then $\boldsymbol{F}_k(P^r_{m,n})$ is connected (Corollary 5.8). However, we do not know how to completely characterize when $\boldsymbol{F}_k(P^r_{m,n})$ is connected, even in the case $r = m$.

As in the case of the determinant, there is a family of *compression subschemes* $\mathfrak{C}_k(\sigma, \tau)$ of $\boldsymbol{F}_k(P^r_{m,n})$, defined in Section 2, that again play an important role in our analysis. For example, any irreducible component of $\boldsymbol{F}_k(P^r_{m,n})$ must intersect a subscheme of the form $\mathfrak{C}_k(\sigma, \tau)$; see Remark 2.5. We prove in Proposition 4.4 that these subschemes are (with a few necessary exceptions) actually components; and we show (Corollary 5.6) that if k is large enough then $\boldsymbol{F}_k(P^r_{m,n})$ is simply a disjoint union of compression components.

In general, however, $\boldsymbol{F}_k(P^r_{m,n})$ has more components than just those of the form $\mathfrak{C}_k(\sigma, \tau)$. In Section 7B, we consider in depth the example of $\boldsymbol{F}_k(P^3_{3,3})$ for various values of k. In particular, we give a complete description of $\boldsymbol{F}_4(P^3_{3,3})$: it consists of nine copies of $\mathbb{P}^1 \times \mathbb{P}^1$, six of $\mathbb{P}^5$, 18 Hirzebruch surfaces $\mathscr{F}_1$, and 36 embedded fat points, arranged into a total of three connected components;

n	2	3	4	5	6	7	8	9
nonempty $\iff k \le$	1	5	11	19	29	41	55	71
singular $\iff k \le$	–	4	9	16	25	36	49	64
connected if $k \le$	–	3	8	13	21	29	40	51
or $k =$					24	34–35	46–48	57, 60–63
disconnected if $k =$	1	4, 5	10, 11	17–19	26–29	37–41	42, 43, 45, 50–55	53–56, 59, 65–71

Table 2. Properties of the Fano scheme $\boldsymbol{F}_k(P^n_{n,n})$ for $n \le 9$.

see Proposition 7.4. In particular, we'll see that $\boldsymbol{F}_4(P^3_{3,3})$ has components not of the form $\mathfrak{C}_4(\sigma, \tau)$, and that this Fano scheme can have embedded primary components.

1C. ***Related work and organization.*** Let us now mention some related research. Fano schemes have been studied in a variety of contexts. The first modern treatment was given in [Altman and Kleiman 1977], where it was proven that the Fano scheme of lines on a smooth cubic hypersurface of dimension at least three is smooth and connected. Sufficient criteria for smoothness and connectedness of Fano schemes of lines on hypersurfaces were given in [Barth and van de Ven 1978], and generalized to higher-dimensional linear spaces in [Langer 1997]. While Fano schemes for generic hypersurfaces X always have dimension equal to the so-called expected dimension, it was proven in [Harris et al. 1998] that this is also true for all smooth hypersurfaces of sufficiently low degree. General properties of Fano schemes of linear spaces on complete intersections have been studied in [Debarre and Manivel 1998].

In the case $r = m = n$, $D^n_{n,n}$ and $P^n_{n,n}$ are both irreducible hypersurfaces. However, the above-mentioned results do not apply to these hypersurfaces, since they are not general, and their degrees are not sufficiently low.

There has also been considerable work classifying the $(\mathrm{GL}_m \times \mathrm{GL}_n)$-equivalence classes of linear spaces of nonfull rank $m \times n$ matrices, i.e., the $\mathrm{GL}_m \times \mathrm{GL}_n$ orbits in the Fano schemes $\boldsymbol{F}_k(D^r_{m,n})$. Such a classification exists for k large relative to m, n, and r [Beasley 1987; de Seguins Pazzis 2013], and for $r \le 4$ [Atkinson 1983; Eisenbud and Harris 1988]. These classifications only apply to a limited range of k, m, n, and r, however, and do not fully describe the geometry of $\boldsymbol{F}_k(D^r_{m,n})$. See also, e.g., [Draisma 2006] for constructions of certain maximal linear spaces of singular matrices. Linear spaces of skew-symmetric matrices have also been studied; see, e.g., [Manivel and Mezzetti 2005], where 2-planes of 6×6 rank four skew-symmetric matrices have been classified. We do not know of any previous work studying $\boldsymbol{F}_k(P^r_{m,n})$.

We use two main tools to study $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$. First, there are natural torus actions on both of these Fano schemes; we are able to get a lot of geometric information by studying the local structure of these Fano schemes at torus fixed points. Torus actions and fixed points are discussed in Section 2; some of our local computations were carried out explicitly using the `Macaulay2` package [Ilten 2012] in Section 7. Our second main tool consists of using deformation theory to calculate tangent space dimensions at select points of the Fano schemes. We carry out these calculations in Section 3.

In Section 4, we study compression spaces in more detail. We put a lot of pieces together in Section 5 to discuss irreducibility, smoothness, and connectedness, in particular completing the proofs of Theorems 1.2, 1.3, and 1.4.

Section 6 is concerned with the Fano scheme of lines; we prove Theorem 1.1 and compute the degrees of $\boldsymbol{F}_1(D^n_{n,n})$ for $n \le 6$. In Section 7, we discuss the explicit examples of $\boldsymbol{F}_k(D^3_{3,3})$ and $\boldsymbol{F}_k(P^3_{3,3})$. In particular, we show that in general neither $\boldsymbol{F}_k(D^r_{m,n})$ nor $\boldsymbol{F}_k(P^r_{m,n})$ are reduced. In Section 8, we conclude with a comparison between the cases of determinants and permanents and present some conjectures and further questions on the structure of our Fano schemes.

2. Compression spaces and torus fixed points

2A. ***Compression spaces.*** The first thing we would like to do is to define some very important subschemes of $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ arising from *compression spaces.*

Definition 2.1 (compare [Eisenbud and Harris 1988]). Fix a natural number $0 \le s \le r-1$. An *s-compression space* is the space of all $m \times n$ matrices over $\mathbb{K}$ sending a fixed $(s+1+n-r)$-dimensional subspace V of $\mathbb{K}^n$ to a fixed s-dimensional subspace W of $\mathbb{K}^m$.

See (3) for an example. Now, every r-dimensional subspace of $\mathbb{K}^n$ meets V in dimension at least $s+1$, so a matrix in an s-compression space "compresses" the image of each r-dimensional subspace into an $(r-1)$-dimensional subspace. In other words, every matrix in a compression space has rank less than r, and we may view this compression space as a point of the Fano scheme $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$, where κ was defined in (1). The set of all s-compression spaces forms a closed subscheme $\mathfrak{C}(s)$ of $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$ of dimension $\delta(s)$. In fact, we have that $\mathfrak{C}(s)$ is isomorphic to

$$\mathrm{Gr}(s+n-r+1, n) \times \mathrm{Gr}(s, m).$$

Indeed, the morphism sending a pair (A, B) of $(s+n-r+1)$- and s-dimensional planes in $\mathbb{K}^n$ and $\mathbb{K}^m$ to the space of matrices which map A to B is bijective, and an inspection of the natural affine charts of $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$ shows that this map is an isomorphism. In Theorem 4.7, we will even compute the degree of $\mathfrak{C}(s)$ as a subscheme of the Grassmannian $\mathrm{Gr}(\kappa(s)+1, mn)$ in its Plücker embedding.

Next, for each $k \le \kappa(s)$, the subscheme $\mathfrak{C}(s)$ of $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$ induces a subscheme $\mathfrak{C}_k(s)$ of $\boldsymbol{F}_k(D^r_{m,n})$ whose points correspond to the k-planes sitting inside some s-compression space. More precisely, let $\mathcal{U}(s)$ denote the restriction of the universal bundle on $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$ to $\mathfrak{C}(s)$. Then there is a natural morphism

$$\rho_k(s) : \mathrm{Gr}(k+1, \mathcal{U}(s)) \to \boldsymbol{F}_k(D^r_{m,n})$$

from the Grassmann bundle $\mathrm{Gr}(k+1, \mathcal{U}(s))$ to $\boldsymbol{F}_k(D^r_{m,n})$ sending a $(k+1)$-dimensional subspace of $\mathbb{K}^{mn}$ to the corresponding point of $\boldsymbol{F}_k(D^r_{m,n})$. We denote the image of this map by $\mathfrak{C}_k(s)$. We call $\mathfrak{C}_k(s)$ an s-compression component of $\boldsymbol{F}_k(D^r_{m,n})$. We will see in Theorem 4.1 and Corollary 4.3 that $\mathfrak{C}_k(s)$ is in fact an irreducible component of $\boldsymbol{F}_k(D^r_{m,n})$.

The permanental case. In contrast to the determinantal case, not every s-compression space consists of matrices with vanishing $r \times r$ permanents. However, the ones that actually correspond to matrices with a fixed $(s+1+n-r) \times (m-s)$ zero submatrix do correspond to points of $\boldsymbol{F}_k(P^r_{m,n})$. We call these *standard* compression spaces. More precisely:

Definition 2.2. Let σ and τ be subsets of $\{1, \dots, n\}$ and $\{1, \dots, m\}$ such that $|\sigma| - |\tau| = n - r + 1$. Let $\mathfrak{C}(\sigma, \tau)$ denote the compression space consisting of those linear maps $\mathbb{K}^n \to \mathbb{K}^m$ which send the standard basis vectors indexed by elements of σ to the subspace of $\mathbb{K}^m$ generated by standard basis vectors indexed by elements of τ. Such compression spaces are called *standard compression spaces*.

An example of a standard compression space is shown in (3). There, $\sigma = \{1, \dots, s+1+n-r\}$ and $\tau = \{m-s+1, \dots, m\}$. A straightforward calculation contained in Proposition 2.3 shows that every point of $\mathfrak{C}(\sigma, \tau)$ lies in $P^r_{m,n}$, and hence $\mathfrak{C}(\sigma, \tau)$ corresponds to a point of $\boldsymbol{F}_k(P^r_{m,n})$ for $k = \kappa(|\tau|)$.

Just as in the determinantal case, $\mathfrak{C}(\sigma, \tau)$ induces subschemes of Fano schemes $\boldsymbol{F}_k(D^r_{m,n})$ for $k \le \kappa(|\tau|)$. Regarding $\mathfrak{C}(\sigma, \tau)$ as a vector space, there is an obvious map $\mathrm{Gr}(k+1, \mathfrak{C}(\sigma, \tau)) \cong Gr(k+1, \kappa(s)+1) \to \mathrm{Gr}(k+1, mn)$ which is an isomorphism onto its image. We call this image $\mathfrak{C}_k(\sigma, \tau)$. But the image of the map is clearly contained in $F_k(P^r_{m,n})$; that is, $\mathfrak{C}_k(\sigma, \tau)$ is a subscheme of $F_k(P^r_{m,n})$ isomorphic to a Grassmannian. We will give conditions in Proposition 4.4 under which $\mathfrak{C}_k(\sigma, \tau)$ is actually a component of $\boldsymbol{F}_k(P^r_{m,n})$.

2B. ***Torus fixed points.*** Next, we would like to take advantage of a natural torus action on our Fano schemes. Let us view points of $\mathbb{P}^{mn-1}$ as $m \times n$ matrices over $\mathbb{K}$, up to simultaneous rescaling. The group $\mathrm{GL}_m \times \mathrm{GL}_n$ then acts naturally on $\mathbb{P}^{mn-1}$: the first factor by inverse matrix multiplication the right, and the second factor by matrix multiplication on the left. (We will assume throughout this paper that $m \geq 2$, since otherwise $D^r_{m,n}$ and $P^r_{m,n}$ become trivial.)

$D^r_{m,n}$ is invariant under this action, but $P^r_{m,n}$ is not. However, both are invariant under the action of the subgroup $T_m \times T_n$ of $\mathrm{GL}_m \times \mathrm{GL}_n$. Here, $T_m \cong (\mathbb{K}^*)^m$ and $T_n \cong (\mathbb{K}^*)^n$ are the standard diagonal tori of GL_m and GL_n that act by rescaling the rows and columns, respectively, of an $m \times n$ matrix. This action of $T_m \times T_n$ extends naturally to actions on the Grassmannian $\mathrm{Gr}(k+1, mn)$ and its subschemes $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$.

When is a closed point of $\mathrm{Gr}(k+1, mn)$ a torus fixed point? Such a point Q corresponds to a $(k+1)$-dimensional linear subspace of $m \times n$ matrices. But any linear relation on the entries that has more than one term cannot be preserved under all possible rescaling of rows and columns. So Q must consist of all matrices obtained by setting all but $k+1$ specified entries to zero. We often represent such a torus fixed point by a matrix with zeroes and $*$s (or zeroes and blanks), where a $*$ denotes an entry that can vary freely.

We now prove that, of the torus fixed points in $\mathrm{Gr}(k+1, mn)$, the ones that lie in $\boldsymbol{F}_k(D^r_{m,n})$ and in $\boldsymbol{F}_k(P^r_{m,n})$ are precisely the ones that are subspaces of *standard compression spaces* (Definition 2.1).

Proposition 2.3. *Let Q be a $(T_m \times T_n)$-fixed point of* $\mathrm{Gr}(k+1, mn)$. *Then the following are equivalent*:

(a) *Q is contained in $\boldsymbol{F}_k(D^r_{m,n})$.*

(b) *Q is contained in $\boldsymbol{F}_k(P^r_{m,n})$.*

(c) *The linear space of matrices corresponding to Q is a subspace of a standard compression space.*

Proof. Represent Q as a matrix of $k+1$ starred entries and $mn-k-1$ zero entries, as explained above. Make a bipartite graph G on vertex sets $V = \{1, \dots, m\}$ and $V' = \{1, \dots, n\}$ with an edge from $i \in V$ to $j \in V'$ whenever entry ij is starred. Now, let us prove that a series of statements are equivalent, the first one being that Q is contained in $\boldsymbol{F}_k(D^r_{m,n})$ (or $\boldsymbol{F}_k(P^r_{m,n})$) and the last being that Q is a subspace of a standard compression space.

First, Q is contained in $\boldsymbol{F}_k(D^r_{m,n})$ (respectively $\boldsymbol{F}_k(P^r_{m,n})$) if and only if, for every subset W of V of size r, G fails to have a matching of W into V'. After all, if G fails to have such a matching, then every term from each $r \times r$ determinant has a zero entry in it; whereas, if G has such a matching, then specializing the r entries corresponding to the edges in that matching to 1, and setting all others to zero, gives a nonzero $r \times r$ determinant (or permanent).

Second, by Lemma 2.4 below, G fails to have a matching of such a $W \subset V$ into V' if and only if there exist some $m-s$ vertices in V with only $r-s-1$ neighbors amongst the vertices in V'. But the existence of some $m-s$ vertices in V with only $r-s-1$ neighbors in V' means precisely that Q has an $(m-s)\times(n-r+s+1)$ block of zeroes, i.e., Q is a subspace of a standard s-compression space. □

Lemma 2.4. *Let $G=(V,V',E)$ be a bipartite graph and r a number with $1\le r\le|V|$. Then the following are equivalent*:

(a) *There is no subset $W\subset V$ with $|W|=r$ which has a matching into V'.*

(b) *There is a subset $S\subset V$ having only $r+|S|-|V|-1$ neighbors in V'.*

Proof. Add $|V|-r$ vertices U to G, each of which is adjacent to every vertex in V. Then condition (a) above is equivalent to V having no matching into $U\cup V'$. By Hall's marriage theorem, this is equivalent to the existence of a subset S of vertices in V with fewer than $|S|$ neighbors in $U\cup V'$. Now, these S vertices have $|V|-r$ neighbors in U, hence fewer than $|S|+r-|V|$ neighbors in V'. □

Remark 2.5. Proposition 2.3 implies that any irreducible component Z of $\boldsymbol{F}_k(D^r_{m,n})$ or $\boldsymbol{F}_k(P^r_{m,n})$ must intersect a subscheme of the form $\mathfrak{C}_k(s)$ or $\mathfrak{C}_k(\sigma,\tau)$, respectively. Indeed, our Fano schemes are projective, hence Z is as well, and thus must contain a torus fixed point. But any torus fixed point is contained in a subscheme of the desired form. This simple observation will play an important role in our proofs.

Finally, we can easily prove when $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ are nonempty just by looking at torus fixed points. See [Dieudonné 1949] for the determinantal case.

Proposition 2.6. *The Fano schemes $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ are nonempty if and only if $k<(r-1)n$.*

Proof. The schemes $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ are nonempty if and only if $k\le\kappa(s)$ for some $s\in\{0,\dots,r-1\}$. After all, if $k\le\kappa(s)$ for some s, then $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ each contain a torus fixed point that is a subspace of a standard s-compression space; but, if not, then $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ have no torus fixed points, so are empty.

Now, we have

$$\kappa(0)=(r-1)m-1 \quad\text{and}\quad \kappa(r-1)=(r-1)n-1.$$

So, if $k<(r-1)n$, then $k\le\kappa(r-1)$ as desired. For the converse, if $k\ge(r-1)n$ then $k>\kappa(0)$ and $k>\kappa(r-1)$. Then, by convexity of κ, we conclude that $k>\kappa(s)$ for all s. □

3. Tangent space calculations

The next thing we would like to do is to calculate the dimension of the tangent space of $\boldsymbol{F}_k(D^r_{m,n})$ and of $\boldsymbol{F}_k(P^r_{m,n})$ at some carefully chosen points. The points we will look at lie in $\mathfrak{C}_k(s)$ or $\mathfrak{C}_k(\sigma,\tau)$, i.e., they are k-planes inside compression spaces; and we will choose them to have a particular form that is favorable to making explicit computations using deformation theory.

Our main results in this section (Theorem 3.1 for the determinant and Theorem 3.2 for the permanent) will each have two parts. First, the tangent space at a *general* point of $\mathfrak{C}_k(s)$ or $\mathfrak{C}_k(\sigma, \tau)$ can be no larger than its dimension at the points we study, since tangent space dimensions are upper semicontinuous. Second, we will show that if certain additional inequalities hold, then every torus fixed point actually has the special form that allows our computation to go through. We can therefore conclude that tangent space dimension at *every* point of $\mathfrak{C}_k(s)$ or $\mathfrak{C}_k(\sigma, \tau)$ can be no larger than the specified dimension.

This section is the technical heart of the paper, and we will reap the rewards in Sections 4 and 5 when we use these tangent space calculations to prove theorems on smoothness and connectedness of $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$.

Here are the two main theorems.

Theorem 3.1. *Fix integers k and s with $1 \le k \le \kappa(s)$ and $0 \le s \le r-1$.*

(i) *For a general point $\eta \in \mathfrak{C}_k(s)$,*

$$\dim T_\eta \boldsymbol{F}_k(D^r_{m,n}) \le \delta(s) + (k+1)(\kappa(s)-k),$$

where $\kappa(s)$ and $\delta(s)$ are defined in (1) and (2).

(ii) *If furthermore both of the conditions*

$$k > \kappa(s) - (m-s-1) \qquad \text{if } s \ne r-1, \tag{4}$$

$$k > \kappa(s) - (n-r+s) \qquad \text{if } s \ne 0 \tag{5}$$

hold, then the dimension bound holds for every *point $\eta \in \mathfrak{C}_k(s)$.*

We have the following analogous result for the permanent:

Theorem 3.2. *Fix integers k and s such that*

$$s = 0,\ r-1 \quad \text{and} \quad 2 \le k \le \kappa(s),$$
$$\text{or} \quad 1 \le s \le r-2 \quad \text{and} \quad 5 \le k \le \kappa(s).$$

Suppose further that

$$s+1+n-r \ge 3 \quad \text{if } s \ne 0 \qquad \text{and} \qquad m-s \ge 3 \quad \text{if } s \ne r-1.$$

Consider a standard compression space $\mathfrak{C}(\sigma, \tau)$ (Definition 2.2) with $|\tau| = s$.

(i) *For a general point $\eta \in \mathfrak{C}_k(\sigma, \tau)$,*

$$\dim T_\eta \boldsymbol{F}_k(P^r_{m,n}) \le (k+1)(\kappa(s)-k).$$

(ii) *If furthermore both of the conditions*

$$k > \kappa(s) - (m-s-2) \qquad \textit{if } s \neq r-1, \tag{6}$$

$$k > \kappa(s) - (n-r+s-1) \qquad \textit{if } s \neq 0 \tag{7}$$

hold, then the dimension bound holds for every *point* $\eta \in \mathfrak{C}_k(\sigma, \tau)$.

Remark 3.3. For Theorem 3.1, the bounds on k in (4) and (5) are sharp. Indeed, if k doesn't satisfy both bounds, then there are torus fixed points in $\mathfrak{C}_k(s)$ which also lie in $\mathfrak{C}_k(s-1)$ or $\mathfrak{C}_k(s+1)$. Since these schemes are not equal by Proposition 4.5, and the dimension of $\mathfrak{C}_k(s)$ is $\delta(s)+(k+1)(\kappa(s)-k)$ by Corollary 4.3, the tangent space dimension at such points must be higher.

We will prove our two main theorems in parallel below, because much of the setup is the same.

We start by recalling a completely general characterization of tangent spaces to Fano schemes. Suppose $X \subset \mathbb{P}^N_K$ is any projective scheme with homogeneous ideal $I \subset S = K[x_0, \ldots, x_N]$, and $\Lambda \subset \mathbb{P}^N$ is a k-plane contained in X defined by the homogeneous ideal $J \subset S$. Write $[\Lambda]$ for the corresponding point of $\boldsymbol{F}_k(X)$ and let $\mathcal{J}$ be the ideal sheaf of Λ. Then the tangent space $T_{[\Lambda]}\boldsymbol{F}_k(X)$ is isomorphic to the space of first-order deformations of Λ in X, which is well known to be isomorphic to $\operatorname{Hom}_{\mathcal{O}_X}(\mathcal{J}, \mathcal{O}_\Lambda)$; see, e.g., [Hartshorne 2010, Theorem 2.4]. In our case, J is generated by linear forms, so J/I as well as S/J are saturated, graded S/I-modules. So $T_{[\Lambda]}\boldsymbol{F}_k(X) \cong \operatorname{Hom}_{\mathcal{O}_X}(\mathcal{J}, \mathcal{O}_\Lambda) \cong \operatorname{Hom}_{S/I}(J/I, S/J)_0$, the space of degree-preserving maps of S/I-modules.

In our specific situation, fix r, m, and n, and let $S = \mathbb{K}[x_{1,1}, \ldots, x_{m,n}]$. Suppose that $J \subset S$ is a linear ideal defining a k-plane $\Lambda \subset \mathbb{P}^{mn-1}$ of matrices lying in the standard s-compression space shown in (3). Let us choose standard linear monomials $z_0, \ldots, z_k$ for S/J; that is, we specify an isomorphism $S/J \cong \mathbb{K}[z_0, \ldots, z_k]$. Then we can regard Λ as a matrix whose entries are linear forms in $z_0, \ldots, z_k$ such that the linear span of these forms has full dimension $k+1$. In other words, Λ has the form

$$\Lambda = \left(\begin{array}{c|c} 0 & B \\ \hline C & D \end{array}\right), \tag{8}$$

where the upper left block of zeroes has size $(m-s)\times(s+1+n-r)$ and the entries of B, C, and D are in $\mathbb{K}[z_0, \ldots, z_k]_1$. See Example 3.5 for a matrix of this form.

The next lemma tells us concretely how to compute the dimensions of the tangent spaces of $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ at the point $[\Lambda]$. Call an $r \times r$ submatrix of an $m \times n$ matrix *anchored* if it involves the last s rows as well as the last $r-s-1$ columns.

Lemma 3.4. *For a k-plane* Λ *as above, let* a_{det} (*respectively* a_{perm}) *be the* $\mathbb{K}$-*dimension of the space of* $(m-s)\times(s+1+n-r)$ *matrices* A *with entries in* $\mathbb{K}[z_0,\dots,z_k]_1$ *such that the* $r\times r$ *anchored determinants* (*respectively permanents*) *of the* $m\times n$ *matrix*

$$Q=\left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right) \tag{9}$$

all vanish. Then

$$\dim T_{[\Lambda]}\boldsymbol{F}_k(D^r_{m,n}) = a_{\text{det}} + (k+1)(\kappa(s)-k),$$
$$\dim T_{[\Lambda]}\boldsymbol{F}_k(P^r_{m,n}) = a_{\text{perm}} + (k+1)(\kappa(s)-k).$$

Proof. Let $I_{\text{det}}\subset S$ and $I_{\text{perm}}\subset S$ be the ideals generated by the determinants and permanents, respectively, of the $r\times r$ submatrices of the $m\times n$ matrix $(x_{i,j})$. We will treat both cases simultaneously. Let I be I_{det} or I_{perm}, and let $J\subset S$ be the linear ideal of Λ. Then J is generated by the coordinates $x_{i,j}$ with $i\le m-s$ and $j\le s+1+n-r$, corresponding to the condition that the upper left block is zero, along with $\kappa(s)-k$ additional independent linear forms f_i on the remaining $x_{i,j}$ that determine the linear relations on the other entries of the matrix. We wish to compute the dimension of $\mathrm{Hom}_{S/I}(J/I,\, S/J)_0$.

To give an element of $\mathrm{Hom}_{S/I}(J/I,\, S/J)_0$, we must specify an element of $(S/J)_1$, i.e., a linear form in $z_0,\dots,z_k$, for each generator $x_{i,j}$ and f_i. This choice of elements of $(S/J)_1$ is subject to the condition that all the relations among the $x_{i,j}$ and the f_i are preserved among their images. Let's examine these relations.

Assume for a moment that $r=m=n$, and suppose we have a map ϕ in $\mathrm{Hom}_{S/I}(J/I,\, S/J)_0$. Among the relations on generators of J/I, there are the Koszul relations of the form $a_1(a_2)-a_2(a_1)$ for $a_i\in J$, which ϕ obviously sends to 0 mod J. Apart from these, then, the only nontrivial relation among generators of J/I comes from writing the determinant or permanent of the $r\times r$ matrix $(x_{i,j})$ as an S-linear combination of the generators of J. In fact, it is easy to choose such an S-linear combination (uniquely up to the Koszul relations), since every term of the $r\times r$ determinant (or permanent) contains at least one position in the $\mathbf{0}$ block and hence one generator of J. Then the fact that ϕ sends the $r\times r$ determinant or permanent to zero means that

$$\sum_{i,j}\pm(b_ic_j)\phi(x_{i,j})=0, \tag{10}$$

where the sum above has one term for each entry $(i,\, j)$ in the upper left zero block. Here, b_i is the determinant or permanent of the submatrix of B gotten by deleting the i-th row, and c_j is the determinant or permanent of the submatrix of C gotten by deleting the j-th column. The reason that no other terms of the determinant or permanent appear is that every other term contains at least two positions in the $\mathbf{0}$ block, and hence, after factoring out some generator of J, will still vanish modulo J. So (10) is the only constraint on the images of the generators of J.

If instead $r < n$, we simply note that the nontrivial relations among the images of the $x_{i,j}$ are just the relation (10) obtained by replacing (8) by each of its $r \times r$ anchored submatrices. (The nonanchored submatrices do not add any constraints, because each term of a nonanchored $r \times r$ determinant or permanent contains at least two positions in the $\mathbf{0}$ block and so is necessarily sent to 0 in S/J.) Then condition (10) applied to the $r \times r$ submatrices of Λ can be rewritten as the condition that the determinants or permanents of the $r \times r$ anchored submatrices of the $m \times n$ matrix in (9) vanish, where $A = (\phi(x_{i,j}))$.

Now we can easily count dimensions for $\mathrm{Hom}_{S/I}(J/I, S/J)_0$. We have just shown that the dimension of the space of choices of $\phi(x_{i,j})$ for $i \leq m-s$ and $j \leq s+1+n-r$ is precisely $a_{\det}$ (respectively a_{perm}). Next, we have exactly $k+1$ dimensions for choosing each $\phi(f_i)$, since after all the f_i do not appear at all in (10) and $\dim(S/J)_1 = k+1$. So the dimension of space of choices for the $\phi(f_i)$ is $(k+1)(\kappa(s)-k)$. Adding, we have proved Lemma 3.4. □

We now prove part (i) of each theorem by applying Lemma 3.4 to some carefully chosen points $[\Lambda]$.

Proof of Theorem 3.1(i). We need to exhibit, for any k and s as in the statement of Theorem 3.1, a matrix of the form (8) such that the dimension $a_{\det}$ defined in Lemma 3.4 is exactly $\delta(s)$. Then Theorem 3.1(i) would follow by upper semicontinuity.

First, some notation: for distinct variables z_i and z_j, define matrices

$$B(z_i, z_j) = \begin{pmatrix} z_i & & \\ z_j & \ddots & \\ & \ddots & z_i \\ & & z_j \\ \hline & 0 & \end{pmatrix} \quad \text{and} \quad C(z_i, z_j) = \left(\begin{array}{cccc|c} z_i & z_j & & & \\ & \ddots & \ddots & & 0 \\ & & z_i & z_j & \end{array}\right) \tag{11}$$

of dimensions $(m-s) \times (r-s-1)$ and $s \times (s+1+n-r)$, respectively. (Thus, by slight abuse of notation, the dimensions will depend on s.) Now let's specify the form that Λ should take, in four cases.

If $s = 0$, so the matrices C and D don't appear, we let $B = B(z_0, z_1) + B'$, where B' is any matrix whose entries are linear forms not involving z_0 or z_1 and such that the linear span of the forms appearing in B has full dimension $k+1$.

Similarly, if $s = r-1$, so that the matrices B and D don't appear, then we let $C = C(z_0, z_1) + C'$, where C' is any matrix whose entries are linear forms not involving z_0 or z_1 and such that the linear span of the forms appearing in C has full dimension $k+1$.

Next, if $s \neq 0, r-1$, and $k = 1$, we let $B = B(z_0, z_1)$ and $C = C(z_0, z_1)$. There are no restrictions on D.

Finally, if $s \neq 0, r-1$, and $k \geq 2$, we let $B = B(z_0, z_1) + B'$ and $C = C(z_1, z_2) + C'$, for any matrices B' and C' that do not involve z_0, z_1, or z_2, and furthermore do not involve any variables in common. There are again no restrictions on D, except that B', C', and D must be chosen so that the linear span of the forms appearing in Λ has full dimension $k+1$.

Example 3.5. For $k = 4$, $s = 2$, $r = m = 5$, and $n = 6$, the following 5×6 matrix over $\mathbb{K} = \mathbb{C}$, say, has the desired form. The $*$ entries represent arbitrary linear forms in z_0, z_1, z_2, z_3, and z_4.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & z_0 & z_3 \\ 0 & 0 & 0 & 0 & z_1 - z_3 & z_0 + 3z_3 \\ 0 & 0 & 0 & 0 & 5z_3 & z_1 + 11z_3 \\ z_1 & z_2 & 0 & 0 & * & * \\ 0 & z_1 & z_2 & z_4 & * & * \end{pmatrix}.$$

Now, in each case, we are done if we can show that the space of matrices A satisfying that the $r \times r$ anchored minors of the matrix Q in (9) vanish has dimension exactly $\delta(s)$. In fact, we can identify a space of dimension $\delta(s)$: take A to be any sum of a matrix in the $\mathbb{K}$-column span of B and a matrix in the $\mathbb{K}$-row span of C. The resulting matrix Q still has rank less than r. Furthermore, B and C have full rank, but on the other hand no nonzero matrix is both in the $\mathbb{K}$-column span of B and the $\mathbb{K}$-row span of C. So the space of matrices A obtained in this way forms a $\mathbb{K}$-vector space of dimension exactly $(r-s-1)(s+1+n-r) + s(m-s) = \delta(s)$. We have thus reduced to proving the converse: if the anchored $r \times r$ minors of Q vanish, then A is the sum of a matrix in the $\mathbb{K}$-column span of B and a matrix in the $\mathbb{K}$-row span of C.

The degenerate cases $s = 0, r-1$ are thus an immediate consequence of the lemma below, whose proof we postpone to the end of the section.

Lemma 3.6. *Let $2 \leq m \leq n$. Consider an $m \times n$ matrix*

$$P = \left(\begin{array}{c} \vec{a} \\ \hline \\ C \\ \\ \end{array} \right)$$

whose entries are linear forms $z_0, \dots, z_k$, where $\vec{a} = (a_i)$ is a $1 \times n$ matrix and C is an $(m-1) \times n$ matrix. Suppose $C = C(z_0, z_1) + C'$, where $C(z_0, z_1)$ is an $(m-1) \times n$ matrix as shown in (11), and C' only involves $z_2, \dots, z_k$.

If the maximal minors of P vanish, then $\vec{a}$ must lie in the $\mathbb{K}$-linear row span of C.

Thus, we may now assume $1 \le s \le r-2$. We have two remaining cases: $k = 1$ and $k \ge 2$. Let us first assume $k \ge 2$.

By subtracting columns of B from A, we may assume that, in the first $r-s-1$ rows of A, no z_0 appears. Also, let us add a copy of the $(r-s)$ row to each row $r-s+1, \dots, m-s$. Then we may assume that each row $r-s+1, \dots, m-s$ in the matrix B has no z_0 or z_1 appearing except for a single z_1 in the last column. We now claim that each row of A is in the $\mathbb{K}$-linear span of the rows of C.

First, let Q' be any $r \times n$ submatrix of Q obtained by choosing the first $r-s-1$ rows, one middle row, and the last s rows of Q. Consider any maximal submatrix $\tilde{Q}$ of Q' that involves the last $r-s-1$ columns of Q'. Write

$$\tilde{Q} = \begin{pmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \mathbf{0} \end{pmatrix}, \tag{12}$$

where $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are submatrices of A, B, and C, respectively. For $i = 1, \dots, r-s$, let b_i be the determinant of the matrix gotten from $\tilde{B}$ by deleting its i-th row, and let $\tilde{c}_i$ be the determinant of the matrix gotten by stacking the i-th row of $\tilde{A}$ to $\tilde{C}$. Let us show that each $\tilde{c}_i$ equals 0.

Since Q had rank at most $r-1$ by assumption, we have $\det(\tilde{Q}) = 0$. Note that

$$0 = \det(\tilde{Q}) = b_1\tilde{c}_1 - b_2\tilde{c}_2 + \cdots \pm b_{r-s}\tilde{c}_{r-s}.$$

Now, each b_i is of the form $z_0^{i-1} z_1^{r-s-i}$ plus terms that have lower total degree in z_0 and z_1. Also, each $\tilde{c}_i$, except possibly $\tilde{c}_{r-s}$, has no z_0 term appearing, by assumption. Then by inspecting the degree in z_0, we conclude that $\tilde{c}_{r-s} = 0$.

Next, we claim that no other $\tilde{c}_i$ has z_1 appearing: if z_1 appears with highest degree d in $\tilde{c}_i$, then the factor $z_0^{i-1} z_1^{r-s-i+d}$ would appear in the summand $b_i\tilde{c}_i$ but nowhere else. So z_0 and z_1 do not appear in any $\tilde{c}_i$. Finally, by inspecting the (z_0, z_1)-degree of each summand, we conclude that each $\tilde{c}_i$ equals 0.

By ranging over all possible choices of Q' and $\tilde{Q}$, we have that

$$\operatorname{rank} \begin{pmatrix} A \\ C \end{pmatrix} = \operatorname{rank} C.$$

Then we may apply Lemma 3.6 to conclude that the rows of A' are actually $\mathbb{K}$-linear combinations of the rows of C.

Finally, assume $1 \le s \le r-2$ and $k = 1$. Again, we want to show that any matrix A satisfying that the $r \times r$ anchored minors of (9) vanish must be a sum of a matrix in the $\mathbb{K}$-column span of $B = B(z_0, z_1)$ and one in the $\mathbb{K}$-row span of $C = C(z_0, z_1)$.

First, note that rows $r-s+1,\dots,m-s$ of B are zero. So we will start by showing that, for any i with $r-s+1 \le i \le m-s$, the i-th row of A is in the $\mathbb{K}$-row span of C. Indeed, let Q' be the $r \times n$ submatrix of Q obtained by choosing the first $r-s-1$ rows, the i-th row, and the last s rows of Q. So the lower right $(s+1)\times(r-s-1)$ block of Q' is zero. Now the first $r-s-1$ rows of Q' are $\mathbb{K}(z_0, z_1)$-linearly independent and their span intersects the span of the last $s+1$ rows trivially. Thus, the last $s+1$ rows have rank at most s. Applying Lemma 3.6 to these last $s+1$ rows, we are done. Furthermore, the same argument shows that columns $s+2,\dots,s+1+n-r$ of A are $\mathbb{K}$-combinations of the columns of B.

After deleting rows $r-s+1,\dots,m-s$ and columns $s+2,\dots,s+1+n-r$, we are left with a singular $r\times r$ matrix Q'. Let A' denote the upper left $(r-s)\times(s+1)$-submatrix of Q'. After adding multiples of the columns of B and the rows of C, we may assume that z_0 appears only in the lower right entry of A'. But z_0 cannot appear there either, for otherwise the monomial z_0^r would appear in $\det(Q')$ with nonzero coefficient, contradicting that $\det(Q') = 0$. So A' has only z_1.

Next, we claim we may reduce to the case that z_1 only appears in the first row and the last column of A'. This is because, if $a_{i,j} = \lambda z_1$, say, is a nonzero entry elsewhere in A', then we can alternately add λ-multiples of rows of C and subtract λ-multiples of columns of B to move λz_1 to any other entry in the same antidiagonal.

But in this case, z_1 doesn't appear at all, since each one would contribute a unique nonzero monomial in z_0^i, z_1^{r-i} to $\det(Q')$, but Q' is singular. Thus, after adding rows of C and columns of B, we have arrived at an equality $A' = 0$, so we are done. This finishes the proof of Theorem 3.1(i). $\square$

Proof of Theorem 3.2(i). Just as in the determinantal case, we need to exhibit, for any k and s as in Theorem 3.2, a matrix of the form (8) such that $a_{\text{perm}} = 0$ in Lemma 3.4. Then Theorem 3.2(i) would follow by upper semicontinuity: we need only invoke the natural $(S_m \times S_n)$-action on $\boldsymbol{F}_k(P^r_{m,n})$ gotten by permuting rows and columns that sends any standard compression component $\mathfrak{C}_k(\sigma,\tau)$ to the particular one shown in (3).

First, for distinct variables z_i, z_j, and z_k, define matrices $B(z_i, z_j, z_k)$ and $C(z_i, z_j, z_k)$, respectively, as

$$\left(\begin{array}{cccc} z_i & & & z_k \\ z_j & \ddots & & \\ z_k & \ddots & z_i & \\ & \ddots & z_j & z_i \\ & & z_k & z_j \\ \hline & & 0 & \end{array}\right) \quad \text{and} \quad \left(\begin{array}{cccccc|c} z_i & z_j & z_k & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & & z_i & z_j & z_k & 0 \\ z_k & & & & z_i & z_j & \end{array}\right), \tag{13}$$

of dimensions $(m-s)\times(r-s-1)$ and $s\times(s+1+n-r)$, respectively.

Let us clarify the degenerate cases of this definition:

$$s=1 \implies C(z_i,z_j,z_k)=(\,z_i\;\;z_j\;\;z_k\;\;0\;\cdots\;\;0\,);$$
$$s=r-2 \implies B(z_i,z_j,z_k)=(\,z_i\;\;z_j\;\;z_k\;\;0\;\cdots\;\;0\,)^T.$$

If $s=2$ then

$$C(z_i,z_j,z_k)=\begin{pmatrix} z_i & z_j & z_k & 0 & \cdots & 0\\ z_k & z_i & z_j & 0 & \cdots & 0\end{pmatrix};$$

and similarly for $B(z_i,z_j,z_k)$ when $r-s-1=2$.

In particular, in each case $C(z_i,z_j,z_k)$ has to have at least 3 columns whenever it is defined; in other words, if $s\neq 0$, we will always assume that $s+1+n-r\geq 3$. Similarly, $B(z_i,z_j,z_k)$ has to have at least 3 rows whenever it is defined; in other words, if $s\neq r-1$, we will always assume that $m-s\geq 3$. This accounts for the conditions in the statement of Theorem 3.2. We treat each case in turn.

First, suppose $s=0$, so $m\geq 3$ and the matrices C and D don't appear. Then let $B=B(z_0,z_1,z_2)+B'$, where B' is any matrix whose entries are linear forms not involving z_0, z_1, or z_2. Similarly, if $s=r-1$—so $n\geq 3$ and B and D don't appear—then we let $C=C(z_0,z_1,z_2)+C'$, where C' is any matrix whose entries are linear forms not involving z_0, z_1, or z_2. In both of these cases, Theorem 3.2(i) is an immediate consequence of the following lemma, whose proof we postpone to the end of the section.

Lemma 3.7. *Let $2\leq m\leq n$ but not $m=n=2$. Consider an $m\times n$ matrix*

$$P=\left(\begin{array}{c}\vec{a}\\ \hline \\ C \\ \\ \end{array}\right)$$

whose entries are linear forms $z_0,\dots,z_k$, where $\vec{a}=(a_i)$ is a $1\times n$ matrix and $C=C(z_0,z_1,z_2)+C'$ is an $(m-1)\times n$ matrix, where $C(z_0,z_1,z_2)$ is as shown in (13) and C' does not involve z_0, z_1, or z_2.

If the maximal permanents of P vanish, then $\vec{a}=0$.

Thus, we are done with the cases $s=0$ and $s=r-1$ modulo the lemma.

So we may assume for the rest of the proof that $1\leq s\leq r-2$. We let $B=B(z_0,z_1,z_2)+B'$ and $C=C(z_3,z_4,z_5)+C'$, for any matrices B' and C' that do not involve $z_0,\dots,z_5$, and furthermore do not involve any variables in common. Let D be any $s\times(r-s-1)$ matrix of linear forms. Again, we want to show $A=0$ if the $r\times r$ anchored permanents of Q in (9) are zero.

Here is the matrix Q:

$$
\begin{array}{cc}
 & \begin{array}{ccc} \qquad s+1 \qquad & \quad n-r \quad & r-s-1 \end{array} \\
\begin{array}{c} \\ r-s \\ \\ \\ m-r \\ \\ s \\ \\ \end{array} &
\left(\begin{array}{cccc|c|ccc}
 & & & & & z_0 & & z_2 \\
 & & & & & z_1 & z_0 & \\
 & & & & & z_2 & z_1 & z_0 \\
 & & & & & & z_2 & z_1 \\ \hline
 & & & & & & & \\ \hline
z_3 & z_4 & z_5 & & & & & \\
 & z_3 & z_4 & z_5 & & & 0 & \\
z_5 & & z_3 & z_4 & & & &
\end{array}\right)
\end{array} \tag{14}
$$

In this picture, A has been broken into four submatrices. If B or C have only one or two columns or rows, respectively, then they should be interpreted as having the degenerate forms described above. (We have also suppressed B' and C' for clarity.)

We break into the following two cases:

- $s = 1$ or $s = r - 1$, and
- $1 < s < r - 2$.

First, suppose $s = 1$. So C is a row matrix. For this case, we need the following lemma, which we will prove at the end of the section.

Lemma 3.8. *Let $r \geq 3$, and suppose Q' is an $m \times r$ matrix with entries in $\mathbb{K}[z_0, \ldots, z_k]_1$ of the form*

$$
Q' = \left(\begin{array}{c|c} A' & B \\ \hline \ell_1 \; \ell_2 & 0 \end{array}\right).
$$

Here ℓ_1 and ℓ_2 are nonzero linear forms in $z_3, \ldots, z_k$ that are not scalar multiples of each other, and $B = B(z_0, z_1, z_2) + B'$, where $B(z_0, z_1, z_2)$ is as in (13) and B' does not contain z_0, z_1, or z_2.

If the $r \times r$ permanents of Q' vanish then, for each $i = 1, \ldots, m-1$, the i-th row of A' is of the form $(a_i\ell_1 \; -a_i\ell_2)$ for some $a_i \in \mathbb{K}$.

Assuming the lemma holds, let us prove that $A = 0$. By applying Lemma 3.8 to each $m \times r$ submatrix of Q gotten by choosing any two of the first three columns and the last $r - 2$ columns, we see that each pair of the first three columns has entries that satisfy the sign relation as in Lemma 3.8. But that is impossible unless each entry in the first three columns of A is zero.

Let $Q_{m,1}$ denote the matrix gotten by deleting the m-th row and first column of Q. Since the first column, say, of Q is zero except for the last entry, which is nonzero, and Q has vanishing $r \times r$ anchored permanents, if follows that the $(r-1)\times(r-1)$ anchored permanents of the matrix $Q_{m,1}$ vanish. Applying Lemma 3.7 to (the transpose of) each $(m-1)\times(r-1)$ submatrix of $Q_{m,1}$ that involves the last $r-2$ columns shows that each column of A is zero.

The case $s = r-2$ is exactly analogous. In this case, the roles of B and C are simply reversed: now B is a column matrix with at least three nonzero entries. Then we apply Lemma 3.8 again to conclude that the first three rows of A must be zero, and then apply Lemma 3.7 to each $(m-1)\times(r-1)$ submatrix of $Q_{1,n}$ to show that every row of A is zero.

For the final case, let $1 < s < r-2$. In this case, our assumptions imply that B (respectively C) has at least two columns (respectively rows); see (14). We will call the $m-r$ rows and $n-r$ columns in (14), if they are present, the *middle* rows and columns. In the diagram, A has been broken into four blocks, and we will show one by one that they are 0.

First suppose $r = m = n$, so the middle rows and columns don't appear. As in the determinantal case, for $i = 1, \dots, m-s$ let b_i be the permanent of the matrix gotten by deleting the i-th row of B, and let $\tilde{c}_i$ be the permanent of the matrix gotten by stacking the i-th row of A to C. Then

$$0 = \operatorname{perm}(Q) = b_1\tilde{c}_1 + b_2\tilde{c}_2 + \cdots + b_{m-s}\tilde{c}_{m-s}.$$

Considering the terms of this equation of the form

$$z_0^{i-1} z_1^{r-s-i} \cdot \gamma,$$

where γ is any monomial not involving the variables appearing in B, shows that, after zeroing out the variables in A that appear in B, $\tilde{c}_i = 0$. Then, by Lemma 3.7, A only contains variables from B.

But a symmetric argument shows that A only contains variables from C. Since no variables appear in both B and C by assumption, we must have $A = 0$.

Now let us drop the assumption that $r = m = n$. Let Q' be the $r \times r$ submatrix of Q gotten by removing the middle $m-r$ rows and $n-r$ columns. Then the $r = m = n$ argument applied to Q' immediately shows that the upper left $(r-s)\times(s+1)$ block of A is zero. Next, we claim that every entry in the first $r-s$ rows of A is zero. Indeed, if in Q' we replace the $(s+1)$-st column with any middle column, we get an $r \times r$ matrix whose upper left $(r-s)\times s$ block is zero and whose lower left $s \times s$ block has nonzero permanent. Then the permanent of the upper right $(r-s)\times(r-s)$ block is zero so, by Lemma 3.7 again, the chosen middle column of A is zero, proving the claim. A symmetric argument shows that every entry in the first $s+1$ columns of A is zero.

Finally, for any middle row i and middle column j, we want to show that $a_{ij} = 0$. But this is immediate from considering the matrix obtained from Q' by replacing the $r-s$ row by i and the $s+1$ column by j, and noting that the permanent of this matrix is zero yet is a nonzero multiple of a_{ij}. This finishes the proof of Theorem 3.2(i). □

Proof of Theorems 3.1(ii) and 3.2(ii). We will show that, if k is sufficiently large, the proofs of Theorems 3.1(i) and 3.2(i) can be applied to any torus fixed point of $\mathfrak{C}_k(s)$ or $\mathfrak{C}_k(\sigma,\tau)$. When we say k is sufficiently large, we mean that conditions (4) and (5) are fulfilled in the determinantal case, or that conditions (6) and (7) are fulfilled in the permanental case.

Part (ii) of each theorem would then follow, because if the conditions (4) and (5) or the conditions (6) and (7) hold, then the bounds in Theorem 3.1(i) and Theorem 3.2(i) would apply to any torus fixed point of $\mathfrak{C}_k(s)$, respectively $\mathfrak{C}_k(\sigma,\tau)$. But the Zariski closure of the torus orbit of any point of $\mathfrak{C}_k(s)$ or $\mathfrak{C}_k(\sigma,\tau)$ certainly contains a torus fixed point. So the bounds would hold for arbitrary points of $\mathfrak{C}_k(s)$, respectively $\mathfrak{C}_k(\sigma,\tau)$.

By the discussion in Section 2, we know that any torus fixed point P of $\mathfrak{C}_k(s)$ or $\mathfrak{C}_k(\sigma,\tau)$ comes from zeroing out entries in the matrix of forms representing a standard compression space. Thus, after permuting rows and columns, we may assume that P has the form (8), and the entries of B, C, and D are either zeroes or $*$s representing distinct variables z_i. Furthermore, the number of zero entries among B, C, and D is exactly $\kappa(s)-k$.

The proofs in both the determinant and permanent cases reduce to the following combinatorics:

Lemma 3.9. *Let $c \le p$ and $p \ge q$, and suppose we have a $p \times q$ array $B=(b_{ij})$ filled with at most $p-c$ zeroes and the rest $*$s. Then we may permute rows and columns so that every entry $b_{i,j}$ with $i-j<c$ is a $*$.*

Assuming the lemma, let us prove Theorem 3.1(ii). Suppose $s \neq r-1$, so B is nonempty. Then $\kappa(s)-k < m-s-1$ by assumption, so B is an $(m-s)\times(r-s-1)$ matrix with at most $m-s-2$ zeroes. Then, by Lemma 3.9 applied with $c=2$, after permuting rows and columns of $B=(b_{i,j})$, the entries $b_{i,i}$ and $b_{i+1,i}$ are nonzero for all i. Then a linear change of coordinates brings B into the form required for Theorem 3.1(i). (The lemma shows that all entries above the diagonal $b_{i,i}$ are nonzero as well, but we don't need this.)

If instead $s \neq 0$, so that the matrix C is nonempty, and $k > \kappa(s)-(s+n-r)$, the same argument shows that there is a permutation of the rows and columns of C so that each of the entries $c_{i,i}$, $c_{i,i+1}$ are nonzero for $i=1,\dots,s$. Again, a linear change of coordinates then brings C into the form required for Theorem 3.1.

The proof of Theorem 3.2(ii) is almost identical, except that we apply Lemma 3.9 with $c=3$ instead of $c=2$, since the matrices in our special permanental form were tridiagonal instead of bidiagonal. Namely: if $s \neq r-1$, so that the matrix B

appears, the assumption that $\kappa(s)-k<m-s-2$ says that B has at most $m-s-3$ zeroes. Then Lemma 3.9 shows that, after permuting, every entry $b_{i,j}$ with $i-j<3$ is nonzero. Similarly, if $s\neq 0$, so that the matrix C appears, the assumption that $\kappa(s)-k<n-r+s-1$ and Lemma 3.9 again imply that, after permuting, every entry $c_{i,j}$ with $j-i<3$ is nonzero. Then we are again done after a linear change of coordinates. So we have proved part (ii) of each theorem, and we are done, apart from the proofs of the supporting lemmas below. □

Proof of Lemma 3.6. Let us write $a_i=\sum\lambda_{i,j}z_j$ for $\lambda\in\mathbb{K}$. Consider the determinant of the submatrix of P consisting of its first m columns. Inspection of the coefficient of the monomial $z_0^iz_1^{m-i}$ for $1\leq i\leq m-1$ gives the equation

$$\lambda_{i,0}=\lambda_{i+1,1}.$$

Likewise, inspection of the monomials z_0^m and z_1^m give equations $\lambda_{1,1}=\lambda_{m,0}=0$. Hence, after applying $\mathbb{K}$-linear row operations to the first row of P, we may assume that $\lambda_{i,j}=0$ for $i\leq m$ and $j=0,1$.

We now show that, under this assumption, $\vec{a}=0$. Indeed, consider again the determinant of the submatrix of P consisting of its first m rows. Inspection of the coefficient of the monomial $z_0^{i-1}z_1^{m-i}z_j$ shows that $\lambda_{i,j}=0$ for $i\leq m$ and $j\geq 2$. So we have proved $a_1=\cdots=a_m=0$. Finally, for any $i>m$, consider the determinant of the submatrix of P consisting of columns $2,\ldots,m,i$. Expanding by the first row, this determinant is $\pm a_i\cdot(z_1^{m-1}+$ terms of lower degree in z_1). The product is zero and the second factor is nonzero, so $a_i=0$. □

Before proving Lemma 3.7, we need the following calculation of the coefficients of the maximal permanents of a tridiagonal matrix:

Lemma 3.10. *Let $p\geq 2$ and consider the $p\times(p+1)$ matrix*

$$Q=\begin{pmatrix} u & v & w & & \\ & \ddots & \ddots & \ddots & \\ & & u & v & w \\ w & & & u & v \end{pmatrix}$$

for independent linear forms u, v, w, with 0 at every other position. Denote the permanent of the submatrix obtained by deletion of the i-th column by q_i. For $1\leq j\leq p-1$, the coefficient of $u^jv^{p-(j+1)}w$ in q_i is equal to

$$\operatorname{coeff}_{u^jv^{p-(j+1)}w}(q_i)=\begin{cases}\overline{p-j} & i=j,\\ 0 & i\neq j.\end{cases}$$

Here, $\overline{p-j}$ just means the integer $p-j$ reduced modulo the characteristic of $\mathbb{K}$.

Proof. Fix i, and let y be any monomial appearing in q_i of the form $u^jv^{p-(j+1)}w$ for some $1\leq j\leq p-1$. We claim that y must select u from each column to the left of the i-th column. If $i = 1$, there is nothing to prove. For $i > 1$, first note that y selects u from the first column. Indeed, if y selects w in the first column, then y must be at least quadratic in w if $i \neq p+1$, or would have the form $v^{p-1}w$ if $i = p+1$. Since y selects u from the first column, it cannot select anything else from the first row of P, forcing it to select u from column 2 if $i > 2$. Continuing in this fashion, we see that y must select u in all columns to the left of the i-th column.

Now, inspection of Q shows that, for each u which y selects to the right of the i-th column, it must select a w lying diagonally to the upper right. Likewise, for y to select a w entry, it must also select a u lying diagonally to the lower left and to the right of the i-th column. Hence, the only monomial of the desired form which appears in q_i is $u^i v^{p-(i+1)} w$, and it may be created by $p - j$ different choices of the position of w. □

Proof of Lemma 3.7. Write $a_i = \sum \lambda_{i,j} z_j$ for $\lambda \in \mathbb{K}$. First suppose $m = 2$, so $n > 2$ by assumption. In this case, we appeal to the explicit primary decomposition of the 2×2 permanents of a $2 \times n$ matrix over any field proved by Laubenbacher and Swanson [2000]. Each component is isolated in this case, and the components are as follows: there are two components that correspond to zeroing out each row; and there are $\binom{n}{2}$ components that correspond to zeroing out a given 2×2 permanent as well as all of the other $2(n-2)$ entries. It immediately follows from this classification that, if at least three entries of the $1 \times n$ matrix C are nonzero, then $\vec{a}$ must be 0.

Now suppose $m > 2$. We begin by considering the permanent of the submatrix of P consisting of its first m columns. Inspection of monomials of the form $z_0^i z_1^{m-i}$ and $z_1^i z_2^{m-i}$ gives equations

$$\begin{aligned} \lambda_{i,0} + \lambda_{i+1,1} &= 0 \quad \text{for } 1 \le i \le m-1, \\ \lambda_{i,1} + \lambda_{i+1,2} &= 0 \quad \text{for } 2 \le i \le m-1, \\ \lambda_{m,1} + \lambda_{1,2} &= 0, \end{aligned}$$

along with $\lambda_{2,2} = \lambda_{1,1} = \lambda_{m,0} = 0$. On the other hand, inspection of the monomials of the form $z_0^i z_1^{m-1-i} z_2$ gives equations

$$\overline{(m-1-i)}\,\lambda_{i,0} + \overline{(m-2-i)}\,\lambda_{i+1,1} + \lambda_{i+2,2} = 0 \quad \text{for } 1 \le i \le m-2$$

by Lemma 3.10. Likewise, inspection of the monomial $z_0 z_1^{m-3} z_2^2$ gives the equation $\overline{m-2}\lambda_{1,2} + \overline{m-3}\lambda_{m,1} + \lambda_{m-1,0}$. Since $\operatorname{char}(\mathbb{K}) \neq 2$, this system of equations implies that $\lambda_{i,j} = 0$ for $1 \le i \le m$ and $j = 1, 2, 3$. For $k = m+1, \dots, n$, consider the first $m-1$ columns and the k-th column of P. The permanent of this $m \times m$ matrix is zero, and all of the entries in the first row are zero except possibly a_k. Expanding the permanent by the first row, we have $a_k = 0$. □

Proof of Lemma 3.8. First, let's prove that, for $i = 1, 2$, each form in column i is a scalar multiple of ℓ_i. Let Q'_{m1} be the matrix obtained from Q' by deleting the last row and first column. Let z_j be any form in the support of ℓ_2, and consider the image in $\mathbb{K}[z_0, \ldots, z_k]/\ell_2 \cong S = \mathbb{K}[z_0, \ldots, \hat{z}_j, \ldots, z_k]$ of each $r \times r$ subpermanent of Q' that involves the last row of Q'. By expanding along the last row, this image is $\overline{\ell_1} \cdot \mathrm{perm}(P)$, where P is an $(r-1) \times (r-1)$ submatrix of Q'_{m1}.

Since $\overline{\ell_1} \neq 0$ by assumption, $\mathrm{perm}(P) = 0$ as an element of S. Then, by Lemma 3.7 applied over the ring S to Q'_{m1}, the first column of Q'_{m1} is zero. In other words, every entry in the second column of Q' is a scalar multiple of ℓ_2. Similarly, every entry in the first column of Q' is a scalar multiple of ℓ_1.

Let $\tilde{Q}$ be the $r \times r$ submatrix consisting of the first $r-1$ rows along with the last row of $\tilde{Q}$. Factoring out ℓ_1 and ℓ_2 from the first two columns of $\tilde{Q}$, we see

$$0 = \mathrm{perm}(\tilde{Q}) = \ell_1 \ell_2 \cdot \beta,$$

where β is a $\mathbb{K}$-linear combination of the maximal subpermanents of B; and in fact these maximal subpermanents are $\mathbb{K}$-linearly independent, since, for example, they each contain a distinct term of highest total (z_0, z_1)-degree. So the coefficients are zero. In other words, for $i = 1, \ldots, r-1$, the 2×2 permanent consisting of the i-th row of A' along with $(\ell_1 \ \ell_2)$ is zero.

Similarly, for $i = r, \ldots, m-1$, let $\tilde{Q}$ be the $r \times r$ submatrix consisting of the first $r-2$ rows, the i-th row, and the last row of $\tilde{Q}$. Then, since the uppermost $(r-2) \times (r-2)$ subpermanent of B is the unique maximal subpermanent of B with a nonzero monomial z_0^{r-2}, but $\mathrm{perm}(\tilde{Q}) = 0$, it follows again that the 2×2 permanent consisting of the i-th row of A' along with $(\ell_1 \ \ell_2)$ is zero. We conclude that each row of A' is of the form $(a\ell_1 \ -a\ell_2)$ for some $a \in \mathbb{K}$. □

Proof of Lemma 3.9. The proof is by induction on $p+q$ with trivial base case. First, if $c > p-q$ then, by the pigeonhole principle, there is a column of all $*$s. Moving that column to be the rightmost one and inducting on the remaining $p \times (q-1)$ matrix, we are done. So we may assume $c \leq p-q$. Let z be the smallest number of zeroes in any column of B. We claim that $z \leq p-c-q+1$. If $z \leq 1$ this is clear, because $p-c-q+1 \geq 1$. Otherwise, since there are at least qz zeroes in B, we have $p-c \geq qz \geq q+z-1$, implying the claim.

Then we can permute rows and columns so that the last column of B has $*$s for its first $p-z$ entries and zeroes for its last z entries. The claim we just proved above says precisely that the last column satisfies the hypotheses of the lemma. Furthermore, the upper left $(p-z) \times (q-1)$ submatrix of B has at most $p-z-c$ zeroes, so we are done by induction. □

4. Results on compression spaces

In this section, we will prove a number of results on compression spaces that will be used in Section 5 to prove results on smoothness and connectedness of our Fano schemes. We will start by proving that the subschemes $\mathfrak{C}_k(s)$ of $\boldsymbol{F}_k(D^r_{m,n})$ are actually irreducible components. This result relies on the tangent space dimension at a general point of $\mathfrak{C}_k(s)$, computed in Theorem 3.1. Similarly, we will prove that, when the conditions on k and s in Theorem 3.2 hold, $\mathfrak{C}_k(\sigma, \tau)$ is a component, where $|\tau| = s$. Finally, we will prove a theorem describing precisely the embedding in projective space of the components $\mathfrak{C}(s)$ of $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$, including a computation of their degrees.

As discussed in Section 2, for each $k \le \kappa(s)$ there is a natural map

$$\rho_k(s) : \mathrm{Gr}(k+1, \mathcal{U}(s)) \to \boldsymbol{F}_k(D^r_{m,n}),$$

where $\mathcal{U}(s)$ is the pullback to $\mathfrak{C}(s)$ of the universal bundle on $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$; the image of this map is $\mathfrak{C}_k(s)$. The next theorem says that $\rho_k(s)$ is generically finite. Since the dimension of the space $\mathrm{Gr}(k+1, \mathcal{U}(s))$ is exactly the general tangent space dimension calculated in Theorem 3.1, it will follow that $\mathfrak{C}_k(s)$ is a component (Corollary 4.3).

Theorem 4.1. *The map $\rho_k(s)$ is generically finite. Furthermore, it is a closed embedding if and only if*

$$\begin{aligned} & k > \kappa(s) - (m-s) && \text{if } s \neq r-1, \\ \text{and} \quad & k > \kappa(s) - (n-r+s+1) && \text{if } s \neq 0. \end{aligned}$$

Furthermore, the image of $\rho_k(s)$ is smooth if and only if the above bounds on k hold.

Before proving the theorem, we will need the following lemma:

Lemma 4.2. *Consider compression spaces P_1 and $P_2 \in \mathfrak{C}(s)$ with $P_1 \neq P_2$. Write $P_1 \cap P_2$ for their intersection as $\kappa(s)$-planes in $\mathbb{P}^{mn-1}$. Then*

$$\dim(P_1 \cap P_2) \le \kappa(s) - \min\{m-s,\ s+1+n-r\}.$$

Furthermore, if $s = 0$ or $s = r-1$ then we can say which term in the minimum to take: we have

$$\begin{aligned} & \dim(P_1 \cap P_2) \le \kappa(s) - (m-s) && \text{if } s = 0, \\ \text{and} \quad & \dim(P_1 \cap P_2) \le \kappa(s) - (s+1+n-r) && \text{if } s = r-1. \end{aligned}$$

Proof. Write $a = s+1+n-r$ for convenience. For $i = 1, 2$, P_i can be characterized as the space of matrices that sends an a-dimensional subspace $V_i \subseteq \mathbb{K}^n$ into an s-dimensional subspace $W_i \subseteq \mathbb{K}^m$. Then $P_1 \cap P_2$ is the space of matrices that:

(i) send V_1 to W_1;

(ii) send V_2 to W_2.

As a consequence of (i) and (ii), we also have that elements of $P_1 \cap P_2$:

(iii) send $V_1 \cap V_2$ to $W_1 \cap W_2$.

We will show that $\dim(P_1 \cap P_2)$ is maximized when $V_1 \cap V_2$ and $W_1 \cap W_2$ are as large as possible.

Let $b = \dim V_1 \cap V_2$. Then (iii) gives $b \cdot \operatorname{codim}(W_1 \cap W_2)$ linear conditions on the matrices in P_i, and (i) and (ii) each give $(a-b)(m-s)$ additional linear conditions. So, because of (iii), if we fix V_1 and V_2, then choosing W_1 and W_2 to have as large intersection as possible gives $\dim(P_1 \cap P_2)$ as large as possible. Similarly, suppose we fix W_1 and W_2 and increase the dimension of $V_1 \cap V_2$ by one. This adds only $\operatorname{codim}(W_1 \cap W_2) \leq 2(m-s)$ conditions in (iii), but removes $2(m-s)$ conditions arising from (i) and (ii). Thus, $\dim(P_1 \cap P_2)$ attains its maximum when $V_1 \cap V_2$ and $W_1 \cap W_2$ are as large as possible.

The only constraint, of course, is that $P_1 \neq P_2$, i.e., we don't have both $V_1 = V_2$ and $W_1 = W_2$. Thus, $\dim(P_1 \cap P_2)$ is largest if either $V_1 = V_2$ and $\dim(W_1 \cap W_2) = s-1$, or $\dim(V_1 \cap V_2) = a-1$ and $W_1 = W_2$. In the first case, by counting linear constraints on the points of $\mathbb{P}^{mn-1}$, we have

$$\dim(P_1 \cap P_2) = (mn-1) - 2(m-s) - (a-1)(m-s) = \kappa(s) - (m-s),$$

and in the second case we have

$$\dim(P_1 \cap P_2) = (mn-1) - a(m-s+1) = \kappa(s) - (s+1+n-r).$$

We conclude that, for any $P_1 \neq P_2$, $\dim(P_1 \cap P_2) \leq \kappa(s) - \min\{m-s,\ s+1+n-r\}$.

Furthermore, if $s = 0$ then $\dim(W_1 \cap W_2) = s-1$ is not possible, and similarly if $s = r-1$ then $\dim(V_1 \cap V_2) = a-1 = n-1$ is not possible (since if $s = r-1$ then V_1 and V_2 are both n-dimensional, so necessarily $\dim(V_1 \cap V_2) = n$ as well). This accounts for the stronger bounds in the second part of the lemma in these special cases. □

Proof of Theorem 4.1. We first show that, if the lower bounds on k hold, the map $\rho_k(s)$ is injective. Indeed, consider any two distinct points $x, y \in \operatorname{Gr}(k+1, \mathcal{U}(s))$ whose image under $\rho_k(s)$ is the same. For any value of k, it is clear that these points must project to distinct points $P, Q \in \mathfrak{C}(s)$. Furthermore, the fact that the images of x and y are equal implies that P and Q contain a common $(k+1)$-dimensional subspace. If we assume that the lower bounds on k hold, then this is impossible by Lemma 4.2. Hence, if the lower bounds on k hold, $\rho_k(s)$ is injective.

We now consider the differential $d\rho_k(s)$ of $\rho_k(s)$. We will show that this is injective everywhere if the bounds on k hold. Then, since $\rho_k(s)$ is injective with injective differential, it is a closed embedding. In particular, if the bounds on k hold, $\mathfrak{C}_k(s)$ is smooth, since it is isomorphic to a Grassmann bundle over a product of Grassmannians. Furthermore, we will show that, for arbitrary k, $d\rho_k(s)$ is injective at certain points. Hence $\rho_k(s)$ must be generically finite.

To start, we describe the map $d\rho_k(s)$. Let Q be the standard compression space compressing the first $s+n-r+1$ standard basis vectors into the subspace generated by the last s standard basis vectors; see (3). In a neighborhood of the fiber over Q, $\mathrm{Gr}(k+1,\mathcal{U}(s))$ trivializes as $\mathbb{A}^{\delta(s)}\times\mathrm{Gr}(k+1,\kappa(s)+1)$, and the tangent space of any point q in the fiber decomposes as a direct sum of

$$T_Q\mathfrak{C}(s)\cong\mathbb{K}^{\delta(s)}\cong(\mathbb{K}^{s+1+n-r}\otimes\mathbb{K}^{r-s-1})\oplus(\mathbb{K}^{m-s}\otimes\mathbb{K}^{s})$$

(tangent directions on the base) and $\mathbb{K}^{(k+1)(\kappa(s)-k)}$ (tangent directions in the fiber). Recall the definition of $\delta(s)$ in (2).

Let $p\in\mathfrak{C}_k(s)$ be the image of $q\in\mathrm{Gr}(k+1,\mathcal{U}(s))$. We may represent p by a matrix P as in (8). The ideal I of this point in $D^r_{m,n}$ is thus given by $x_{i,j}$ for $i\le m-s$ and $j\le s+1+n-r$, along with $\kappa(s)-k$ additional independent linear forms f_i as in Theorem 3.1. The tangent space $T_p\mathfrak{C}_k(s)$ is contained in the tangent space $T_p\boldsymbol{F}_k(D^r_{m,n})$. As in the proof of Lemma 3.4, tangent vectors to $\boldsymbol{F}_k(D^r_{m,n})$ at p may be described by mapping the $x_{i,j}$ and f_i to linear forms modulo J. There are no restrictions on the choice of the images of f_i, but the images of the $x_{i,j}$ may be constrained; denote these images by $y_{i,j}$.

Now, the summand $\mathbb{K}^{(k+1)(\kappa(s)-k)}\subset T_q\,\mathrm{Gr}(k+1,\mathcal{U}(s))$ maps isomorphically to the subspace of $T_p\mathfrak{C}_k(s)$ with all $y_{i,j}=0$. Indeed, deforming q within the fiber over Q amounts to perturbing the f_i. On the other hand, a tangent vector of the form

$$(t\otimes u,v\otimes w)\in(\mathbb{K}^{s+1+n-r}\otimes\mathbb{K}^{r-s-1})\oplus(\mathbb{K}^{m-s}\otimes\mathbb{K}^{s})$$

maps to a tangent vector with the image $y_{i,j}$ of $x_{i,j}$ given by

$$y_{i,j}=(Bu)_it_j+v_i(wC)_j.$$

Here, we have chosen a basis for $T_Q\mathfrak{C}(s)$ exactly as in the proof of Theorem 3.1(i): we may deform Q within $\mathfrak{C}(s)$ by performing $(m-s)m$ independent row and $(s+1+n-r)(r-s-1)$ independent column operations. It follows from our description of the differential $d\rho_k(s)$ that $d\rho_k(s)$ is injective at q if the rows of B, as well as the columns of C, are linearly independent. This latter condition is clearly the case if either the requisite bounds on k hold or q is a generic point in the fiber over Q. Furthermore, if the bounds on k hold, then applying the $(\mathrm{GL}_m\times\mathrm{GL}_n)$-action on $\mathrm{Gr}(k+1,\mathcal{U}(s))$ induced by the natural action on $\mathfrak{C}(s)$ shows that the differential is injective everywhere. We conclude that $\rho_k(s)$ is generically finite, and is furthermore a closed embedding if the bounds on k hold.

Finally, if the bounds on k fail to hold, $\mathfrak{C}_k(s)$ must be singular. Indeed, we will exhibit points P in $\mathfrak{C}_k(s)$ with tangent space dimension equal to

$$\dim T_P\mathfrak{C}_k(s)=(k+1)(\kappa(s)-k)+(k+1)(m-s)(s+n-r+1).\tag{15}$$

This exceeds the dimension of $\mathrm{Gr}(k+1,\mathcal{U}(s))$, which equals the dimension of $\mathfrak{C}_k(s)$. Thus, $\mathfrak{C}_k(s)$ is singular at these points.

Suppose that $k \le \kappa(s) - (m-s)$ and $s \ne r-1$. Consider any point $p \in \mathfrak{C}_k(s)$ represented by a matrix P as in (8) with the first column of B identically 0. We may describe tangent vectors of $\mathfrak{C}_k(s)$ at p as above. There is a point $q \in \mathrm{Gr}(k+1, \mathcal{U}(s))$ mapping to p which lies in the fiber over the standard compression space in $\mathfrak{C}(s)$, which compresses the subspace generated by the standard basis vectors in positions $2, \dots, s+n-r+2$ into the subspace generated by the last s standard basis vectors. Inspecting the differential of $\rho_k(s)$ at q shows that its image contains all tangent vectors with $y_{i,j} = 0$ for $j \ne 1$, that is, the $y_{i,1}$ may be arbitrary. Considering preimages of p in other fibers over $\mathfrak{C}_k(s)$ shows that all $y_{i,j}$ may be arbitrary and, hence, the tangent space dimension at p is as in (15).

If instead $k \le \kappa(s) - (n-r+s+1)$ and $s \ne 0$, a similar argument involving zeroing out the first row of C completes the proof. □

Corollary 4.3. *For each $k \le \kappa(s)$, the variety $\mathfrak{C}_k(s)$ is an irreducible component of $\boldsymbol{F}_k(D^r_{m,n})$, of dimension $\delta(s) + (k+1)(\kappa(s)-k)$. If conditions (4) and (5) in Theorem 3.1 are satisfied, then every point of $\mathfrak{C}_k(s)$ is a smooth point of $\boldsymbol{F}_k(D^r_{m,n})$.*

Proof. Since the map $\rho_k(s)$ is generically finite by Theorem 4.1, the variety $\mathfrak{C}_k(s)$ has the same dimension as the Grassmann bundle $\mathrm{Gr}(k+1, \mathcal{U}(s))$ over $\mathfrak{C}(s)$. The dimension of $\mathfrak{C}(s)$ is $\delta(s)$ (see (2)) and the rank of $\mathcal{U}(s)$ is $\kappa(s)+1$; hence, the dimension of $\mathfrak{C}_k(s)$ is $\delta(s) + (k+1)(\kappa(s)-k)$. But, by Theorem 3.1(i), this is an upper bound on the tangent space dimension of $\boldsymbol{F}_k(D^r_{m,n})$ at a general point of $\mathfrak{C}_k(s)$. Hence, $\mathfrak{C}_k(s)$ must be an irreducible component of $\boldsymbol{F}_k(D^r_{m,n})$.

Furthermore, if the bounds (4) and (5) on k are satisfied, it follows from Theorem 3.1(ii) that the tangent space dimension of any point of $\boldsymbol{F}_k(D^r_{m,n})$ equals the dimension of $\mathfrak{C}_k(s)$. □

By the same token, for numbers k and s satisfying the conditions in Theorem 3.2, the subscheme $\mathfrak{C}_k(\sigma, \tau)$ is actually a component of $\boldsymbol{F}_k(P^r_{m,n})$.

Proposition 4.4. *Fix integers k and s with $0 \le s \le r-1$, $k \le \kappa(s)$. Let σ and τ be subsets of $\{1, \dots, n\}$ and $\{1, \dots, m\}$ of sizes $s+1+n-r$ and s, respectively. Suppose that*

$$s = 0,\ r-1 \quad \textit{and} \quad 2 \le k \le \kappa(s), \qquad \textit{or} \qquad 1 \le s \le r-2 \quad \textit{and} \quad 5 \le k \le \kappa(s),$$

and that

$$s+1+n-r \ge 3 \quad \textit{if } s \ne 0 \qquad \textit{and} \qquad m-s \ge 3 \quad \textit{if } s \ne r-1.$$

Then $\mathfrak{C}_k(\sigma, \tau)$ is an irreducible component of $\boldsymbol{F}_k(P^r_{m,n})$.

Proof. By definition (see the discussion in Section 2), $\mathfrak{C}_k(\sigma, \tau)$ is isomorphic to $\mathrm{Gr}(k+1, \kappa(s)+1)$ and thus has dimension $(k+1)(\kappa(s)-k)$. If the assumptions hold, we can apply the first part of Theorem 3.2 to conclude that the tangent space dimension of $\boldsymbol{F}_k(P^r_{m,n})$ at a general point of $\mathfrak{C}_k(\sigma,\tau)$ is at most $(k+1)(\kappa(s)-k)$. But this is the dimension of $\mathfrak{C}_k(\sigma,\tau)$, hence it must be an irreducible component of $\boldsymbol{F}_k(P^r_{m,n})$. □

The next result ensures that the compression components $\mathfrak{C}_k(s)$ and $\mathfrak{C}_k(\sigma,\tau)$, respectively, of $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ that we have now identified are distinct.

Proposition 4.5. *Let* $k \geq 1$.

(i) $\mathfrak{C}_k(s) = \mathfrak{C}_k(s')$ *if and only if* $s = s'$.

(ii) *For any* $\sigma, \sigma' \subset \{1, \ldots, n\}$ *and* $\tau, \tau' \subset \{1, \ldots, m\}$ *with*

$$|\sigma| - |\tau| = |\sigma'| - |\tau'| = n - r + 1,$$

we have $\mathfrak{C}_k(\sigma,\tau) = \mathfrak{C}_k(\sigma',\tau')$ *if and only if* $\sigma = \sigma'$ *and* $\tau = \tau'$.

Proof. Part (ii) is more or less immediate, since we can certainly find a k-plane in the standard compression space $\mathfrak{C}(\sigma,\tau)$ with points having nonzero entries in all but the $(|\tau|+1+n-m)(m-|\tau|)$ required entries. Such a k-plane clearly cannot lie inside some other standard compression space $\mathfrak{C}(\sigma',\tau')$, since the matrices in $\mathfrak{C}(\sigma',\tau')$ must have some new entries that are required to be zero.

Let us prove (i). For every pair of numbers s' and s with $0 \leq s' < s < r$, we will show that there is a line in an s-compression space that doesn't lie in any s'-compression space. That will show that $\mathfrak{C}_1(s')$ and $\mathfrak{C}_1(s)$ are distinct. Furthermore, it would follow that any k-plane inside that s-compression space that contains that line doesn't lie in any s'-compression space, thus showing that $\mathfrak{C}_k(s)$ and $\mathfrak{C}_k(s')$ are distinct for all $k > 1$ too.

Indeed, for any $1 \leq s \leq r-1$, consider the line in $\mathbb{P}^{mn-1}$ of matrices $P(z_0, z_1)$ of the form given in (8), setting $D = 0$ and setting $B = B(z_0, z_1)$ and $C = C(z_0, z_1)$. These matrices are defined in (11).

We now show that this point of $\mathfrak{C}_1(s)$ does not lie in $\mathfrak{C}_1(s')$ for any $s' < s$. In other words, we wish to show that, for all subspaces A of $\mathbb{K}^n$ of dimension $s'+1+n-r$, there is a choice of z_0 and z_1 such that A meets the kernel of the matrix $P(z_0, z_1)$ in dimension at most $n-r$.

Indeed, the kernel of $P(z_0, z_1)$ is the $(n-r+1)$-dimensional space spanned by the rows of the matrix

$$\left(\begin{array}{cccc|c} z_1^s & -z_0 z_1^{s-1} & \cdots & \pm z_0^s & \\ & & & I_{n-r} & \mathbf{0} \end{array}\right).$$

So, if A does not already contain the span of the last $n-r$ vectors, we are done. If it does then, since the rational normal curve of the first row sweeps out an $(s+1)$-dimensional space but A is only $(s'+1+n-r)$-dimensional, a general choice of z_0 and z_1 will have the property that A meets the kernel of $P(z_0, z_1)$ in dimension at most $n-r$, as desired. □

Proposition 4.6. *If $\boldsymbol{F}_k(P^r_{m,n})$ is nonempty, then it is reducible.*

Proof. We may assume that $n>3$, since the case of $n=2$ is just $\boldsymbol{F}_1(P^2_{2,2})\cong\boldsymbol{F}_1(D^2_{2,2})$, which has two components, each a copy of $\mathbb{P}^1$.

Suppose first that $k>1$. Then the reducibility follows from Propositions 4.4 and 4.5. Indeed,

$$\mathfrak{C}_k(\{1,\dots,n\},\{1,\dots,r-1\}) \quad\text{and}\quad \mathfrak{C}_k(\{1,\dots,n\},\{2,\dots,r\})$$

are distinct irreducible components.

Suppose next that $k=1$ and $r>2$. We will consider subschemes

$$Z=\mathfrak{C}_1(\{1,\dots,n\},\tau) \quad\text{and}\quad Z'=\mathfrak{C}_1(\{1,\dots,n\},\tau'),$$

where $|\tau|=|\tau'|=r-1$ and we take τ, τ' to have smallest possible nonempty intersection, that is, $|\tau\cap\tau'|=\max\{1,2(r-1)-m)\}$. Each of these schemes has dimension $2(\kappa(r-1)-1)$, and their intersection has codimension $2(r-2)n$ or $2(m-(r-1))n$ in Z, Z'. Applying Lemma 3.4 to a point in Z or Z' whose $r-1$ nonzero rows form the matrix $C(z_0,z_1)$ in (11) shows that the tangent space dimension of $\boldsymbol{F}_k(P^r_{m,n})$ at general points of Z and Z' is bounded above by $2(\kappa(r-1)-1)+\delta(r-1)$. Indeed, a straightforward calculation shows that the permissible matrices A in Lemma 3.4 are precisely those matrices whose $m-r+1$ rows are linear combinations of the $r-1$ vectors obtained from the rows of $C(z_0,z_1)$ by flipping the sign of z_1, yielding a total dimension of $\delta(r-1)$.

Now, suppose that $Y=\boldsymbol{F}_1(P^r_{m,n})$ is irreducible. Then the codimension of Z, Z' in $\boldsymbol{F}_k(P^r_{m,n})$ is bounded above by $\delta(r-1)=(r-1)(m-(r-1))$. Since we must have

$$\operatorname{codim}_Z Z\cap Z'\le\operatorname{codim}_Y Z',$$

we get

$$\begin{aligned} 2(r-2)n &\le (r-1)(m-(r-1)) \quad \text{if } |\tau\cap\tau'|=1,\\ 2(m-(r-1))n &\le (r-1)(m-(r-1)) \quad \text{if } |\tau\cap\tau'|=2(r-1)-m. \end{aligned}$$

But both of these are impossible if $r>2$, so $\boldsymbol{F}_k(P^r_{m,n})$ cannot be irreducible.

Finally, suppose that $k=1$ and $r=2$. Then $P^2_{m,n}$ is reducible by [Laubenbacher and Swanson 2000], so $\boldsymbol{F}_1(P^2_{m,n})$ must be as well. Indeed, it is immediate from the classification of minimal primary components of $P^2_{m,n}$ [Laubenbacher and Swanson 2000, Theorem 4.1] that every point of $P^2_{m,n}$ is contained in a line in $P^2_{m,n}$. So there is a surjective map from the universal bundle $\mathcal{U}$ over $\boldsymbol{F}_1(P^2_{m,n})$ to $P^2_{m,n}$. Hence, if $\boldsymbol{F}_1(P^2_{m,n})$ were irreducible, then $\mathcal{U}$ and $P^2_{m,n}$ would be too. $\square$

The last result in this section studies the degrees of the components $\mathfrak{C}(s)$ of $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$ for $s=0,\dots,r-1$.

Theorem 4.7. *The s-compression spaces form a subscheme $\mathfrak{C}(s)$ of the Fano scheme $\boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$ isomorphic to $\mathrm{Gr}(s+n-r+1, n) \times \mathrm{Gr}(m-s, m)$. The embedding is via the Segre product of the $(m-s)$-fold Veronese of $\mathrm{Gr}(s+n-r+1, n)$ in its Plücker embedding, and the $(s+n-r+1)$-fold Veronese of $\mathrm{Gr}(m-s, m)$ in its Plücker embedding, possibly up to a linear projection. Thus, $\mathfrak{C}(s)$ has degree*

$$\binom{\delta(s)}{s(m-s)} d_1 \cdot (m-s)^{(s+n-r+1)(r-s-1)} \cdot d_2 \cdot (s+n-r+1)^{s(m-s)},$$

where

$$d_1 = \frac{((s+n-r+1)(r-s-1))!}{\prod_{1\le i\le r-s-1<j\le n}(j-i)} \quad \textit{and} \quad d_2 = \frac{(s(m-s))!}{\prod_{1\le i\le s<j\le m}(j-i)}.$$

Proof. Let $t = s+n-r+1$. An s-compression space consists of the matrices that send the row span of some full-rank $t \times n$ matrix $A = (a_{i,j})$ to the orthogonal complement of the row span of a full-rank $(m-s) \times m$ matrix $B = (b_{i,j})$. We would like to express the Plücker coordinates of this point of $\mathfrak{C}(s)$ in terms of the Plücker coordinates of A and of B.

The condition that a linear map $M = (x_{i,j})$ sends A to the orthogonal complement of B is precisely the condition $BMA^T = 0$; that is,

$$(BMA^T)_{i,j} = \sum_{u=1}^{m}\sum_{v=1}^{n} b_{i,u}a_{j,v}x_{u,v} = 0$$

for $1 \le i \le m-s$ and $1 \le j \le t$. Thus, the compression space defined by A and B can be represented by the $t(m-s) \times mn$ matrix C whose rows are indexed by (i, j) and columns by (u, v), and whose $((i, j), (u, v))$ entry is $b_{i,u}a_{j,v}$.

Write $\{p_I\}_{|I|=t}$ and $\{q_J\}_{|J|=m-s}$ for the maximal minors of A and B, respectively. The maximal minors of C are polynomials of bidegree $(t(m-s), t(m-s))$ in the entries of A and in the entries of B and, by construction, they are invariant under the action of GL_t and GL_{m-s} on A and B, respectively. It follows that they are polynomials of bidegree $m-s$ and t in the Plücker coordinates of A and of B, respectively (whenever they are nonzero), since the Plücker coordinates generate the rings of GL-invariants.

We do not have a general formula for the maximal minors of C, but can describe the minors in special cases, and this will be enough to prove the theorem. Recall that a maximal minor of C corresponds to choosing exactly $t(m-s)$ entries $x_{u,v}$ of M. Call a choice of exactly $m-s$ entries in each of t columns of M *A-pure*; call a choice of exactly t entries in each of $m-s$ rows of M *B-pure*. In the following, we will give a formula expressing A-pure and B-pure minors as monomials in the p_I and q_J.

Suppose we have an B-pure choice of entries, which after permuting columns can be described as follows. Let $I_1, \dots, I_{m-s}$ be order-t subsets of $\{1, \dots, n\}$. Then the $t(m-s)$ entries of M

$$\{(u, v) : u = 1, \dots, m-s, v \in I_i\} \tag{16}$$

are B-pure. In other words, we choose t entries in each of the first $m-s$ rows of M as given by the sets $I_1, \dots, I_{m-s}$.

We claim in Lemma 4.8 below that the maximal minor of C that corresponds to the choice of sets $I_1, \dots, I_{m-s}$ in the way we just described is exactly

$$\pm p_{I_1} \cdots p_{I_{m-s}} q_{\{1,\dots,m-s\}}{}^t.$$

Of course, the argument applies, after permuting columns, to any B-pure maximal minor; and a symmetric argument, which we omit, applies to any A-pure maximal minor. Call a monomial in the p_I and q_J *pure* if it involves a unique p_I or if it involves a unique q_J. Then it would follow immediately from the claim that *every pure monomial appears as a maximal minor of C*.

Now, by the above discussion, we see that the map

$$X = \mathrm{Gr}(t, n) \times \mathrm{Gr}(m-s, m) \to \boldsymbol{F}_{\kappa(s)}(D^r_{m,n})$$

taking a pair of subspaces A, B to the compression space mapping A into the orthogonal complement of B is given by a linear system L which is a subspace of the complete linear system $|\mathcal{O}_X(m-s, t)|$. Since every pure monomial appears in L, it immediately follows that the map determined by L is an embedding.

The degree of $\mathrm{Gr}(a, b)$ in its Plücker embedding is known classically as

$$(a(b-a))!/\prod_{1\le i\le a<j\le b}(j-i);$$

see, e.g., [Mukai 1993]. Thus, d_1 and d_2 are just the degrees of $\mathrm{Gr}(s+n-r+1, n)$ and $\mathrm{Gr}(m-s, m)$ in their Plücker embeddings. It follows that the degree of X embedded by $|\mathcal{O}_X(m-s, t)|$ is given by the formula in the theorem, since the Hilbert polynomial of a Segre product is the product of the Hilbert polynomials. But since L is a subspace of $|\mathcal{O}_X(m-s, t)|$ that also gives an embedding, the degree of X embedded via L is the same. $\square$

Lemma 4.8. *The maximal minor of C corresponding to the sets $I_1, \dots, I_{m-s}$ as in (16) is*

$$\pm p_{I_1} \cdots p_{I_{m-s}} q_{\{1,\dots,m-s\}}{}^t.$$

We give a running illustration of the proof that follows in Example 4.9.

Proof of Lemma 4.8. The rows of the $t(m-s) \times t(m-s)$ submatrix C' of C in question are indexed by pairs (i, j) with $i = 1, \dots, m-s$ and $j = 1, \dots, t$.

Order them lexicographically. The columns of C' are indexed by entries (i, j) with $i = 1, \dots, t$, $j \in I_i$. Order these lexicographically as well; the determinant is of course preserved up to sign. Then C' can be regarded as a block matrix with $(m-s)^2$ blocks $G_{i,j}$ of size $t \times t$,

$$C' = \begin{pmatrix} G_{1,1} & \cdots & G_{1,m-s} \\ \vdots & \ddots & \vdots \\ G_{m-s,1} & \cdots & G_{m-s,m-s} \end{pmatrix},$$

where $G_{i,j}$ is the $t \times t$ matrix

$$G_{i,j} = \begin{pmatrix} b_{i,j}a_{1,I_{j,1}} & \cdots & b_{i,j}a_{1,I_{j,t}} \\ \vdots & \ddots & \vdots \\ b_{i,j}a_{t,I_{j,1}} & \cdots & b_{i,j}a_{t,I_{j,t}} \end{pmatrix}.$$

Here, $I_{j,l}$ denotes the l-th element of I_j in order.

But then C' evidently factors into the product of two matrices which can again be regarded as having $(m-s)^2$ blocks of size $t \times t$. The first matrix has (i, j)-block equal to $b_{i,j} \cdot \mathrm{Id}_t$. In the second matrix, each diagonal block (i, i) equal to the I_i-submatrix of A, and every off-diagonal block is zero. Thus the determinant of C' is

$$p_{I_1} \cdots p_{I_{m-s}} q_{\{1,\dots,m-s\}}{}^t. \qquad \square$$

Example 4.9. Let $r = m = n = 4$ and $s = 1$, so $t = 2$. A maximal submatrix C' of C corresponds to a choice of 6 entries of the 4×4 matrix X. Suppose we pick the B-pure submatrix given by

$$\begin{pmatrix} * & * & \cdot & \cdot \\ * & \cdot & * & \cdot \\ \cdot & * & * & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

i.e., the entries $(1, 1)$, $(1, 2)$, $(2, 1)$, $(2, 3)$, $(3, 2)$, $(3, 3)$. Then C' is

$$\begin{array}{c} \\ (1,1) \\ (1,2) \\ (2,1) \\ (2,2) \\ (3,1) \\ (3,2) \end{array}
\begin{array}{c}
\begin{array}{cccccc} (1,1) & (1,2) & (2,1) & (2,3) & (3,2) & (3,3) \end{array} \\
\begin{pmatrix}
b_{11}a_{11} & b_{11}a_{12} & b_{12}a_{11} & b_{12}a_{13} & b_{13}a_{12} & b_{13}a_{13} \\
b_{11}a_{21} & b_{11}a_{22} & b_{12}a_{21} & b_{12}a_{23} & b_{13}a_{22} & b_{13}a_{23} \\
b_{21}a_{11} & b_{21}a_{12} & b_{22}a_{11} & b_{22}a_{13} & b_{23}a_{12} & b_{23}a_{13} \\
b_{21}a_{21} & b_{21}a_{22} & b_{22}a_{21} & b_{22}a_{23} & b_{23}a_{22} & b_{23}a_{23} \\
b_{31}a_{11} & b_{31}a_{12} & b_{32}a_{11} & b_{32}a_{13} & b_{33}a_{12} & b_{33}a_{13} \\
b_{31}a_{21} & b_{31}a_{22} & b_{32}a_{21} & b_{32}a_{23} & b_{33}a_{22} & b_{33}a_{23}
\end{pmatrix}
\end{array}$$

and it factors as the product

$$\begin{pmatrix} b_{11} & & b_{12} & & b_{13} & \\ & b_{11} & & b_{12} & & b_{13} \\ b_{21} & & b_{22} & & b_{23} & \\ & b_{21} & & b_{22} & & b_{23} \\ b_{31} & & b_{32} & & b_{33} & \\ & b_{31} & & b_{32} & & b_{33} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & & & & \\ a_{21} & a_{22} & & & & \\ & & a_{11} & a_{13} & & \\ & & a_{21} & a_{23} & & \\ & & & & a_{12} & a_{13} \\ & & & & a_{22} & a_{23} \end{pmatrix}.$$

This shows that $\det C' = p_{12}p_{13}p_{23}q_{123}^2$, as claimed in Lemma 4.8.

5. Smoothness and connectedness

5A. ***Smoothness and connectedness for $\boldsymbol{F_k(D^r_{m,n})}$.*** Now we are equipped to describe which Fano schemes $\boldsymbol{F}_k(D^r_{m,n})$ are smooth and irreducible. We restate the theorem:

Theorem 1.2. *Let $1 \le k < (r-1)n$.*

(i) *The Fano scheme $\boldsymbol{F}_k(D^r_{m,n})$ is smooth if and only if $k > (r-2)n$.*

(ii) *$\boldsymbol{F}_k(D^r_{m,n})$ is irreducible if and only if $m \ne n$ and $k > (r-2)n + m - r + 1$.*

Proof. For part (i), suppose that $k > (r-2)n$. Then, for $0 \le s \le r-2$, we have $k > \kappa(s) - (m-s-1)$. Indeed, the function $\kappa(s) - (m-s-1)$ is convex as a function of s, and yields $(r-2)n$ and $(r-2)m$ when evaluated at $s = r-2$ and $s = 0$, respectively. Since

$$\kappa(s) - (n-r+s) = \kappa(s-1) - (m-(s-1)-1),$$

it follows that conditions (4) and (5) of Theorem 3.1 hold for all $0 \le s \le r-1$. By Corollary 4.3, we may conclude that all torus fixed points of $\boldsymbol{F}_k(D^r_{m,n})$ are smooth, so $\boldsymbol{F}_k(D^r_{m,n})$ must be smooth.

If on the other hand $k \le (r-2)n$, consider any point of $\boldsymbol{F}_k(D^r_{m,n})$ corresponding to an $m \times n$ matrix of linear forms whose first $(m-r)+2$ rows contain at most one nonzero entry. This is a point of both $\mathfrak{C}_k(r-1)$ and $\mathfrak{C}_k(r-2)$, which are distinct irreducible components by Corollary 4.3 and Proposition 4.5. Hence $\boldsymbol{F}_k(D^r_{m,n})$ is not smooth there.

For part (ii), assume $k > (r-2)n + (m-r) + 1$ and $m \ne n$. Then we claim that the only compression component appearing in $\boldsymbol{F}_k(D^r_{m,n})$ is $\mathfrak{C}_k(r-1)$. Indeed, we have that $\kappa(r-2) = (r-2)n + m - r + 1$ and $\kappa(r-2) - \kappa(0) = (r-2)(n-m-1) \ge 0$. Since $\kappa(s)$ is a convex function, we have $k > \kappa(s)$ for $s = 0, \dots, r-2$. Now Proposition 2.3 shows that all torus fixed points of $\boldsymbol{F}_k(D^r_{m,n})$ must lie in $\mathfrak{C}_k(r-1)$. Since these points are smooth by part (i), $\mathfrak{C}_k(r-1)$ is the only irreducible component of $\boldsymbol{F}_k(D^r_{m,n})$.

On the other hand, if $k \le (r-2)n+(m-r)+1$, then $k \le \kappa(r-2)$ so $\mathfrak{C}_k(r-2)$ is also a component. Likewise, if $n=m$ then both $\mathfrak{C}_k(0)$ and $\mathfrak{C}_k(r-1)$ are components, since $k \le (r-1)n-1$ by assumption, and $\kappa(0)=\kappa(r-1)=(r-1)n-1$. □

Corollary 5.1. *If $k > (r-2)n$, then $\boldsymbol{F}_k(D^r_{m,n})$ is the disjoint union of compression space components $\mathfrak{C}_k(s)$.*

Proof. The Fano scheme $\boldsymbol{F}_k(D^r_{m,n})$ is smooth by Theorem 1.2(i). Let Z be any irreducible component of $\boldsymbol{F}_k(D^r_{m,n})$. By Remark 2.5, Z intersects a compression space component $\mathfrak{C}_k(s)$. But since $\boldsymbol{F}_k(D^r_{m,n})$ is smooth, we must have $Z=\mathfrak{C}_k(s)$. □

Remark 5.2. It also follows from [Beasley 1987] that, under the assumption $k \ge \kappa(r-2)$, the only components of $\boldsymbol{F}_k(D^r_{m,n})$ are $\mathfrak{C}_k(s)$ for $s \in \{r-1,\ r-2,\ 1,\ 0\}$. Furthermore, [de Seguins Pazzis 2013, Theorem 10] implies that if $k \ge \kappa(r-3)$, then the only components of $\boldsymbol{F}_k(D^r_{m,n})$ are $\mathfrak{C}_k(s)$ for $s \in \{r-1,\ r-2,\ r-3,\ 2,\ 1,\ 0\}$.

We can also give a near-complete description of when $\boldsymbol{F}_k(D^r_{m,n})$ is connected.

Theorem 5.3. *Suppose that $1 \le k < (r-1)n$.*

(i) *If there is some $s=0,\dots,r-2$ such that $k \le \kappa(s)$ and such that the conditions (4) and (5) in Theorem 3.1 hold, then $\boldsymbol{F}_k(D^r_{m,n})$ is disconnected.*

(ii) *If not, then $\boldsymbol{F}_k(D^r_{m,n})$ is either connected or has exactly two connected components. Furthermore, $\boldsymbol{F}_k(D^r_{m,n})$ is connected if*

$$k > \kappa(0) \quad \text{or} \quad k \le \kappa(0)-(m-r+1)(r-1).$$

In particular, if $r=m$ and either condition (4) or condition (5) fails to hold for every $s=0,\dots,r-2$ such that $k \le \kappa(s)$, then $\boldsymbol{F}_k(D^r_{m,n})$ is connected.

Proof. Given $s \in \{0,\dots,r-2\}$ such that $k \le \kappa(s)$ and (4) and (5) hold, we want to show that $\boldsymbol{F}_k(D^r_{m,n})$ is disconnected. Since $k \le \kappa(s)$, the scheme $\boldsymbol{F}_k(D^r_{m,n})$ contains an irreducible component $\mathfrak{C}_k(s)$. Since (4) and (5) hold by assumption, all of the points of $\mathfrak{C}_k(s)$ are smooth by Corollary 4.3. Thus $\mathfrak{C}_k(s)$ is itself a connected component of $\boldsymbol{F}_k(D^r_{m,n})$. On the other hand, since $k \le (r-1)n-1=\kappa(r-1)$, the scheme $\boldsymbol{F}_k(D^r_{m,n})$ also has an irreducible component $\mathfrak{C}_k(r-1)$, distinct from $\mathfrak{C}_k(s)$ by Proposition 4.5. Hence $\boldsymbol{F}_k(D^r_{m,n})$ is disconnected.

We now prove part (ii). Suppose instead that, for every $s=0,\dots,r-2$ such that $k \le \kappa(s)$, at least one of conditions (4) and (5) holds. Let us first show that the compression components $\mathfrak{C}_k(s)$ form either one or two connected components; since all irreducible components of $\boldsymbol{F}_k(D^r_{m,n})$ meet some compression component by Remark 2.5, the same would be true for $\boldsymbol{F}_k(D^r_{m,n})$.

Since $k \le \kappa(r-1)$, we again have that $\mathfrak{C}_k(r-1)$ is an irreducible component of $\boldsymbol{F}_k(D^r_{m,n})$. Now let $S \subseteq \{0,\dots,r-2\}$ be the set of numbers s such that

$\kappa(s)-(m-s-1)\geq k$. We claim that a number s belonging to S means precisely that both $\mathfrak{C}_k(s)$ and $\mathfrak{C}_k(s+1)$ appear as components of $\boldsymbol{F}_k(D^r_{m,n})$, and the two components meet. Indeed, they would intersect at a torus fixed point of the form

$$\begin{pmatrix} 0 & \cdots & 0 & 0 & & \\ \vdots & \ddots & \vdots & \vdots & & \\ 0 & \cdots & 0 & 0 & & \\ 0 & \cdots & 0 & & & \\ & & & & & \\ & & & & & \end{pmatrix},$$

where the upper left $(m-s)\times(s+1+n-r)$ block, along with the top upper left $(m-s-1)\times(s+2+n-r)$ block, is zero, and furthermore all but $k+1$ of the unmarked entries are zero. (The point is that the number of blank entries in the matrix above is $\kappa(s)-(m-s-1)+1\geq k+1$, i.e., we can actually fit $k+1$ independent forms into the blanks.)

Furthermore, the only compression space components appearing in $\boldsymbol{F}_k(D^r_{m,n})$, apart from $\mathfrak{C}_k(r-1)$, have the form $\mathfrak{C}_k(s)$ and $\mathfrak{C}_k(s+1)$ for $s\in S$. This is because, if $k\leq\kappa(s)$ for some $0<s<r-1$ with $s\notin S$ then, since (4) doesn't hold, (5) must hold. Then we must have

$$k\leq\kappa(s)-(n-r+s)=\kappa(s-1)-(m-(s-1)-1),$$

so $s-1\in S$.

One can check that $\kappa(s)-(m-s-1)$ is a convex function of s and that it always takes a value when $s=r-2$ at least as large as that when $s=0$. We have three cases: If S is empty, then $\mathfrak{C}_k(r-1)$ is the only compression component, so $\boldsymbol{F}_k(D^r_{m,n})$ is connected. Next, S could be a single interval $S=\{a,\ a+1,\ldots,r-2\}$. In that case, by our characterization of S, the compression components are precisely $\mathfrak{C}_k(a),\ldots,\mathfrak{C}_k(r-1)$, and they are again all connected. Otherwise, S must have the form $s=\{0,\ldots,a\}\sqcup\{b,\ldots,r-2\}$. In this last case, the compression components are $\mathfrak{C}_k(0),\ldots,\mathfrak{C}_k(a+1)$, $\mathfrak{C}_k(b),\ldots,\mathfrak{C}_k(r-1)$, and they form at most two connected components.

Now we prove that the extra conditions specified in (ii) show that the union of the compression components is connected (and so $\boldsymbol{F}_k(D^r_{m,n})$ is connected). First, if $k>\kappa(0)$, then the last case above can't occur, so $\boldsymbol{F}_k(D^r_{m,n})$ is connected. If instead $k\leq\kappa(0)-(m-r+1)(r-1)$, then $\mathfrak{C}_k(0)$ and $\mathfrak{C}_k(r-1)$ both appear, and they must intersect. Indeed, they meet at a torus fixed point in which the first $1+n-r$ columns and first $m-(r-1)$ rows are zero. Together with the previous paragraph, this implies that $\boldsymbol{F}_k(D^r_{m,n})$ actually has only one connected component.

In particular, if $m = r$, then $\boldsymbol{F}_k(D^r_{m,n})$ is always connected. We have already proved this if $k > \kappa(0)$. If $k \le \kappa(0)$ then, by the assumption of part (ii), condition (4) holds, i.e., $k \le \kappa(0) - (m-1) = \kappa(0) - (m-r+1)(r-1)$, so we are again done by the previous analysis. □

Theorem 1.3, which characterizes exactly when $\boldsymbol{F}_k(D^m_{m,n})$ is connected, follows immediately from specializing Theorem 5.3 to the case $r = m$.

Remark 5.4. In the special case $r = m$, the proof of Theorem 5.3 implies the following statement: *if $\boldsymbol{F}_k(D^m_{m,n})$ is disconnected, then it has an irreducible component which is its own connected component.* This statement is not always true when $r < m$. Indeed, consider the example $n = m = 8$, $r = 7$, $k = 40$. Then $\boldsymbol{F}_k(D^r_{m,n})$ consists only of the components $\mathfrak{C}_k(0)$, $\mathfrak{C}_k(1)$, $\mathfrak{C}_k(r-2)$, $\mathfrak{C}_k(r-1)$ by Remark 5.2. The components $\mathfrak{C}_k(0)$ and $\mathfrak{C}_k(1)$ intersect, as do $\mathfrak{C}_k(r-2)$ and $\mathfrak{C}_k(r-1)$, and there are no other intersections. Hence the Fano scheme is disconnected, but there is no irreducible connected component. This makes it difficult to detect disconnectedness of $\boldsymbol{F}_k(D^r_{m,n})$ without complete knowledge of its irreducible components.

In fact, we don't know if there is a situation in which the compression components form two distinct connected components but some other component of $\boldsymbol{F}_k(D^r_{m,n})$ connects them. The first case that is unknown to us is $n = m = 5$, $k = 10$, and $r = 4$. In this case, the compression components split into two connected components as in the proof of Theorem 5.3, but k is not large enough to use Remark 5.2 to guarantee that no other types of components appear.

We have seen that connectedness of $\boldsymbol{F}_k(D^r_{m,n})$ can be nonmonotonic with k (Table 1). But it is nevertheless true that, if k is sufficiently small, then $\boldsymbol{F}_k(D^r_{m,n})$ is connected.

Corollary 5.5. *If*

$$k \le m(r-2) - \frac{((n-m)-(r-2))^2}{4},$$

then the Fano scheme $\boldsymbol{F}_k(D^r_{m,n})$ is connected.

Proof. If we pick k such that $k \le \kappa(s) - (m-s-1)$ for all s, then we would ensure that the set S in the proof of Theorem 5.3 is $\{0, \dots, r-2\}$ and hence $\boldsymbol{F}_k(D^r_{m,n})$ is connected. Now,

$$\kappa(s) - (m-s-1) = s^2 + s(n-m-r+2) + m(r-2)$$

is minimized when $s = -(n-m-r+2)/2$, when it takes on value

$$m(r-2) - \frac{((n-m)-(r-2))^2}{4}. \qquad \square$$

5B. ***Smoothness and connectedness for $\boldsymbol{F}_k(P^r_{m,n})$.*** We can now determine exactly when $\boldsymbol{F}_k(P^r_{m,n})$ is smooth, as stated in Theorem 1.4.

Theorem 1.4. *Let $k \geq 1$. The Fano scheme $\boldsymbol{F}_k(P^r_{m,n})$ is smooth if and only if $n = 2$ or $k > (r-2)n+1$.*

Proof. If $n = 2$, then $\boldsymbol{F}_k(P^r_{m,n})$ is either empty or the Fano scheme of lines on a smooth quadric surface, which is the union of two disjoint copies of $\mathbb{P}^1$. Suppose instead that $n > 2$ and $k > (r-2)n+1$. Suppose there is an s-compression component $\mathfrak{C}_k(\sigma, \tau)$ appearing in $\boldsymbol{F}_k(P^r_{m,n})$, i.e., suppose $k \leq \kappa(s)$. We would like to apply Theorem 3.2 to show that all the points of $\mathfrak{C}_k(\sigma, \tau)$ are smooth points of $\boldsymbol{F}_k(P^r_{m,n})$. That would mean that all torus fixed points of $\boldsymbol{F}_k(P^r_{m,n})$ are smooth by Proposition 2.3, and so $\boldsymbol{F}_k(P^r_{m,n})$ must be smooth.

First, let's check that the hypotheses of Theorem 3.2 hold for our choice of k, r, and s. If $r = 2$, then $k \geq 2$ by our assumption on k, and $s = 0$ or $s = r-1$ are the only possibilities. On the other hand, if $r > 2$, then we have $k \geq 5$ as desired. Next, we claim that $s+1+n-r \geq 3$ if $s \neq 0$, and $m-s \geq 3$ if $s \neq r-1$. The first condition could only fail if $s = 1$ and $r = m = n$; but in this case $\kappa(1) = (r-2)n+1 < k$, contradicting that $k \leq \kappa(1)$. Similarly, the second condition could only fail if $s = r-2$ and $r = m$, but in this case $\kappa(r-2) = (r-2)n+1 < k$, again a contradiction.

Next, we always have that for $0 \leq s \leq r-2$, $k > \kappa(s) - (m-s-2)$. Indeed, the function $\kappa(s) - (m-s-2)$ is convex as a function of s, and yields $(r-2)n+1$ and $(r-2)m+1$ when evaluated at $s = r-2$ and $s = 0$, respectively. Furthermore, since

$$\kappa(s) - (n-r+s-1) = \kappa(s-1) - (m-(s-1)-2),$$

it follows that conditions (6) and (7) of Theorem 3.2 hold for all $0 \leq s \leq r-1$.

Hence, we can apply Theorem 3.2 to all compression components appearing in $\boldsymbol{F}_k(P^r_{m,n})$. By Theorem 3.2, at any torus fixed point of $\boldsymbol{F}_k(P^r_{m,n})$ which is an s-compression space, the tangent space has dimension bounded above by $(k+1)(\kappa(s)-k)$. But, by Proposition 4.4, this point is in a compression space component of exactly this dimension. We conclude that all torus fixed points of $\boldsymbol{F}_k(P^r_{m,n})$ are smooth, so $\boldsymbol{F}_k(P^r_{m,n})$ must be smooth.

Suppose instead that $k \leq (r-2)n+1$ and $n \neq 2$; we must show that $\boldsymbol{F}_k(P^r_{m,n})$ is singular. First, let us assume that $k > 1$. Then, by Proposition 4.4, $\mathfrak{C}_k(r-1)$ is an irreducible component and has dimension $(k+1)(\kappa(r-1)-k)$. Consider a point P of $\mathfrak{C}_k(r-1)$ corresponding to an $m \times n$ matrix whose first $m-r+2$ rows have the form

$$\begin{pmatrix} 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & z_0 & z_1 \end{pmatrix}, \tag{17}$$

that is, the first $m-r+1$ rows are zero, and in the $(m-r+2)$-nd row, all entries vanish except the two on the right. Now, perturbing the relations among the entries of P not in the first $m-r+1$ rows gives $(k+1)(\kappa(r-1)-k)$ tangent directions as in the proof of Lemma 3.4. However, we may also perturb P by changing its first $m-r+2$ rows to

$$\begin{pmatrix} 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \varepsilon z_0 & -\varepsilon z_1 \\ 0 & \cdots & 0 & z_0 & z_1 \end{pmatrix}$$

to get an additional tangent vector. Hence, $\boldsymbol{F}_k(P^r_{m,n})$ is not smooth at this point.

To conclude, we must deal with the case $k=1$, still supposing that $n \neq 2$ and $k \leq (r-2)n+1$. First, suppose that $r>2$ and consider the point $P \in \mathfrak{C}_1(r-1)$ corresponding to the $m \times n$ matrix whose bottom row is $(\, z_0 \;\; z_1 \;\; 0 \;\; \cdots \;\; 0\,)$, with 0 entries everywhere else. Then by Lemma 3.4, the dimension of the tangent space to $\boldsymbol{F}_k(P^r_{m,n})$ at P is $2(mn-2)$. Consider instead the point $P' \in \mathfrak{C}_1(r-1)$ corresponding to the $m \times n$ matrix

$$\begin{pmatrix} A \\ C \end{pmatrix}$$

with A an $(m-r+1)\times n$ block of zeroes and $C = C(z_0, z_1)$ (for $s = r-1$); see (11). A straightforward calculation with Lemma 3.4 gives $\dim T_{P'}\boldsymbol{F}_k(P^r_{m,n}) < 2(mn-2)$. Hence, $\boldsymbol{F}_k(P^r_{m,n})$ must be singular at P.

Finally, we now assume that $k=1$ and $r=2$. Let $P \in \boldsymbol{F}_k(P^r_{m,n})$ correspond to the $m \times n$ matrix of (17). There is a first-order deformation of this linear space given by perturbing P to

$$\begin{pmatrix} 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \varepsilon z_0 & -\varepsilon z_1 \\ 0 & \cdots & \varepsilon z_0 & z_0 & z_1 \end{pmatrix}.$$

This cannot be lifted to second order: there are obstruction terms $\varepsilon^2 z_0^2$ and $-\varepsilon^2 z_0 z_1$ which cannot simultaneously be canceled out. Indeed, the only possibility for canceling either obstruction is by including an order-two term at the upper left of the lower 2×3 block of the above matrix, and these terms do not agree. Hence, $\boldsymbol{F}_1(P^r_{m,n})$ is not smooth at P. □

Corollary 5.6. *If $k > (r-2)n+1$, then all components of $\boldsymbol{F}_k(P^r_{m,n})$ are of the form $\mathfrak{C}_k(\sigma, \tau)$.*

Proof. If $k > (r-2)n+1$ then, by Theorem 1.4, $\boldsymbol{F}_k(P^r_{m,n})$ is smooth. By the arguments in the proof of Theorem 1.4 above, each $\mathfrak{C}_k(\sigma, \tau)$ is actually an irreducible component. Since any irreducible component of $\boldsymbol{F}_k(P^r_{m,n})$ intersects some $\mathfrak{C}_k(\sigma, \tau)$, all components must have this form. □

Our result on the connectedness of $\boldsymbol{F}_k(P^r_{m,n})$ is similar to the case for $\boldsymbol{F}_k(D^r_{m,n})$:

Theorem 5.7. *Suppose that* $1 \le k < (r-1)n$.

(i) *Suppose there is some* $s = 0, \ldots, r-1$ *satisfying the conditions of Theorem 3.2 and satisfying (6) and (7) as well. Then* $\boldsymbol{F}_k(P^r_{m,n})$ *is disconnected.*

(ii) *Conversely,* $\boldsymbol{F}_k(P^r_{m,n})$ *is connected if* $k \le \max\{\kappa(0), \kappa(r-2)\}$ *and*:

  (a) *for each integer* s *with* $0 < s < r-1$ *satisfying* $k \le \kappa(s)$, *we have* $k \le \kappa(s) - \min\{m-s-1, n-r+s\}$; *and*

  (b) *if* $k \le \kappa(0)$, *then* $k \le \kappa(0) - (m-r+1)(r-1)$.

Proof of Theorem 5.7. If the hypotheses in (i) are met then, by Theorem 3.2 and Proposition 4.4, $\mathfrak{C}_k(\sigma, \tau)$ is an irreducible component whose points are all smooth points of $\boldsymbol{F}_k(P^r_{m,n})$ for any $\sigma \subset \{1, \ldots, n\}$, $\tau \subset \{1, \ldots, m\}$ with $|\tau| = s$ and $|\sigma| = s+1+n-r$. Each of these irreducible components $\mathfrak{C}_k(\sigma, \tau)$ is thus a connected component of $\boldsymbol{F}_k(P^r_{m,n})$, so this Fano scheme is not connected.

Suppose on the other hand that the hypotheses in (ii) hold. Then, as in the proof of Theorem 5.3, we may connect any irreducible component of $\boldsymbol{F}_k(P^r_{m,n})$ to some $\mathfrak{C}_k(\sigma, \tau)$, where $|\tau| = r-1$. Hence, to show connectedness, we must simply connect all subschemes of the form $\mathfrak{C}_k(\sigma, \tau)$ with $|\tau| = r-1$.

If $k \le \kappa(0)$, then we may consider the subscheme $\mathfrak{C}_k(\{1, \ldots, n-r+1\}, \{\})$. By assumption (ii)(b), this clearly intersects every subscheme of the form $\mathfrak{C}_k(\sigma, \tau)$ where $|\tau| = r-1$. Likewise, if $k \le \kappa(r-2)$, we may assume $r > 2$ for otherwise we are done by the previous case. Then, by assumption (ii)(a), we have

$$k \le \kappa(r-2) - \min\{m-r+1, n-2\} \le \kappa(r-2) - (m-r+1).$$

Then we may connect subschemes of the form $\mathfrak{C}_k(\sigma, \tau)$, where $|\tau| = r-1$, via subschemes of the form $\mathfrak{C}_k(\sigma', \tau')$, where $|\tau'| = r-2$. Hence, in both cases, $\boldsymbol{F}_k(P^r_{m,n})$ is connected. □

Corollary 5.8. *If*

$$k \le m(r-2) - \frac{((n-m)-(r-2))^2}{4},$$

then the Fano scheme $\boldsymbol{F}_k(P^r_{m,n})$ *is connected.*

Proof. Similarly to the proof of Corollary 5.5, the bound on k implies that $k \le \kappa(s) - (m-s-1)$ for all s. Thus all compression components $C_k(\sigma, \tau)$ appear. Further, just as in the proof of Theorem 5.7, for each $s = 0, \ldots, r-2$, the components of the form $\mathfrak{C}_k(\sigma, \tau)$ with $|\tau| = s+1$ are connected via the components of the form $\mathfrak{C}_k(\sigma', \tau')$ with $|\tau'| = s$. □

6. Fano schemes of lines

To start out this section, we will prove Theorem 1.1, giving a complete description of the components of $\boldsymbol{F}_1(D^r_{m,n})$, expanding on [Eisenbud and Harris 1988, Corollary 2.2].

Theorem 1.1. *The Fano scheme $\boldsymbol{F}_1(D^r_{m,n})$ has exactly r irreducible components, of dimensions*

$$\delta(s) + 2(\kappa(s) - 1) \quad \text{for } 0 \le s \le r-1.$$

In particular, if $m = n$, then each irreducible component of $\boldsymbol{F}_1(D^r_{n,n})$ has dimension $(n-r)(r-2) + 2nr - n - 5$. If $r > 2$, then all components intersect pairwise. Furthermore, if $r = m = n$, then $\boldsymbol{F}_1(D^n_{n,n})$ is a reduced local complete intersection.

Remark 6.1. To clarify, we are claiming in Theorem 1.1 that $\boldsymbol{F}_1(D^r_{m,n})$ has exactly r *minimal* primary components; there may well be embedded components in addition to these. On the other hand, in the hypersurface case $r = m = n$, our theorem implies that $\boldsymbol{F}_1(D^r_{m,n})$ is reduced, so in that case there can't be any embedded components.

Proof of Theorem 1.1. By considering the Kronecker canonical form of a pencil of matrices [Gantmacher 1959, Section XII.4], it follows that every point of $\boldsymbol{F}_1(D^r_{m,n})$ is contained in a subscheme of the form $\mathfrak{C}_1(s)$ for $s = 0, \ldots, r-1$. These form exactly r distinct components of $\boldsymbol{F}_1(D^r_{m,n})$ of the desired dimension, by Corollary 4.3 and Proposition 4.5. In particular, if $m = n$, then each of the r components has dimension $\delta(s) + 2(\kappa(s) - 1) = (n-r)(r-2) + 2nr - n - 5$. It is also easy to see that, unless $r = 2$, any two components $\mathfrak{C}_1(s)$ and $\mathfrak{C}_1(s')$ intersect at a torus fixed point that has all zero entries in its upper left $(m-s) \times (s+1+n-r)$ block as well as in its upper left $(m-s') \times (s'+1+n-r)$ block.

If $m = n = r$, then $D^r_{m,n}$ is a hypersurface of degree n, and $\boldsymbol{F}_1(D^n_{n,n})$ is the zero locus of a global section of a rank $n+1$ vector bundle on $\mathrm{Gr}(2, n^2)$; see, e.g., [Eisenbud and Harris 2013, Proposition 8.4]. Moreover, each of the n components $\boldsymbol{F}_1(D^r_{m,n})$ has dimension $2n^2 - n - 5$, implying that $\boldsymbol{F}_1(D^n_{n,n})$ has codimension $n+1$ in $\mathrm{Gr}(2, n^2)$. Therefore, $\boldsymbol{F}_1(D^r_{m,n})$ is a local complete intersection. Furthermore, it must be reduced, since it is a local complete intersection and is generically reduced by Theorem 3.1. □

Remark 6.2. We would like to prove a similar result for the Fano scheme $\boldsymbol{F}_1(P^r_{m,n})$, especially in the case $r = m = n$. A modification of the argument used to prove Theorem 3.2 may be used to show that, for a general point η of any $\mathfrak{C}_1(\sigma, \tau)$,

$$\dim T_\eta \boldsymbol{F}_1(P^r_{m,n}) \le \delta(s) + 2(\kappa(s) - 1).$$

n	Degree of $\boldsymbol{F}_1(D^n_{n,n})$
2	4
3	2754
4	97943936
5	91842552457500
6	1905481100678765027040

Table 3. Degrees of Fano schemes of lines.

In the special case $r = m = n$, the right-hand side again simplifies to $2n^2 - n - 5$, which, as above is the expected dimension of $\boldsymbol{F}_1(P^n_{n,n})$. To complete the argument that $\boldsymbol{F}_1(P^n_{n,n})$ has the expected dimension (and is thus a reduced local complete intersection) one would need to show that every irreducible component of $\boldsymbol{F}_1(P^n_{n,n})$ contains one of the subschemes $\mathfrak{C}_1(\sigma, \tau)$.

Our computer calculations show that $\boldsymbol{F}_1(P^n_{n,n})$ indeed has the expected dimension for $n = 3$ and 4. We conjecture that in fact $\boldsymbol{F}_1(P^n_{n,n})$ has the expected dimension for all n; see Conjecture 8.1.

Proposition 6.3. *The Fano scheme $\boldsymbol{F}_1(D^n_{n,n})$ has degree equal to*

$$\int_{\mathrm{Gr}(2,n^2)} c_{n+1}(\mathrm{Sym}^n\, S^*) \cdot \mathscr{L}^{2n^2-n-5},$$

where S denotes the tautological rank-2 subbundle and $\mathscr{L}$ the tautological line bundle of $\mathrm{Gr}(2, n^2)$. *For $n \leq 6$, these values are recorded in Table 3.*

Proof. By Theorem 1.1, $\boldsymbol{F}_1(D^n_{n,n})$ has the expected codimension of $n+1$ in $\mathrm{Gr}(2, n^2)$. Hence, [Eisenbud and Harris 2013, Proposition 8.4] implies that the Chow class of $\boldsymbol{F}_1(D^n_{n,n})$ in the Chow ring of $\mathrm{Gr}(2, n^2)$ is just $c_{n+1}(\mathrm{Sym}^n\, S^*)$, and the claim follows.

The degrees for $n \leq 6$ may now be explicitly computed using Schubert calculus. We do this using the `Schubert2` package of *Macaulay2* [Grayson and Stillman 1996]:

```
loadPackage "Schubert2"
G = flagBundle({2,n^2-2});
(S,Q)=G.Bundles;
c= chern(n+1,symmetricPower(n,dual S));
L=chern_1 tautologicalLineBundle G;
integral (c*L^(2*n^2-n-5))
```

These are shown in Table 3. □

	Component	Dimension	Smooth?
$\boldsymbol{F}_1(D^3_{3,3})$	$\mathfrak{C}_1(0)$	10	No
	$\mathfrak{C}_1(1)$	10	No
	$\mathfrak{C}_1(2)$	10	No
$\boldsymbol{F}_2(D^3_{3,3})$	$\mathfrak{C}_2(0)$	11	No
	$\mathfrak{C}_2(1)$	10	No
	$\mathfrak{C}_2(2)$	11	No
	$\mathfrak{C}^*$	8	Yes
$\boldsymbol{F}_3(D^3_{3,3})$	$\mathfrak{C}_3(0)$	10	Yes
	$\mathfrak{C}_3(1)$	8	Yes
	$\mathfrak{C}_3(2)$	10	Yes
$\boldsymbol{F}_4(D^3_{3,3})$	$\mathfrak{C}_4(0)$	7	Yes
	$\mathfrak{C}_4(1)$	4	Yes
	$\mathfrak{C}_4(2)$	7	Yes
$\boldsymbol{F}_5(D^3_{3,3})$	$\mathfrak{C}_5(0)$	2	Yes
	$\mathfrak{C}_5(2)$	2	Yes

Table 4. Irreducible components of $\boldsymbol{F}_k(D^3_{3,3})$ for $1 \le k \le 5$.

Remark 6.4. We cannot use these methods to compute the degree of $\boldsymbol{F}_k(D^n_{n,n})$ for $k > 1$. Indeed, the dimension of $\boldsymbol{F}_k(D^n_{n,n})$ is larger than the expected dimension, as a quick computation shows.

Remark 6.5. If one can show that $\dim \boldsymbol{F}_1(P^n_{n,n}) = \dim \boldsymbol{F}_1(D^n_{n,n})$, then the above degree computations hold for the Fano scheme $\boldsymbol{F}_1(P^n_{n,n})$ as well. In particular, they hold for the cases $n = 3$ and 4, as in Remark 6.2.

7. Fano schemes for 3 x 3 matrices

7A. ***The Fano schemes $\boldsymbol{F}_k(D^3_{3,3})$.*** Using the results of the previous sections, we may glean quite a bit of information about $\boldsymbol{F}_k(D^3_{3,3})$, $1 \le k \le 5$. It is never irreducible, it is smooth if and only if $k = 4$ or 5, and it is disconnected again if and only if $k = 4$ or 5. Furthermore, the degree of $\boldsymbol{F}_5(D^3_{3,3})$ is 18, and the degree of $\boldsymbol{F}_1(D^3_{3,3})$ is 2754.

We may use the results of [Atkinson 1983] to actually describe all nonembedded irreducible components of $\boldsymbol{F}_k(D^3_{3,3})$. Let $\mathfrak{C}^*$ be the $\mathrm{GL}_3 \times \mathrm{GL}_3$-orbit closure of a general 3×3 antisymmetric matrix in $\boldsymbol{F}_2(D^3_{3,3})$. It follows from [Atkinson 1983] that any point of $\boldsymbol{F}_k(D^3_{3,3})$ is either a subspace of a compression space or, in the case $k = 2$, a point of $\mathfrak{C}^*$; see also [Eisenbud and Harris 1988, Theorem 1.1].

Proposition 7.1. *The irreducible components of $\boldsymbol{F}_k(D^3_{3,3})$ are exactly the ones described in Table 4.*

Proof. That these are exactly the components follows from the above discussion, Corollary 4.3, and Proposition 4.5. The dimension calculations, with the exception of $\mathfrak{C}^*$, follow from Corollary 4.3 as well. The dimension of $\mathfrak{C}^*$ follows from a straightforward calculation of the stabilizer of $\mathrm{GL}_3 \times \mathrm{GL}_3$ at the point corresponding to a general antisymmetric matrix, which has dimension 10.

Regarding smoothness, the components $\mathfrak{C}_k(s)$ are smooth if and only if $k \geq 3$, by Theorem 4.1. To show that the component $\mathfrak{C}^*$ is smooth, it suffices to show that it is smooth at its torus fixed points. We claim that the only fixed points in $\mathfrak{C}^*$ are $(S_3 \times S_3)$-equivalent to the point P in Example 7.3 below. Indeed, let Q be a general 3×3 antisymmetric matrix. A straightforward calculation shows that, for any A, $B \in \mathrm{GL}_3$, the space of linear forms spanned by the first row of $A \cdot Q \cdot B$ is at most two-dimensional. Hence, the $\mathrm{GL}_3 \times \mathrm{GL}_3$ orbit of Q does not intersect the Plücker chart containing the torus fixed point

$$P' = \begin{pmatrix} * & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

so the orbit closure $\mathfrak{C}^*$ does not contain this fixed point, or any fixed point $(S_3 \times S_3)$-equivalent to it. Furthermore, $\mathfrak{C}^*$ cannot contain a fixed point $(S_3 \times S_3)$-equivalent to

$$P'' = \begin{pmatrix} * & * & 0 \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}.$$

Indeed, P'' is in the same $GL_3 \times GL_3$ orbit as

$$\begin{pmatrix} z_0 & z_1 & z_2 \\ 0 & 0 & z_2 \\ 0 & 0 & 0 \end{pmatrix},$$

whose T-orbit closure contains P'. The same argument excludes any torus fixed points $(S_3 \times S_3)$-equivalent to P'^T or P''^T. Hence, all fixed points in $\mathfrak{C}^*$ are of the desired form.

Now, an explicit calculation as in the example below shows that $\mathfrak{C}^*$ is smooth at P. Hence, $\mathfrak{C}^*$ is smooth at all its torus fixed points, and thus smooth. □

Remark 7.2. Even though all irreducible components of $\boldsymbol{F}_3(D^3_{3,3})$ are smooth, the Fano scheme itself is not smooth, since the components have nonempty intersection.

Example 7.3 (a formal neighborhood in $\boldsymbol{F}_2(D^3_{3,3})$). We will analyze a neighborhood of the torus fixed point

$$P = \begin{pmatrix} * & * & 0 \\ * & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

in $\boldsymbol{F}_2(D^3_{3,3})$. In this case, it is advantageous to consider a *formal* neighborhood, since the defining equations are easier to calculate.

First, let us assume that $\operatorname{char}\mathbb{K}=0$. Using *Macaulay2* [Grayson and Stillman 1996] and the package VersalDeformations [Ilten 2012], we calculate a formal neighborhood of this fixed point:

```
loadPackage "VersalDeformations";
A=QQ[x_1..x_9];
M=genericMatrix(R,3,3);
S=A/ideal det M;
J=ideal(x_3,x_5..x_9);
(F,R,G,C)=localHilbertScheme(gens J);
```

Primary decomposition of the obstruction equations gives two 11-dimensional components, one 10-dimensional component, four 8-dimensional components, and one 7-dimensional component:

```
decomp=primaryDecomposition(sub(ideal sum G,QQ[gens ring G_0]));
apply(decomp,i->dim i)
```

Closer inspection shows that the 11- and 10-dimensional components are nonembedded, as is one 8-dimensional component. Furthermore, these components are all smooth. The other three 8-dimensional components are embedded in the 10-dimensional component, and the 7-dimensional component is embedded in the two 11-dimensional components.

We draw two conclusions. First, $\boldsymbol{F}_2(D^3_{3,3})$ is nonreduced. Second, all irreducible components of $\boldsymbol{F}_2(D^3_{3,3})$ (with their reduced structure) are smooth at P. In particular, $\mathfrak{C}^*$ is smooth at P.

Even if $\operatorname{char}\mathbb{K}>0$, the above calculation shows that $\mathfrak{C}^*$ is smooth at P. Indeed, it still follows from the calculation that $\boldsymbol{F}_2(D^3_{3,3})$ contains a 10-dimensional subscheme Z contained in $\mathfrak{C}^*$ and smooth at P. But, since $\dim\mathfrak{C}^*=10$, this is just $\mathfrak{C}^*$.

7B. ***The Fano schemes $\boldsymbol{F}_k(P^3_{3,3})$.*** Using the results of the previous sections, we may also say quite a bit about $\boldsymbol{F}_k(P^3_{3,3})$ for $1\le k\le 5$. It is never irreducible, and it is smooth if and only if $k=5$. It is disconnected if $k=5$ and connected if $k\le 3$. By the discussion below, we will see that it is disconnected if $k=4$. Furthermore, $\boldsymbol{F}_5(P^3_{3,3})$ consists just of 6 points, and hence has degree 6, while, by Remark 6.5, the degree of $\boldsymbol{F}_1(P^3_{3,3})$ is 2754.

4-planes on the 3×3 *permanental hypersurface.* As an example of the rich geometry that can occur, we will now give a detailed description of $\boldsymbol{F}_4(P^3_{3,3})$. We will do this by looking at the local structure of $\boldsymbol{F}_4(P^3_{3,3})$ around torus fixed points. For simplicity, we assume $\operatorname{char}\mathbb{K}=0$, although we expect the calculation to hold in arbitrary characteristic.

Let's start by considering the fixed point

$$P=\begin{pmatrix}0&0&*\\0&0&*\\*&*&*\end{pmatrix}.$$

There are nine fixed points in the $S_3\times S_3$ orbit of P. Explicit calculation on the corresponding Plücker chart shows that, around P, $\boldsymbol{F}_4(P^3_{3,3})$ is cut out by the equations $r_1r_2=s_1s_2=0$ in $\operatorname{Spec}\mathbb{K}[r_1,r_2,s_1,s_2]$, with a parametrization given by

$$\begin{pmatrix}s_1z_0+r_1z_2-r_1s_1z_4 & s_2z_0-r_1z_3+r_1s_2z_4 & z_0\\ -s_1z_1+r_2z_2+r_2s_1z_4 & -s_2z_1-r_2z_3-r_2s_2z_4 & z_1\\ z_2 & z_3 & z_4\end{pmatrix}.$$

On this chart, $\boldsymbol{F}_4(P^3_{3,3})$ thus decomposes into four copies of $\mathbb{A}^2$.

Each copy of $\mathbb{A}^2$ compactifies in $\boldsymbol{F}_4(P^3_{3,3})$ to a $\mathbb{P}^1\times\mathbb{P}^1$, as can be seen by considering transition maps to the charts containing the other fixed points in the $S_3\times S_3$ orbit of P. In fact, one can parametrize these components explicitly. For example, for the component given locally by $r_2=s_2=0$, we can identify a point $((u_0:u_1),(v_0:v_1))$ with the space of matrices that are orthogonal to the four matrices

$$\begin{pmatrix}0&0&0\\0&1&0\\0&0&0\end{pmatrix},\quad \begin{pmatrix}0&0&0\\u_0&0&u_1\\0&0&0\end{pmatrix},\quad \begin{pmatrix}0&v_0&0\\0&0&0\\0&v_1&0\end{pmatrix},\quad \begin{pmatrix}u_0v_0&0&-u_1v_0\\0&0&0\\-u_0v_1&0&u_1v_1\end{pmatrix}$$

with respect to the entrywise inner product. On the affine chart of this $\mathbb{P}^1\times\mathbb{P}^1$ given by $u_0\neq 0$, $v_0\neq 0$, there is an isomorphism with the $r_2=s_2=0$ copy of $\mathbb{A}^2$ from above, gotten by setting $s_1=u_1/u_0$ and $r_1=v_1/v_0$. The other three standard affine charts of this $\mathbb{P}^1\times\mathbb{P}^1\subset\boldsymbol{F}_4(P^3_{3,3})$ are just copies of $\mathbb{A}^2$ containing other fixed points in the $S_3\times S_3$ orbit of P.

The $S_3\times S_3$ action on rows and columns gives a total of nine copies of $\mathbb{P}^1\times\mathbb{P}^1$; these are exactly the irreducible components containing a fixed point equivalent to P. The previous parametrization shows that they are all embedded in a linear subspace of $\operatorname{Gr}(5,9)$ by $\mathbb{O}(2,2)$, and hence each has degree 8. Together, they form a connected component of $\boldsymbol{F}_4(P^3_{3,3})$. Its dual intersection complex, drawn on the fundamental domain of a torus, is pictured in Figure 1. The fixed points, corresponding to the nine squares, are labeled.

Next, consider the fixed point

$$Q=\begin{pmatrix}0&0&0\\0&*&*\\*&*&*\end{pmatrix}.$$

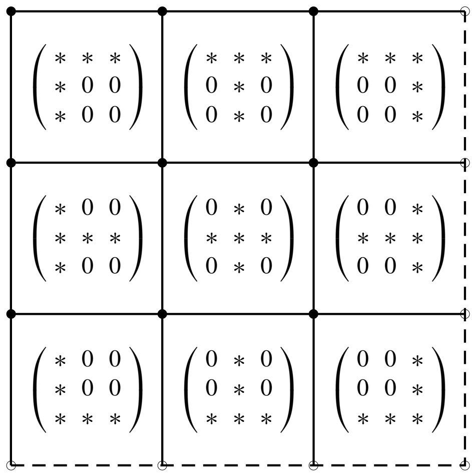


Figure 1. Dual intersection complex for the first connected component of $\boldsymbol{F}_4(P^3_{3,3})$.

There are 18 fixed points in the $S_3 \times S_3$ orbit of Q and, after transposition, one gets 18 more fixed points. Together with the orbit of P above, this covers all fixed points in $\boldsymbol{F}_4(P^3_{3,3})$.

Explicit calculation on the Plücker chart containing Q shows that Q is contained in 4 components of $\boldsymbol{F}_4(P^3_{3,3})$. One component is just $\mathfrak{C}_4(\{1,2,3\},\{2,3\})$, which is isomorphic to $\mathrm{Gr}(5,6) \cong \mathbb{P}^5$, linearly embedded and hence of degree 1. There is an embedded component which is just a fat point supported on Q. The other two components are two-dimensional, locally cut out by the equation $t_1t_2 = 0$ in $\operatorname{Spec} \mathbb{K}[t_1, t_2, t_3]$, with a parametrization given by

$$\begin{pmatrix} -t_1z_2 & t_2t_3z_3 + t_1z_3 - t_3z_0 & -t_2t_3z_4 + t_1z_4 + t_3z_1 \\ -t_2z_2 & z_0 & z_1 \\ z_2 & z_3 & z_4 \end{pmatrix}.$$

By again considering transition maps to other charts, one sees that both of these components compactify in $\boldsymbol{F}_4(P^3_{3,3})$ to $\mathscr{F}_1$, the first Hirzebruch surface.

As above, we can parametrize these components explicitly. Consider the Cox ring $\mathbb{K}[u_0, u_1, v, w]$ of $\mathscr{F}_1$, with the degrees of u_0, u_1, v, w given by the columns of the matrix $\left(\begin{smallmatrix} 1&1&1&0 \\ 0&0&1&1 \end{smallmatrix}\right)$. For the component given locally by $t_2 = 0$, we can identify a point $(u_0 : u_1 : v : w)$ with the space of matrices orthogonal to the four matrices

$$\begin{pmatrix} 0&0&0 \\ 1&0&0 \\ 0&0&0 \end{pmatrix}, \quad \begin{pmatrix} u_0&0&0 \\ 0&0&0 \\ u_1&0&0 \end{pmatrix}, \quad \begin{pmatrix} 0&u_0w&0 \\ 0&v&0 \\ 0&u_1w&0 \end{pmatrix}, \quad \begin{pmatrix} 0&0&u_0w \\ 0&0&-v \\ 0&0&-u_1w \end{pmatrix}$$

Connected component	Type	#	Degree
I	$\mathbb{P}^1 \times \mathbb{P}^1$	9	8
II and III	$\mathbb{P}^5$	3	1
	$\mathscr{F}_1$	9	8
	Fat point	18	–

Table 5. Primary decomposition for the connected components of $\boldsymbol{F}_4(P^3_{3,3})$.

with respect to the entrywise inner product. A parametrization of the other component is gotten by an appropriate permutation of the above. It follows from this parametrization that both copies of $\mathscr{F}_1$ are embedded by a divisor of type $(3, 2)$ in the given basis of the Picard group, which is simply the anticanonical class. In particular, each of these components has degree 8.

By letting $S_3 \times S_3$ act, we get a total of nine copies of $\mathscr{F}_1$, each of which contains four fixed points in the $S_3 \times S_3$ orbit of Q. The dual intersection complex is the simplicial complex which is the union of the boundary complexes of three disjoint 2-simplices, with intersections taking place in torus invariant divisors with self-intersection zero. The $S_3 \times S_3$ action also gives us 3 copies of $\mathbb{P}^5$, and each $\mathscr{F}_1$ intersects exactly one of these in a line and the other two in a point. Together, these form a second connected component of $\boldsymbol{F}_4(P^3_{3,3})$.

Finally, transposing all matrices gives a third connected component, isomorphic to the second. We conclude that $\boldsymbol{F}_4(P^3_{3,3})$ is neither reduced nor connected though, in their reduced structures, every irreducible component is smooth. Summing this all up, we have the following:

Proposition 7.4. *The Fano scheme $\boldsymbol{F}_4(P^3_{3,3})$ has exactly three connected components. Their primary decompositions are as described in Table 5.*

8. Comparisons and further questions

In our study of Fano schemes of determinants and permanents, we have seen that a good understanding of the components $\mathfrak{C}_k(s)$ and $\mathfrak{C}_k(\sigma, \tau)$ induced by compression spaces can already say a lot about the geometry of the entire Fano scheme. One striking difference between $\boldsymbol{F}_k(D^r_{m,n})$ and $\boldsymbol{F}_k(P^r_{m,n})$ is that, while the components $\mathfrak{C}(s)$ are positive-dimensional, the components $\mathfrak{C}(\sigma, \tau)$ are isolated points. This has the following consequences:

- For large values of k, the dimension of $\boldsymbol{F}_k(D^r_{m,n})$ is greater than that of $\boldsymbol{F}_k(P^r_{m,n})$. In other words, $D^r_{m,n}$ contains "more" high-dimensional linear spaces than $P^r_{m,n}$.
- For large values of k, $\boldsymbol{F}_k(D^r_{m,n})$ has fewer irreducible components than $\boldsymbol{F}_k(P^r_{m,n})$.

For the case of lines on the determinantal and permanental hypersurfaces, however, we conjecture that $P^n_{n,n}$ behaves exactly as $D^n_{n,n}$ does; we have verified that this is the case for $n = 3, 4$.

Conjecture 8.1. *The Fano scheme* $\boldsymbol{F}_1(P^n_{n,n})$ *has dimension* $2n^2 - n - 5$, *that is, the expected dimension, and is a reduced local complete intersection.*

Even though the compression subschemes $\mathfrak{C}_k(s)$ and $\mathfrak{C}(\sigma, \tau)$ give a surprising amount of information about the global structure of our Fano schemes, one way in which they fail to tell the whole story is that they don't by themselves detect connectedness, as far as we know. See Remark 5.4 for a detailed discussion. In particular, we would like to know:

Question 8.2. Is the Fano scheme $\boldsymbol{F}_{10}(D^4_{5,5})$ connected? (See Remark 5.4.) More generally, is there a Fano scheme $\boldsymbol{F}_k(D^r_{m,n})$ which is connected but whose compression components $\mathfrak{C}_k(s)$ are not by themselves connected?

A frequently asked question regarding a moduli space is whether it is *rational* or *unirational*. The components $\mathfrak{C}_k(s)$ of $\boldsymbol{F}_k(D^r_{m,n})$ are, by construction, unirational; furthermore, the single exotic component $\mathfrak{C}^*$ of $\boldsymbol{F}_2(D^3_{3,3})$ is also obviously unirational. We suspect that all irreducible components of $\boldsymbol{F}_k(D^r_{m,n})$ may well be unirational. Rationality is more delicate: we do not even know if the components $\mathfrak{C}_k(s)$ are rational, although we can prove that they are for the cases $s = 0$ and $r - 1$.

Question 8.3. Are the irreducible components of $\boldsymbol{F}_k(D^r_{m,n})$ unirational? Are they rational?

Finally, our choice of the schemes $D^r_{m,n}$ and $P^r_{m,n}$ was motivated by Valiant's conjecture. They can be simultaneously generalized, however, in the following interesting way. Fix any character χ of the symmetric group S_r on r letters. The χ-*immanant* of an $r \times r$ matrix (x_{ij}) is the degree-r polynomial with $r!$ terms,

$$\mathrm{Imm}_\chi = \sum_{\sigma \in S_r} \chi(\sigma) x_{1\sigma(1)} \cdots x_{r\sigma(r)}.$$

Thus the χ-immanant specializes to the determinant and permanent by choosing χ to be the sign or trivial character, respectively. Now consider the scheme $D^{\chi,r}_{m,n} \subset \mathbb{P}^{mn-1}$ cut out by the $r \times r$ χ-immanants of a general $m \times n$ matrix. As χ varies, we have a family of schemes that interpolate between $D^r_{m,n}$ and $P^r_{m,n}$.

Question 8.4. What is the structure of the Fano scheme $\boldsymbol{F}_k(D^{\chi,r}_{m,n})$? Does $\boldsymbol{F}_k(D^{\chi,r}_{m,n})$ behave like $\boldsymbol{F}_k(D^r_{m,n})$ or $\boldsymbol{F}_k(P^r_{m,n})$?

Acknowledgments

We would like to thank Allen Knutson, Luke Oeding, and Bernd Sturmfels for helpful discussions, and Bernd Sturmfels for suggesting this problem to us. We are grateful to the referee for very detailed comments. Chan was supported by a NSF Postdoctoral fellowship. Ilten was partially supported by an AMS–Simons travel grant.

References

[Altman and Kleiman 1977] A. B. Altman and S. L. Kleiman, "Foundations of the theory of Fano schemes", *Compositio Math.* **34**:1 (1977), 3–47. MR 58 #27967 Zbl 0414.14024

[Atkinson 1983] M. D. Atkinson, "Primitive spaces of matrices of bounded rank, II", *J. Austral. Math. Soc. Ser. A* **34**:3 (1983), 306–315. MR 84h:15017 Zbl 0521.15009

[Barth and van de Ven 1978] W. Barth and A. van de Ven, "Fano varieties of lines on hypersurfaces", *Arch. Math.* (*Basel*) **31**:1 (1978), 96–104. MR 80j:14004 Zbl 0383.14003

[Beasley 1987] L. B. Beasley, "Null spaces of spaces of matrices of bounded rank", pp. 45–50 in *Current trends in matrix theory* (Auburn, AL, 1986), edited by F. Uhlig and R. Grone, North-Holland, New York, 1987. MR 88h:15021 Zbl 0656.15008

[Bürgisser et al. 2011] P. Bürgisser, J. M. Landsberg, L. Manivel, and J. Weyman, "An overview of mathematical issues arising in the geometric complexity theory approach to VP $\neq$ VNP", *SIAM J. Comput.* **40**:4 (2011), 1179–1209. MR 2861717 Zbl 1252.68134

[Debarre and Manivel 1998] O. Debarre and L. Manivel, "Sur la variété des espaces linéaires contenus dans une intersection complète", *Math. Ann.* **312**:3 (1998), 549–574. MR 99j:14048 Zbl 0913.14015

[Dieudonné 1949] J. Dieudonné, "Sur une généralisation du groupe orthogonal à quatre variables", *Arch. Math.* **1** (1949), 282–287. MR 10,586l Zbl 0032.10601

[Draisma 2006] J. Draisma, "Small maximal spaces of non-invertible matrices", *Bull. London Math. Soc.* **38**:5 (2006), 764–776. MR 2007j:17010 Zbl 1112.15019

[Eisenbud and Harris 1988] D. Eisenbud and J. Harris, "Vector spaces of matrices of low rank", *Adv. in Math.* **70**:2 (1988), 135–155. MR 89j:14010 Zbl 0657.15013

[Eisenbud and Harris 2013] D. Eisenbud and J. Harris, "3264 and all that: intersection theory in algebraic geometry", unpublished manuscript, 2013, Available at http://isites.harvard.edu/fs/docs/icb.topic720403.files/book.pdf.

[Gantmacher 1959] F. R. Gantmacher, *The theory of matrices, II*, Chelsea, New York, 1959. MR 21 #6372c Zbl 0085.01001

[Grayson and Stillman 1996] D. R. Grayson and M. E. Stillman, "Macaulay2, a software system for research in algebraic geometry", 1996, Available at http://www.math.uiuc.edu/Macaulay2/.

[Harris et al. 1998] J. Harris, B. Mazur, and R. Pandharipande, "Hypersurfaces of low degree", *Duke Math. J.* **95**:1 (1998), 125–160. MR 99j:14043 Zbl 0991.14018

[Hartshorne 1966] R. Hartshorne, "Connectedness of the Hilbert scheme", *Inst. Hautes Études Sci. Publ. Math.* 29 (1966), 5–48. MR 35 #4232 Zbl 0171.41502

[Hartshorne 2010] R. Hartshorne, *Deformation theory*, Graduate Texts in Mathematics **257**, Springer, New York, 2010. MR 2011c:14023 Zbl 1186.14004

[Ilten 2012] N. O. Ilten, "Versal deformations and local Hilbert schemes", *J. Softw. Algebra Geom.* **4** (2012), 12–16. MR 2947667

[Kirkup 2008] G. A. Kirkup, "Minimal primes over permanental ideals", *Trans. Amer. Math. Soc.* **360**:7 (2008), 3751–3770. MR 2009a:13017 Zbl 1152.13021

[Landsberg 2013] J. M. Landsberg, "Geometric complexity theory: an introduction for geometers", preprint, 2013. arXiv 1305.7387v2

[Langer 1997] A. Langer, "Fano schemes of linear spaces on hypersurfaces", *Manuscripta Math.* **93**:1 (1997), 21–28. MR 98g:14048 Zbl 0904.14024

[Laubenbacher and Swanson 2000] R. C. Laubenbacher and I. Swanson, "Permanental ideals", *J. Symbolic Comput.* **30**:2 (2000), 195–205. MR 2001i:13039 Zbl 0966.13007

[Manivel and Mezzetti 2005] L. Manivel and E. Mezzetti, "On linear spaces of skew-symmetric matrices of constant rank", *Manuscripta Math.* **117**:3 (2005), 319–331. MR 2006b:14088 Zbl 1084.14050

[Mukai 1993] S. Mukai, "Curves and Grassmannians", pp. 19–40 in *Algebraic geometry and related topics* (Inchon, 1992), edited by J.-H. Yang and Y. Namikawa, Int. Press, Cambridge, MA, 1993. MR 95i:14032 Zbl 0846.14030

[de Seguins Pazzis 2013] C. de Seguins Pazzis, "The classification of large spaces of matrices with bounded rank", preprint, 2013. arXiv 1004.0298v3

[Valiant 1979] L. G. Valiant, "Completeness classes in algebra", pp. 249–261 in *Conference record of the eleventh annual ACM symposium on theory of computing* (Atlanta, GA, 1979), edited by M. J. Fischer et al., ACM, New York, 1979. MR 83e:68046

Communicated by David Eisenbud
Received 2014-06-10 Revised 2015-01-15 Accepted 2015-02-23

mtchan@math.harvard.edu *Department of Mathematics, Harvard University, Cambridge, MA 02138, United States*

nilten@sfu.ca *Department of Mathematics, Simon Fraser University, Burnaby BC V5A1S6, Canada*

ALGEBRA AND NUMBER THEORY 9:3 (2015)
dx.doi.org/10.2140/ant.2015.9.681

Triple intersection formulas for isotropic Grassmannians

Vijay Ravikumar

Let X be an isotropic Grassmannian of type B, C, or D. In this paper we calculate K-theoretic Pieri-type triple intersection numbers for X: that is, the sheaf Euler characteristic of the triple intersection of two arbitrary Schubert varieties and a special Schubert variety in general position. We do this by determining explicit equations for the projected Richardson variety corresponding to the two arbitrary Schubert varieties, and show that it is a complete intersection in projective space. The K-theoretic Pieri coefficients are alternating sums of these triple intersection numbers, and we hope they will lead to positive Pieri formulas for isotropic Grassmannians.

1. Introduction 681
2. Preliminaries 1: types B and C 685
3. Preliminaries 2: type D 688
4. Result 1: defining $Z_{P,T}$ in type D 690
5. Result 2: $\psi(\pi^{-1}(Y_{P,T})) \subset Z_{P,T}$ 696
6. Result 3: $Z_{P,T} \subset \psi(\pi^{-1}(Y_{P,T}))$ 701
7. The Grothendieck ring 711
8. Computing triple intersection numbers 712
9. Computing Pieri coefficients 719
Acknowledgments 721
References 722

1. Introduction

When studying the ordinary cohomology of an (isotropic) Grassmannian, a *triple intersection number* refers to the number of intersection points of three Schubert varieties in general position. By convention, this number is zero when the triple intersection has positive dimension. Algebraically this number is given as the pushforward of the product of three Schubert classes to the cohomology ring of a single point.

MSC2010: primary 14N15; secondary 19E08, 14M15.
Keywords: triple intersection numbers, isotropic Grassmannian, orthogonal Grassmannian, submaximal Grassmannian, Richardson variety, projected Richardson variety, Pieri rule, K-theoretic Pieri formula, K-theoretic triple intersection.

Given three Schubert varieties in general position, let Z denote their scheme-theoretic triple intersection. The corresponding *K-theoretic triple intersection number* is the sheaf Euler characteristic of Z; that is, the pushforward of the product of the three Schubert classes to the Grothendieck ring of a point. We denote this number by $\chi(Z)$. If Z is finite, then, just as in cohomology, $\chi(Z)$ is equal to the number of points in Z (since these finitely many points are reduced, by Kleiman's transversality theorem [1974]). If Z has positive dimension however, then $\chi(Z)$ can be a nonzero (and possibly negative) integer.

In either setting, the triple intersection numbers determine the structure constants for multiplication with respect to the Schubert basis. These structure constants are known as *Littlewood–Richardson coefficients*, and in ordinary cohomology they are equal to triple intersection numbers. In K-theory however, the Littlewood–Richardson coefficients are alternating sums of triple intersection numbers.

An arbitrary Schubert class can be written as an integer polynomial in certain *special Schubert classes*, which (in cohomology) are closely related to the Chern classes of the tautological quotient bundle on the Grassmannian in question. A triple intersection number is said to be of *Pieri-type* if one of the three Schubert classes is a special Schubert class. Similarly, a *Pieri coefficient* refers to a Littlewood–Richardson coefficient occurring in the product of an arbitrary Schubert class and a special Schubert class.

In this paper, we determine K-theoretic Pieri-type triple intersection numbers for all isotropic Grassmannians of types B, C, and D. Our results generalize [Buch and Ravikumar 2012], in which similar calculations are carried out for the cominuscule Grassmannians, that is, for the type-A Grassmannian $\mathrm{Gr}(m, \mathbb{C}^N)$, the maximal odd orthogonal Grassmannian $\mathrm{OG}(m, \mathbb{C}^{2m+1})$, and the Lagrangian Grassmannian $\mathrm{LG}(m, \mathbb{C}^{2m})$.

1A. ***Methods and results.*** Let ω be a skew-symmetric or symmetric nondegenerate bilinear form on $\mathbb{C}^N$, where $N \geq 2$. Fix a basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_N$ for $\mathbb{C}^N$ that is *isotropic* in the sense that

$$\omega(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{i+j,N+1} \quad \text{for } 1 \leq i \leq j \leq N.$$

Note that if ω is symmetric, then $\omega(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{i+j,N+1}$ for all i and j in the integer interval $[1, N]$. If ω is skew-symmetric (which can only happen when N is even), then $\omega(\boldsymbol{e}_i, \boldsymbol{e}_j) = -\delta_{i+j,N+1}$ for $i > j$.

For any subspace $\Sigma \subset \mathbb{C}^N$, we define $\Sigma^\perp := \{\boldsymbol{w} \in \mathbb{C}^N : \omega(\boldsymbol{v}, \boldsymbol{w}) = 0 \text{ for all } \boldsymbol{v} \in \Sigma\}$. We say Σ is *isotropic* if $\Sigma \subset \Sigma^\perp$. Given a positive integer $m \leq N/2$, the isotropic Grassmannian $\mathrm{IG}_\omega(m, \mathbb{C}^N)$ is defined as

$$\mathrm{IG}_\omega(m, \mathbb{C}^N) := \{\Sigma \in \mathrm{Gr}(m, \mathbb{C}^N) : \Sigma \subset \Sigma^\perp\}. \tag{1}$$

This projective variety parametrizes isotropic m-planes in $\mathbb{C}^N$. It is said to have Lie-type C when ω is skew-symmetric (in which case N is even), Lie-type B when ω is symmetric and N is odd, and Lie-type D when ω is symmetric and N is even.

In order to define Schubert varieties in $X := \mathrm{IG}_\omega(m, \mathbb{C}^N)$, we must fix some flags on $\mathbb{C}^N$. We define the *standard flag* $E_\bullet$ on $\mathbb{C}^N$ by $E_j := \langle e_1, \dots, e_j \rangle$, the span of the first j basis vectors. In types B and C, we define the opposite flag $E_\bullet^{\mathrm{op}}$ by $E_j^{\mathrm{op}} := \langle e_{N+1-j}, \dots, e_N \rangle$, the span of the last j basis vectors. A more complicated type-D definition is given in Section 3.

Given $\Sigma \in X$, the *Schubert symbol of* Σ relative to $E_\bullet$,

$$\mathfrak{s}(\Sigma) := \{c \in [1, N] : \Sigma \cap E_c \supsetneq \Sigma \cap E_{c-1}\},$$

records the steps c in $E_\bullet$ at which the intersection $\Sigma \cap E_c$ jumps dimension. Note that the set $\mathfrak{s}(\Sigma)$ has cardinality m, and that if $c \in \mathfrak{s}(\Sigma)$ then $N + 1 - c \notin \mathfrak{s}(\Sigma)$, since Σ is isotropic. In general, a subset $P \subset [1, N]$ of cardinality m is a *Schubert symbol*[1] if $c + d \neq N + 1$ for any $c, d \in P$. We let $\Omega(X)$ denote the set of all Schubert symbols for X.

Given a Schubert symbol P, we define the *Schubert variety* $X_P := X_P(E_\bullet)$ to be the closure in X of the *Schubert cell* $X_P^\circ(E_\bullet) := \{\Sigma \in X : \mathfrak{s}(\Sigma) = P\}$. We say X_P is a Schubert variety relative to the flag $E_\bullet$. We also define the *opposite Schubert variety* X^P to be the unique Schubert variety relative to the opposite flag $E_\bullet^{\mathrm{op}}$ that intersects X_P at a single point. For the special Schubert varieties, we adopt an additional indexing convention, writing $X_{(r)}$ to denote the special Schubert variety of codimension r in X. Given Schubert symbols P and T, we write $T \preceq P$ if $X_T \subset X_P$. The resulting partial order on the set of Schubert symbols, known as the *Bruhat order*, is described combinatorially in Sections 2 and 3.

The transverse intersection of two Schubert varieties is known as a *Richardson variety*. Associated to Schubert symbols P and T we have a Richardson variety $Y_{P,T} := X_P \cap X^T$, which is nonempty if and only if $T \preceq P$. Since $[\mathcal{O}_{X_P}] \cdot [\mathcal{O}_{X^T}] = [\mathcal{O}_{Y_{P,T}}]$ (see, e.g., [Brion 2005]), the K-theoretic Pieri-type triple intersection numbers can be written

$$\chi([\mathcal{O}_{Y_{P,T}}] \cdot [\mathcal{O}_{X_{(r)}}]), \tag{2}$$

where $\chi : K(X) \to \mathbb{Z}$ is the sheaf Euler characteristic map. These numbers are nonzero only when $T \preceq P$.

We can reinterpret this triple intersection number by means of the following incidence relation, which consists of the two-step isotropic flag variety $\mathrm{IF}_\omega(1, m, \mathbb{C}^N)$ whose natural projections we denote by ψ (to the Grassmannian $IG_\omega(1, \mathbb{C}^N)$) and

[1] Schubert symbols are sometimes known as *jump sequences*, and in [Buch et al. 2009] they are referred to as *index sets*.

π (to the Grassmannian $IG_\omega(m, \mathbb{C}^N)$).

$$\begin{tikzcd} \mathrm{IF}_\omega(1,m,\mathbb{C}^N) \arrow[r, "\psi"] \arrow[d, "\pi"'] & \mathrm{IG}_\omega(1,\mathbb{C}^N) \subset \mathbb{P}^{N-1} \\ \mathrm{IG}_\omega(m,\mathbb{C}^N) & \end{tikzcd}$$

In particular, we make use of the projected Richardson variety $\psi(\pi^{-1}(Y_{P,T})) \subset \mathbb{P}^{N-1}$, which is the (projectivization of the) union of all m-planes in the Richardson variety $Y_{P,T}$. Projected Richardson varieties like $\psi(\pi^{-1}(Y_{P,T}))$ have a number of nice geometric properties. He and Lam [2011] relate these varieties to the K-theory of affine Grassmannians, and it has been proved by Billey and Coskun [2012], and by Knutson, Lam, and Speyer [Knutson et al. 2014], that they are Cohen–Macaulay with rational singularities and that the projection map is cohomologically trivial, in the sense that $\psi_*[\mathbb{O}_{\pi^{-1}(Y_{P,T})}] = [\mathbb{O}_{\psi(\pi^{-1}(Y_{P,T}))}]$. By this last fact, along with the projection formula, the calculation of the triple intersection number (2) amounts to showing that $\psi(\pi^{-1}(Y_{P,T}))$ is a complete intersection in $\mathbb{P}^{N-1}$ and determining the equations that define it.

A description of the projected Richardson variety $\psi(\pi^{-1}(Y_{P,T}))$ is carried out in [Buch et al. 2009], but in the special case that the Schubert symbols P and T satisfy a relation $P \to T$. Roughly speaking, this relation signifies that T shows up in some cohomological Pieri product involving P. The relation $P \to T$ requires that $T \preceq P$, and $T \preceq P$ is a more general condition. We note that for $P \not\to T$, the K-theoretic triple intersection numbers $\chi([\mathbb{O}_{Y_{P,T}}] \cdot [\mathbb{O}_{X_{(r)}}])$ need not vanish (in contrast to the cohomological triple intersection numbers), and are therefore essential ingredients for the K-theoretic Pieri coefficients.

When X is a Grassmannian of Lie type B or C, Buch, Kresch, and Tamvakis [Buch et al. 2009] define a complete intersection $Z_{P,T} \subset \mathbb{P}^{N-1}$ for Schubert symbols $T \preceq P$, and prove that the projected Richardson variety $\psi(\pi^{-1}(Y_{P,T}))$ is contained in it. They attempt to extend the definition of $Z_{P,T}$ to the type-D Grassmannian, but use an erroneous definition of Schubert varieties, resulting in a definition of $Z_{P,T}$ that only makes sense in the special case that $P \to T$.

The first result of this paper, presented in Section 4, is to provide a corrected definition of $Z_{P,T}$ in the general setting that $T \preceq P$, and to show that it is a complete intersection of linear and quadratic hypersurfaces. This process involves new combinatorics of Schubert symbols, such as the notion of an *exceptional cut*.

The second result, presented in Section 5, is that $\psi(\pi^{-1}(Y_{P,T})) \subset Z_{P,T}$ for any Schubert symbols $T \preceq P$ in a Grassmannian of Lie-type B, C, or D. This result generalizes [Buch et al. 2009, Lemma 5.1], in which this statement is proved in types B and C only.

The third result, presented in Section 6, is that given a type-B, C, or D Grassmannian and arbitrary Schubert symbols $T \preceq P$, we have $Z_{P,T} \subset \psi(\pi^{-1}(Y_{P,T}))$. We prove this result by constructing a smaller Richardson variety contained in $Y_{P,T}$ that projects surjectively onto $Z_{P,T}$.

Combining these results, we arrive at the main theorem of this paper:

Theorem 1.1. *Let X be a Grassmannian of Lie-type B, C, or D. For any Schubert symbols $T \preceq P$, we have $Z_{P,T} = \psi(\pi^{-1}(Y_{P,T}))$.*

By Theorem 1.1, we know exactly which equations define the projected Richardson variety in all three Lie types. These equations allow for a pleasant calculation of the triple intersection numbers, which we carry out in Section 8. In Section 9, we describe how K-theoretic Pieri coefficients are calculated as alternating sums of these triple intersection numbers. Taken together, the results of this paper complete the story of Pieri-type triple intersection numbers for Grassmannians. We hope this approach will soon lead to a positive Pieri formula.

2. Preliminaries 1: types B and C

2A. ***Schubert symbols.*** Let $X := \mathrm{IG}_\omega(m, \mathbb{C}^N)$ be a Grassmannian of type C or B, where $N := 2n$ or $N := 2n+1$, depending on whether X is of type C or B respectively. In the former case, we will also denote X by $\mathrm{SG}(m, 2n)$ and refer to it as a *symplectic Grassmannian*. In the latter case, we will also denote X by $\mathrm{OG}(m, 2n+1)$ and refer to it as an *odd orthogonal Grassmannian*. Recall that for Schubert symbols T and P in $\Omega(X)$, the relation $T \preceq P$ signifies that $X_T \subset X_P$. This partial order on the set of Schubert symbols has a simple combinatorial description.

Given Schubert symbols $T = \{t_1 < \cdots < t_m\}$ and $P = \{p_1 < \cdots < p_m\}$, we write $T \leq P$ whenever $t_i \leq p_i$ for $1 \leq i \leq m$. By [Buch et al. 2009, Proposition 4.1] we have the following lemma:

Lemma 2.1. *Provided X is of Lie-type B or C, we have $T \leq P$ if and only if $T \preceq P$.*

For any Schubert symbol $P \in \Omega(X)$, let $\overline{P} = \{c \in [1, N] : N+1-c \in P\}$ and $[P] = P \cup \overline{P}$. Also let $|P|$ denote the codimension of the Schubert variety X_P in X.

For each Schubert symbol P, there is a unique dual symbol $P^\vee$ with the property that for any Schubert symbol T, $X_P(E_\bullet) \cap X_T(E_\bullet^{\mathrm{op}})$ is equal to a single point if and only if $T = P^\vee$. The opposite Schubert symbol X^P defined in the introduction is therefore equal to $X_{P^\vee}(E_\bullet^{\mathrm{op}})$. The following lemma, from [Buch et al. 2009, Proposition 4.2], gives a simple description the dual symbol $P^\vee$:

Lemma 2.2. *Provided X is of Lie-type B or C, we have $P^\vee = \overline{P}$ for all Schubert symbols $P \in \Omega(X)$.*

2B. ***Richardson diagrams.*** It is a well-known fact (following from Borel's fixed-point theorem [1956]) that $T \preceq P$ if and only if $X_P \cap X^T$ is nonempty. This variety $Y_{P,T} := X_P \cap X^T$ is connected, and hence reduced and irreducible by Kleiman's transversality theorem [1974] (see also [Richardson 1992]). It is known as a *Richardson variety*.

Given Schubert symbols $T \leq P$, we define the *Richardson diagram* $D(P,T) = \{(j,c) : t_j \leq c \leq p_j\}$, which we represent as an $m \times N$ matrix with a $*$ for every entry in $D(P,T)$ and zeros elsewhere. We say a matrix $(a_{i,j})$ has *shape* $D(P,T)$ if its dimensions are $m \times N$ and $a_{j,c} = 0$ for all $(j,c) \notin D(P,T)$. Given a matrix of shape $D(P,T)$ whose row vectors are independent and orthogonal, its rowspace will be an element of $Y_{P,T}$.

Example 2.3. Any rank-m matrix of shape $D(P,P)$ will have rowspace $\Sigma_P := \langle \boldsymbol{e}_{p_1}, \dots, \boldsymbol{e}_{p_m} \rangle$, which is the only element of $Y_{P,P}$.

Example 2.4. Suppose $P = \{2,3,4,10\}$ and $T = \{1,2,4,6\}$ in SG(4, 10). Suppose $(a_{i,j})$ is a rank-m matrix of shape $D(P,T)$. The rowspace of $(a_{i,j})$ will be in $Y_{P,T}$ if and only if $a_{1,1}a_{4,10} + a_{1,2}a_{4,9} = 0$, $a_{2,2}a_{4,9} + a_{2,3}a_{4,8} = 0$, and $a_{4,7} = 0$. We leave it to the reader to write down explicit entries satisfying these equations. The diagram $D(P,T)$ is

$$\begin{pmatrix} * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & * & * & * & * \end{pmatrix}.$$

Given a Schubert symbol $P = \{p_1, \dots, p_m\}$, let $p_0 = 0$ and $p_{m+1} = N+1$. We won't consider these as actual elements in the Schubert symbol P, but the notation will be useful. Define a *visible cut* through $D(P,T)$ to be any integer $c \in [0,N]$ such that no row of $D(P,T)$ contains stars in both column c and column $c+1$; i.e., such that $p_i \leq c < t_{i+1}$ for some i. We will consider $c = 0$ and $c = N$ to be visible cuts. Define an *apparent cut* to be any integer $c \in [0,N]$ such that c or $N-c$ is a visible cut. In types B and C we define a *cut* in $D(P,T)$ to be synonymous with an *apparent cut*. Let $\mathscr{C}_{P,T}$ be the set of cuts in $D(P,T)$.

An integer c is a *zero column* of $D(P,T)$ if $p_j < c < t_{j+1}$ for some j, since in this case column c of $D(P,T)$ has no stars. An entry (j,c) in $D(P,T)$ is a *lone star* if either

(i) $c \in T$ and c is a cut in $D(P,T)$, or

(ii) $c \in P$ and $c-1$ is a cut in $D(P,T)$.

The simplest example of a lone star occurs when $t_j = p_j = c$ for some j. In this case row j and column c of $D(P,T)$ each contain a single star at (j,c). We define the set $\mathscr{L}_{P,T} \subset [1,N]$ to be the set of integers c such that either

(i) c is a zero column in $D(P, T)$, or

(ii) there exists a lone star in column $N+1-c$.

Finally, we define the set

$$\mathcal{Q}_{P,T} := \begin{cases} [0,n] \cap \mathcal{C}_{P,T} & \text{if } X \text{ is of type } C, \\ ([0,n] \cap \mathcal{C}_{P,T}) \cup \{n+1\} & \text{if } X \text{ is of type } B. \end{cases}$$

Example 2.5. Continuing with Example 2.4, the set of cuts $\mathcal{C}_{P,T}$ is equal to $\{0,3,4,5,6,7,10\}$. Of these, 0, 3, 4, 5, and 10 are visible cuts. Furthermore, 5 is a zero column, $(3,4)$ is a lone star, $\mathcal{L}_{P,T} = \{5,7\}$, and $\mathcal{Q}_{P,T} = \{0,3,4,5\}$.

In types B and C, lone stars take a particularly simple form. Namely:

Proposition 2.6. *Let X be an isotropic Grassmannian of Lie-type B or C. Suppose (j,c) is a lone star in $D(P,T)$. If $c = t_j$ and t_j is an apparent cut, or if $c = p_j$ and $p_j - 1$ is an apparent cut, then either $N+1-c$ is a zero column or $t_j = p_j = c$.*

Proof. Suppose $c = t_j$ is an apparent cut in $D(P,T)$, and that $N+1-c$ is not a zero column. Since $N+1-c$ is not in T *and* not a zero-column, it follows that $N-c$ is not a visible cut. But then c must be a visible cut, so $c = p_j$. A similar argument holds if we start by assuming $c = p_j$. □

Since zero columns are flanked by cuts, we have the following immediate result:

Corollary 2.7. *In types B and C, if $c \in \mathcal{L}_{P,T}$ then c and $c-1$ are both cuts.*

2C. ***The projected Richardson variety.*** We now define a subvariety of $\mathbb{P}^{N-1}$ that will play a key role in the calculation of triple intersection numbers. Let $x_1, \dots, x_N \in (\mathbb{C}^N)^*$ be the dual basis to the isotropic basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_N \in \mathbb{C}^N$. Let $f_0 = 0$, and for $1 \le c \le n$ let $f_c = x_1 x_N + \cdots + x_c x_{N+1-c}$. For example, $f_1 = x_1 x_N$ and $f_2 = x_1 x_N + x_2 x_{N-1}$. In addition, if X is type B, let $f_{n+1} = x_1 x_{2n+1} + \cdots + x_n x_{n+2} + \frac{1}{2} x_{n+1}^2$. Given Schubert symbols $T \preceq P$, let $Z_{P,T} \subset \mathbb{P}^{N-1}$ denote the subvariety defined by the vanishing of the polynomials $\{f_c \mid c \in \mathcal{Q}_{P,T}\} \cup \{x_c \mid c \in \mathcal{L}_{P,T}\}$. We note that, in the type-B case, $Z_{P,T}$ must satisfy the equation $f_{n+1} = 0$ and hence lie in $\mathrm{OG}(1, 2n+1)$, the quadric hypersurface of isotropic lines in $\mathbb{P}^{2n}$.

In fact, $Z_{P,T}$ is a complete intersection in $\mathbb{P}^{N-1}$ cut out by the polynomials:

(a) $f_d - f_c = x_{c+1} x_{N-c} + \cdots + x_d x_{N+1-d}$ if c and d are consecutive elements of $\mathcal{Q}_{P,T}$ such that $d - c \ge 2$.

(b) x_c if $c \in \mathcal{L}_{P,T}$.

We will prove this fact for all three Lie types in Proposition 4.15.

Recall that we have projections π and ψ from the flag variety $\mathrm{IF}_\omega(1, m, \mathbb{C}^N)$ to X and $\mathrm{IG}_\omega(1, \mathbb{C}^N)$, respectively. The variety $\pi^{-1}(Y_{P,T})$ is a Richardson variety in $\mathrm{IG}_\omega(1, m, \mathbb{C}^N)$, and its image $\psi(\pi^{-1}(Y_{P,T}))$ is known as a *projected Richardson* *variety*. We shall prove that the projected Richardson variety $\psi(\pi^{-1}(Y_{P,T}))$ is in fact equal to $Z_{P,T}$. One inclusion is straightforward:

Lemma 2.8. *Given Schubert symbols $T \preceq P$ for a Grassmannian X of Lie-type B or C, we have $\psi(\pi^{-1}(Y_{P,T})) \subset Z_{P,T}$.*

A proof of Lemma 2.8 can be found in [Buch et al. 2009, Lemma 5.1]. This proof is correct for types B and C, but does not go through in type D due to an erroneous definition of the Bruhat order. We supply a corrected proof for all three Lie types in Section 5.

Example 2.9. Continuing with Example 2.5, suppose $\boldsymbol{M}$ is a matrix with shape $D(P,T)$ and independent, isotropic row vectors. Note that any vector in the rowspace of $\boldsymbol{M}$ must satisfy the quadratic equation $x_1x_{10} + x_2x_9 + x_3x_8 = 0$ and the linear equations $x_5 = 0$ and $x_7 = 0$, which are precisely the equations defining $Z_{P,T}$. By Lemma 2.8, *any* vector contained in an m-plane $\Sigma \in Y_{P,T}$ satisfies these equations.

3. Preliminaries 2: type *D*

Consider $\mathbb{C}^{2n+2}$, endowed with a nondegenerate symmetric bilinear form. Let $X := \mathrm{OG}(m, 2n+2)$ denote the *even orthogonal Grassmannian* of isotropic m-planes in $\mathbb{C}^{2n+2}$, where $1 \le m \le n+1$. In this section we describe the Bruhat order for even orthogonal Grassmannians, which is more complicated than in types B and C. We note that definition (1) implies that $\mathrm{OG}(n+1, 2n+2)$ is disconnected. Although we won't go into it here, this fact can help to give a geometric intuition behind the Bruhat order on even orthogonal Grassmannians (see [Ravikumar 2013, Chapter 5] for a detailed description).

For any Schubert symbol $P \in \Omega(X)$, let $\overline{P} = \{c \in [1, 2n+2] : 2n+3-c \in P\}$ and $[P] = P \cup \overline{P}$. As before, let $|P|$ denote the codimension of the Schubert variety X_P in X. We define $\mathfrak{t}(P) \in \{0,1,2\}$ as follows. If $n+1 \in [P]$, then we let $\mathfrak{t}(P)$ be congruent mod 2 to the number of elements in $[1, n+1] \setminus P$. In other words, if $\#([1,n+1] \setminus P)$ is even then $\mathfrak{t}(P) = 0$, and if $\#([1,n+1] \setminus P)$ is odd then $\mathfrak{t}(P) = 1$. Finally, if $\{n+1, n+2\} \cap P = \varnothing$, we set $\mathfrak{t}(P) = 2$.[2]

Proposition 3.1 [Buch et al. 2013, Proposition A.2]. *Given Schubert symbols P and T in $\Omega(\mathrm{OG}(m, 2n+2))$, we have $T \preceq P$ if and only if*

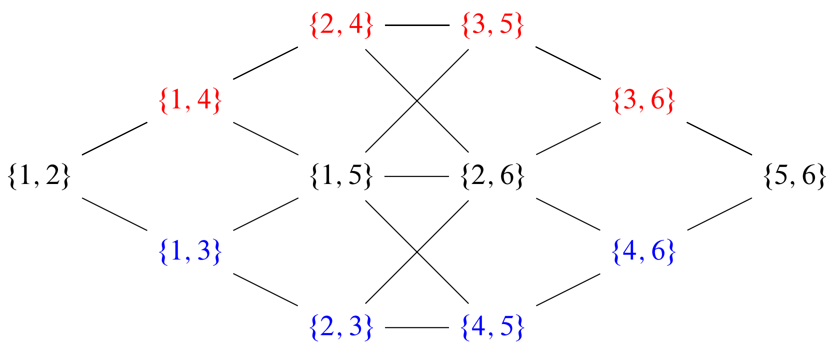
(i) *$T \le P$, and*

(ii) *if there exists $c \in [1,n]$ such that $[c+1, n+1] \subset [P] \cap [T]$ and $\#P \cap [1,c] = \#T \cap [1,c]$, then we have $\mathfrak{t}(P) = \mathfrak{t}(T)$.*

[2] $\mathfrak{t}(P)$ differs slightly from the function $\mathrm{type}(P)$ in [Buch et al. 2013]. Namely, $\mathrm{type}(P) \equiv \mathfrak{t}(P) + 1 \pmod 3$.

By Proposition 3.1, the type-D Bruhat order is not simply the $\leq$ ordering. The following example illustrates the difference.

Example 3.2. The $\preceq$ partial order is shown below for the Schubert symbols on $\mathrm{OG}(2, \mathbb{C}^6)$, which are colored by type. Notice that there are six "missing" edges, which would have occurred had we used the (incorrect) $\leq$ ordering.



We define the opposite flag $E_\bullet^{\mathrm{op}}$ by $E_j^{\mathrm{op}} = \langle e_{2n+3-j}, \dots, e_{2n+2} \rangle$ for $j \neq n+1$ and

$$E_{n+1}^{\mathrm{op}} = \begin{cases} \langle e_{n+2}, e_{n+3}, \dots, e_{2n+2} \rangle & \text{if } n \text{ is odd,} \\ \langle e_{n+1}, e_{n+3}, \dots, e_{2n+2} \rangle & \text{if } n \text{ is even.} \end{cases}$$

This definition guarantees that $E_\bullet$ and $E_\bullet^{\mathrm{op}}$ lie in the same connected component of the variety of complete isotropic flags on $\mathbb{C}^{2n+2}$ (endowed with a nondegenerate symmetric bilinear form), which is disconnected.

Let ι be the permutation of $\{1, \dots, 2n+2\}$ that interchanges $n+1$ and $n+2$ and leaves all other numbers fixed. Given a type-D Schubert symbol $P = \{p_1, \dots, p_m\}$, let $\iota(P) = \{\iota(p_1), \dots, \iota(p_m)\}$. From [Buch et al. 2009, p. 43], we have the following description of the dual symbol $P^\vee$:

Lemma 3.3. *Given a Schubert symbol* $P \in \Omega(\mathrm{OG}(m, 2n+2))$, *we have*

$$P^\vee = \begin{cases} \bar{P} & \textit{when } n \textit{ is odd,} \\ \iota(\bar{P}) & \textit{when } n \textit{ is even.} \end{cases}$$

If the type-D definitions of opposite flags and dual Schubert symbols appear confusing, the following observation may offer some relief:

Observation 3.4. *Let* $X := \mathrm{IG}_\omega(m, N)$ *be a Grassmannian of type* B, C, *or* D. *For any Schubert symbol* P, *we have*

$$X^\circ_{P^\vee}(E_\bullet^{\mathrm{op}}) = \{\Sigma \in \mathrm{IG} : \Sigma \cap \langle e_{p_i}, \dots, e_N \rangle \supsetneq \Sigma \cap \langle e_{p_i+1}, \dots, e_N \rangle\}.$$

Observation 3.4 is obvious unless we are working in $\mathrm{OG}(m, 2n+2)$ and n is even. We illustrate that case in the following example:

Example 3.5. Consider $\mathrm{OG}(1,6)$, and let $P = \{4\}$. Then $P^\vee = \{4\}$, $E_3^{\mathrm{op}} = \langle \boldsymbol{e}_3, \boldsymbol{e}_5, \boldsymbol{e}_6\rangle$, and $E_4^{\mathrm{op}} = \langle \boldsymbol{e}_3, \boldsymbol{e}_4, \boldsymbol{e}_5, \boldsymbol{e}_6\rangle$. By definition,

$$X^\circ_{P^\vee}(E_\bullet^{\mathrm{op}}) = \{\Sigma \in \mathrm{IG} : \Sigma \cap E_4^{\mathrm{op}} \supsetneq \Sigma \cap E_3^{\mathrm{op}}\},$$

which is equal to the set of points in $\mathbb{P}^5$ of the form $\langle(0,0,0,1,*,*)\rangle$, in agreement with Observation 3.4.

By Observation 3.4, any element of the Schubert cell $X^\circ_{P^\vee}(E_\bullet^{\mathrm{op}})$ is the rowspace of an isotropic $m \times N$ matrix $(a_{i,j})$ with $a_{i,p_i} = 1$ for $1 \le i \le m$ and $a_{i,j} = 0$ for $j < p_i$.

Example 3.6. Consider $\mathrm{OG}(3,10)$, and let $P = \{1,4,5\}$. In this case, $P^\vee = \{5,7,10\}$. We can write any element of $X^\circ_{P^\vee}(E_\bullet^{\mathrm{op}})$ as the rowspace of a matrix of the form

$$\begin{pmatrix} 1 & * & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & * & * \end{pmatrix}.$$

Thus $X_{T^\vee}(E_\bullet^{\mathrm{op}}) = X^T$ for any Schubert symbol T, since $X_T \cap X_{T^\vee}(E_\bullet^{\mathrm{op}})$ is a single point. We define the Richardson variety $Y_{P,T} := X_P \cap X^T$. As before, $Y_{P,T} \neq \varnothing$ if and only if $T \preceq P$. We define the Richardson diagram $D(P,T) := \{(j,c) : t_j \le c \le p_j\}$ for any Schubert symbols $T \le P$. This definition holds when $T \npreceq P$, but in this case there cannot exist a matrix of shape $D(P,T)$ whose row vectors are independent and orthogonal (a fact we shall prove in Proposition 4.12).

Example 3.7. There are no matrices $(a_{i,j})$ of shape $D(\{2,5,7,8\},\{1,3,4,6\})$ whose rows span an element of $\mathrm{OG}(4,10)$, because the isotropic relations on the entries are inconsistent:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{2,3} & a_{2,4} & a_{2,5} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{3,4} & a_{3,5} & a_{3,6} & a_{3,7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{4,6} & a_{4,7} & a_{4,8} & 0 & 0 \end{pmatrix}.$$

We leave it to the reader to verify this fact, as well as the fact that $\{1,3,4,6\} \npreceq \{2,5,7,8\}$ in type D.

4. Result 1: defining $Z_{P,T}$ in type D

Let $X := \mathrm{OG}(m, 2n+2)$ be a type-D Grassmannian, let $N := 2n+2$, and let $T \preceq P$ be Schubert symbols in $\Omega(X)$. Visible cuts, apparent cuts, lone stars, and zero columns in $D(P,T)$ are defined exactly as in types B and C. Similarly, $\mathscr{C}_{P,T}$

continues to denote the set of all cuts in $D(P,T)$, and $\mathscr{L}_{P,T}$ continues to denote the set of integers $c \in [1, 2n+2]$ such that either c is a zero column or column $2n+3-c$ contains a lone star. However, in order to define the subvariety $Z_{P,T} \subset \mathbb{P}^{2n+1}$, we must define a new type of cut in $D(P,T)$.

4A. *Exceptional cuts.* If for some i we have $p_i = n+2 \le t_{i+1}$ or $t_i = n+1 \ge p_{i-1}$, we let $n+1$ be a cut in $D(P,T)$, which we will refer to as an *exceptional center cut*. This cut will induce a lone star in column $n+2$ or $n+1$ respectively.

Example 4.1. $P = \{2,4\}$ and $T = \{1,2\}$ in $\mathrm{OG}(2,6)$. $D(P,T)$ is shown below, and has an exceptional center cut. As a result, $(2,4)$ is a lone star, and $3 \in \mathscr{L}_{P,T}$.

$$\left(\begin{array}{ccc|ccc} * & * & 0 & 0 & 0 & 0 \\ 0 & * & * & * & 0 & 0 \end{array}\right).$$

There are additional exceptional cuts in $D(P,T)$. Let $c \in [1,n]$ be a *cut candidate* if $[c+1, n+1] \subset [P] \cap [T]$ and $\#(T \cap [1,c]) = \#(P \cap [1,c]) + 1$. If $\mathfrak{t}(T) \neq \mathfrak{t}(P)$, then c and $N+1-c$ will also be cuts in $D(P,T)$ for each cut candidate c. We'll refer to these as *exceptional cuts* as well. We give several examples of diagrams with exceptional cuts, as the definition is somewhat complicated.

Example 4.2. $P = \{3,6\}$ and $T = \{2,3\}$ in $\mathrm{OG}(2,6)$. $D(P,T)$ is shown below, and $\mathscr{C} = \{0,1,2,3,4,5,6\}$. Of these, 2, 3 (the center cut), and 4 are exceptional cuts. $(1,2)$, $(1,3)$, and $(2,3)$ are all lone stars, and $\mathscr{L}_{P,T} = \{1,4,5\}$.

$$\left(\begin{array}{cc|c|c|cc} 0 & * & * & 0 & 0 & 0 \\ 0 & 0 & * & * & * & * \end{array}\right).$$

Example 4.3. $P = \{3,4,7\}$ and $T = \{1,3,4\}$ in $\mathrm{OG}(3,8)$. $D(P,T)$ is shown below, and $\mathscr{C} = \{0,1,2,3,4,5,6,7,8\}$. Of these, 2, 3, 4 (the center cut), 5, and 6 are exceptional cuts. By finding all the lone stars, one can check that $\mathscr{L}_{P,T} = \{2,5,6,8\}$.

$$\left(\begin{array}{cc|c|c|c|c|cc} * & * & * & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * & * & * & * & 0 \end{array}\right).$$

Example 4.4. $P = \{4,6,8\}$ and $T = \{1,3,5\}$ in $\mathrm{OG}(3,8)$. $D(P,T)$ is shown below, and $\mathscr{C} = \{0,2,6,8\}$. Of these, 2 and 6 are exceptional cuts, and $\mathscr{L}_{P,T} = \varnothing$.

$$\left(\begin{array}{cc|cccc|cc} * & * & * & * & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & * & * & * & * \end{array}\right).$$

Example 4.5. $P = \{4,5,8,9\}$ and $T = \{1,3,4,6\}$ in $\mathrm{OG}(4,10)$. $D(P,T)$ is shown below, and $\mathscr{C} = \{0,1,2,8,9,10\}$. Of these, 2 and 8 are exceptional cuts, and $\mathscr{L}_{P,T} = \{2,10\}$.

$$\left(\begin{array}{cc|cccccc|cc} * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * & * & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * & * & * & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * & * & * & * & 0 \end{array}\right).$$

We relate certain features of the Richardson diagram $D(P,T)$ to the type-D Bruhat order and to the existence of exceptional cuts.

Lemma 4.6. *For any Schubert symbol P, the following conditions are equivalent*:

(1) $[c+1, n+1] \subset [P]$.

(2) $\#([c+1, N-c] \cap P) = n+1-c$.

Proof. Note that $n+1-c = \#([c+1,n+1])$. Because P is an isotropic Schubert symbol, there can be at most $n+1-c$ elements in $[c+1, N-c] \cap P$. Since $[c+1, n+1] \subset [P]$, there are at least that many. $\square$

Lemma 4.7. *Given Schubert symbols $T \leq P$ such that $[c+1, n+1] \subset [T] \cap [P]$, we have*

$$\#([1,c] \cap T) - \#([1,c] \cap P) = \#([N+1-c, N] \cap P) - \#([N+1-c, N] \cap T).$$

Proof. By Lemma 4.6, $\#([c+1, N-c] \cap P) = \#([c+1, N-c] \cap T) = n+1-c$. It follows that

$$\begin{aligned} \#([1,c] \cap P) + \#([N+1-c, N] \cap P) \\ &= m - (n+1-c) \\ &= \#([1,c] \cap T) + \#([N+1-c, N] \cap T). \end{aligned} \qquad \square$$

Lemma 4.7 says that whenever $[c+1, n+1] \subset [T] \cap [P]$, the number of rows crossing from column c to column $c+1$ of $D(P,T)$ is equal to the number of rows crossing from column $N-c$ to column $N+1-c$ of $D(P,T)$. We therefore have the following corollary:

Corollary 4.8. *Given $c \in [1,n]$, suppose $[c+1, n+1] \subset [T] \cap [P]$ and $\mathfrak{t}(P) \neq \mathfrak{t}(T)$ for Schubert symbols $T \leq P$.*

(1) *The first four of the following statements are equivalent*, *and any of them implies the last*:

- $\#([1,c] \cap T) = \#([1,c] \cap P)$.
- $\#([N-c+1, N] \cap T) = \#([N-c+1, N] \cap P)$.
- *c is a visible cut in $D(P,T)$.*

- *$N-c$ is a visible cut in $D(P,T)$.*
- *$T \not\leq P$.*

(2) *The following statements are equivalent*:

- *c and $N-c$ are exceptional cuts in $D(P,T)$.*
- $\#([1,c]\cap T) = \#([1,c]\cap P)+1$.
- $\#([N-c+1,N]\cap P) = \#([N-c+1,N]\cap T)+1$.
- *$D(P,T)$ has exactly one row crossing from column c to column $c+1$.*
- *$D(P,T)$ has exactly one row crossing from column $N-c$ to column $N-c+1$.*

We finish this section by proving that several important properties of Richardson diagrams carry over to the type-D case. In particular, we extend Corollary 2.7 to type D, and then prove in Corollary 4.13 that $(P\cup T)\cap \mathscr{L}_{P,T} = \varnothing$ (a fact that is obvious in types B and C). Once these facts are established, we will be ready to define $Z_{P,T}$.

First we observe that, for any $T\leq P$, $D(P,T)$ and $D(\overline{T},\overline{P})$ have the same cut candidates, by Lemma 4.7. It follows that:

Observation 4.9. $180°$ *rotation of the diagram $D(P,T)$ preserves all cuts, including exceptional cuts. In other words, $\mathscr{C}_{P,T} = \mathscr{C}_{\overline{T},\overline{P}}$.*

We can now prove the type-D version of Corollary 2.7:

Proposition 4.10. *In type D, if $c\in\mathscr{L}_{P,T}$, then c and $c-1$ are both in $\mathscr{C}_{P,T}$.*

Proof. If c is a zero column then the result is clear. Otherwise, it must be the case that $(i, N+1-c)$ is a lone star for some i. By Observation 4.9, we can assume without loss of generality that $N+1-c\leq n+1$.

Case 1: *$N+1-c = t_i$ and t_i is a cut in $D(P,T)$.* We claim that t_i-1 must be a cut as well. If $t_i = p_i$, then $p_{i-1} < t_i$, and we are done. Thus, we only need to consider the case that t_i is an exceptional cut in $D(P,T)$.

If t_i-1 is not a visible cut, then $p_{i-1}\geq t_i$. In fact, if $t_i = n+1$, then $p_{i-1} = t_i$, since that is the only way the exceptional center cut can arise. On the other hand, if $t_i\neq n+1$, then, since $\#([1,t_i]\cap T) = \#([1,t_i]\cap P)+1$, row i of $D(P,T)$ is the *only* row crossing the exceptional cut t_i. In this case too we must have $p_{i-1} = t_i$.

We therefore have $t_i\in[T]\cap[P]$. Furthermore, since row $i-1$ is the only row crossing from column t_i-1 to column t_i, we have $\#([1,t_i-1]\cap T) = \#([1,t_i-1]\cap P)+1$. Thus t_i-1 is also an exceptional cut in $D(P,T)$.

Case 2: *$N+1-c = p_i$ and p_i-1 is a cut in $D(P,T)$.* We claim that p_i must be a cut as well. As before, we can assume that p_i-1 is an exceptional cut in $D(P,T)$.

If p_i is not a visible cut, then $t_{i+1}\leq p_i$. In fact, we must have $t_{i+1} = p_i$, since row i is the only row crossing the exceptional cut p_i-1.

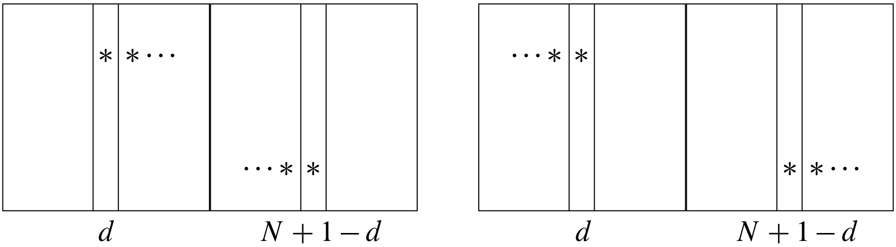


Figure 1. Conflicting lone stars in $D(P,T)$ in Proposition 4.12.

If $p_i = n+1$, then, since $t_{i+1} = p_i$, the diagram $D(P,T)$ has the exceptional center cut $n+1$, and we are done.

If $p_i \neq n+1$, then, since $t_{i+1} = p_i$, row $i+1$ must be the only row crossing from column p_i to column p_i+1. Hence, $\#([1,p_i]\cap T) = \#([1,p_i]\cap P)+1$. Thus p_i is also an exceptional cut in $D(P,T)$. $\square$

Given Schubert symbols $T \leq P$ in $\Omega(\mathrm{OG}(m,2n+2))$ such that $\mathfrak{t}(T)\neq\mathfrak{t}(P)$, we define a *critical window* in $D(P,T)$ to be an interval $[c+1,N-c]$ such that c and $N-c$ are visible cuts in $D(P,T)$, and $[c+1,n+1]\subset[T]\cap[P]$.

Lemma 4.11. *Given Schubert symbols $T < P$ in $\Omega(\mathrm{OG}(m,2n+2))$, we have $T \not\preceq P$ if and only if a critical window exists in $D(P,T)$.*

Proof. If $T\not\preceq P$, then $\mathfrak{t}(P)\neq\mathfrak{t}(T)$ and there exists $c\in[1,n]$ such that $[c+1,n+1]\subset[T]\cap[P]$ and $\#[1,c]\cap P = \#[1,c]\cap T$. By Corollary 4.8, both c and $N-c$ are visible cuts in $D(P,T)$, and hence $[c+1,N-c]$ is a critical window. Conversely, if $D(P,T)$ has a critical window, then it is clear that $T\not\preceq P$. $\square$

The fact that $(P\cup T)\cap\mathscr{L}_{P,T}=\varnothing$ follows easily from the next proposition:

Proposition 4.12. *Given $T<P$ in $\mathrm{OG}(m,2n+2)$, $T\not\preceq P$ if and only if there exists an integer $d\in[1,N]$ such that $D(P,T)$ has lone stars in columns d and $N+1-d$.*

Proof. Suppose columns d and $N+1-d$ of $D(P,T)$ both contain lone stars, and assume $d\leq n+1$. If $d=t_i$ for some i, then $N+1-d=p_j$ for some j, as shown in the left side of Figure 1. It follows that $t_i\neq p_i$, so t_i must be an exceptional cut in $D(P,T)$. Thus $[t_i+1,n+1]\subset[T]\cap[P]$ and $\mathfrak{t}(T)\neq\mathfrak{t}(P)$. Furthermore, row i must be the only row crossing from column t_i to column t_i+1, and hence $p_{i-1}<t_i$, implying that t_i-1 is a visible cut. Therefore, $[t_i,p_j]$ is a critical window in $D(P,T)$. On the other hand, if $d=p_i$ for some i, then $N+1-d=t_j$ for some j, as shown in the right side of Figure 1. In this case, p_i-1 must be an exceptional cut, p_i must be a visible cut, and $[p_i+1,t_i-1]$ must be a critical window in $D(P,T)$. By Lemma 4.11, it follows that $T\not\preceq P$.

On the other hand, if $T \not\preceq P$, then by Lemma 4.11 there exists a critical window $[c+1, N-c]$ in $D(P,T)$. We claim that there exists a (possibly smaller) critical window of the form $[t_i, p_j]$ for some i and j. To see why, note that if $[c+1, N-c]$ does not have the form $[t_i, p_j]$, then either $c+1 = t_i = p_i$ or $N-c = t_j = p_j$. Either way, $[c+2, N-c-1]$ is a smaller critical window. However, this process of shrinking can't continue indefinitely. In particular, if $[c+1, N-c] = [n+1, n+2]$, then the fact that $[n+1, n+2]$ is a critical window implies that $t_i = n+1$ and $p_i = n+2$ for some i.

Finally, note that whenever $[t_i, p_j]$ is a critical window in $D(P,T)$ it must be the case that t_i and $p_j - 1$ are exceptional cuts. Thus (i, t_i) and (j, p_j) are lone stars and $t_i + p_j = N+1$, completing the proof. □

Corollary 4.13. *Given $T \preceq P$, we have $(P \cup T) \cap \mathscr{L}_{P,T} = \varnothing$.*

Proof. Suppose $c \in (P \cup T) \cap \mathscr{L}_{P,T}$. Since c is not a zero column, there exists a lone star in column $N+1-c$ of $D(P,T)$. By Proposition 4.10, c and $c-1$ are both cuts in $D(P,T)$. But then column c contains a lone star as well, since $c \in (P \cup T)$, contradicting Proposition 4.12. □

4B. ***A complete intersection.*** The quadratic equation characterizing isotropic vectors in $\mathbb{C}^{2n+2}$ is $f^D_{n+1} := x_1 x_N + \cdots + x_{n+1} x_{n+2} = 0$. We once again let $\mathscr{Q}_{P,T} = ([0,n] \cap \mathscr{C}) \cup \{n+1\}$. We let $Z_{P,T} \subset \mathbb{P}^{2n+1}$ denote the subvariety cut out by the familiar polynomials $\{f_c \mid c \in \mathscr{Q}_{P,T}\} \cup \{x_c \mid c \in \mathscr{L}_{P,T}\}$, where we let $f_{n+1} = f^D_{n+1}$.

It is not immediately obvious that $Z_{P,T}$ is a complete intersection in $\mathbb{P}^{N-1}$, or even that it is an irreducible subvariety. To prove these facts, we need the following lemma, which we prove in all three Lie types:

Lemma 4.14. *Given Schubert symbols $T \preceq P$ for a Grassmannian X of Lie-type B, C, or D, if $c-1$ and c are both in $\mathscr{Q}_{P,T}$, then $c \in \mathscr{L}_{P,T}$ or $N+1-c \in \mathscr{L}_{P,T}$.*

Proof. If we are working in type B, and $c = n+1$, then $n+1$ must be a zero column, and hence be in $\mathscr{L}_{P,T}$. Otherwise, we can assume that $c \leq \lfloor N/2 \rfloor$. If either c or $N+1-c$ is a zero column in $D(P,T)$ then we are done, so assume neither column is empty.

Note that if $c-1$ is an exceptional cut, then $c \in [T] \cap [P]$. Otherwise, either $c-1$ or $N+1-c$ is a visible cut in $D(P,T)$, and hence $c \in T$ or $N+1-c \in P$ respectively, since neither c nor $N+1-c$ is a zero column in $D(P,T)$.

In all of these cases $c \in [T] \cup [P]$. Therefore (j, c) or $(j, N+1-c)$ is a lone star for some j. It follows that $N+1-c$ or c is in $\mathscr{L}_{P,T}$. □

We can now prove that $Z_{P,T}$ is a complete intersection in types B, C, and D.

Proposition 4.15. *Given Schubert symbols $T \preceq P$ for a Grassmannian X of Lie-type B, C, or D, the variety $Z_{P,T}$ is a complete intersection in $\mathbb{P}^{N-1}$ cut out by the following polynomials*:

(a) $f_d - f_c = x_{c+1}x_{N-c} + \cdots + x_d x_{N+1-d}$, *if c and d are consecutive elements of $\mathcal{Q}_{P,T}$ such that $d - c \geq 2$.*

(b) x_c, *if $c \in \mathcal{L}_{P,T}$.*

Proof. Let $I_{P,T} \subset \mathbb{C}[x_1, \ldots, x_N]$ be the ideal generated by the polynomials of types (a) and (b) mentioned in the statement of this proposition. Note that each of these polynomials is irreducible, and that by Corollary 2.7 and Proposition 4.10 no variable x_i appears in multiple generators.

It follows that $\mathbb{C}[x_1, \ldots, x_N]/I_{P,T}$ is a tensor product over $\mathbb{C}$ of finitely many integral domains. Since $\mathbb{C}$ is algebraically closed, $\mathbb{C}[x_1, \ldots, x_N]/I_{P,T}$ must itself be an integral domain, by [Springer 2009, Lemma 1.5.2]. Hence $I_{P,T}$ is a prime ideal.

Let $I'_{P,T}$ be the ideal generated by the polynomials used to define $Z_{P,T}$: namely, $\{f_c \mid c \in \mathcal{Q}_{P,T}\} \cup \{x_c \mid c \in \mathcal{L}_{P,T}\}$. Note that each of the generators of $I_{P,T}$ is a linear combination of these defining polynomials, and is therefore contained in $I'_{P,T}$.

On the other hand, note that whenever $d-1$ and d are elements of $\mathcal{Q}_{P,T}$, the polynomial $f_d - f_{d-1} = x_d x_{N+1-d}$ is contained in $I_{P,T}$, by Lemma 4.14. Thus if $c < d$ are *any* consecutive elements of $\mathcal{Q}_{P,T}$, then $f_d - f_c \in I_{P,T}$. Now, supposing f_c is one of the quadratic polynomials defining $Z_{P,T}$, let $\{0 = c_0 < c_1 < \cdots < c_s = c\}$ be the complete list of cuts between 0 and c. It follows that $f_c = f_c - f_0 = (f_{c_s} - f_{c_{s-1}}) + (f_{c_{s-1}} - f_{c_{s-2}}) + \cdots + (f_{c_1} - f_{c_0}) \in I_{P,T}$, and therefore that $I'_{P,T} \subset I_{P,T}$.

We have shown that $I'_{P,T} = I_{P,T}$, and hence that $Z_{P,T}$ is the zero set of a prime ideal. It follows that the polynomials used to define $I_{P,T}$ also cut out $Z_{P,T}$ as a complete intersection in $\mathbb{P}^{N-1}$. □

5. Result 2: $\psi(\pi^{-1}(Y_{P,T})) \subset Z_{P,T}$

Let $X := \mathrm{OG}(m, 2n+2)$ and $N := 2n+2$. We would like to show that any vector lying in any subspace $\Sigma \in Y_{P,T}$ satisfies the equations defining $Z_{P,T}$. The equations involving exceptional cuts are the most difficult to verify, so we'll address them first.

Let $Y^\circ_{P,T} = X^\circ_P(E_\bullet) \cap X^\circ_{T^\vee}(E^{\mathrm{op}}_\bullet)$. It is a dense open subset of $Y_{P,T}$ (see [Richardson 1992]), so we can restrict our attention to $\psi(\pi^{-1}(Y^\circ_{P,T}))$:

Proposition 5.1. *Consider Schubert symbols $T \preceq P$ for* $\mathrm{OG}(m, 2n+2)$, *and suppose $c \in [1, n]$ is an exceptional cut in $D(P, T)$. Then $f_c(\boldsymbol{w}) = 0$ for all $\boldsymbol{w} \in \psi(\pi^{-1}(Y^\circ_{P,T}))$.*

Proof. Since c is exceptional, we know $[c+1, n+1] \subset [T] \cap [P]$, $\#P \cap [1, c] + 1 = \#T \cap [1, c]$, and $\mathfrak{t}(P) \neq \mathfrak{t}(T)$.

Let $\ell = n + 1 - c$, and let $E^{(c)} = E_{N-c}/E_c$, which we identify with the span $\langle \boldsymbol{e}_{c+1}, \dots, \boldsymbol{e}_{N-c} \rangle$. Finally, let $\alpha = \#P \cap [1, c]$.

Suppose Σ is an element of $Y^{\circ}_{P,T}$. Since $\Sigma \in X^{\circ}_{P}$, we have $\dim(\Sigma \cap E_c) = \alpha$. Similarly, since $\Sigma \in X^{\circ}_{T^\vee}(E^{\mathrm{op}}_{\bullet})$, we have $\dim(\Sigma \cap E^{\mathrm{op}}_c) = m - (\alpha + \ell + 1)$. Furthermore, $\dim(\Sigma \cap E_{N-c}) = \alpha + \ell$ and $\dim(\Sigma \cap E^{\mathrm{op}}_{N-c}) = m - (\alpha + 1)$. Finally, we know that

$$\begin{aligned}
\dim(\Sigma \cap E^{(c)}) &= \dim(\Sigma \cap E_{N-c} \cap E^{\mathrm{op}}_{N-c}) \\
&\geq \dim(\Sigma \cap E_{N-c}) + \dim(\Sigma \cap E^{\mathrm{op}}_{N-c}) - m \\
&= (\alpha + \ell) + (m - (\alpha + 1)) - m \\
&= \ell - 1.
\end{aligned}$$

Therefore we can choose vectors $\boldsymbol{u}_1$ through $\boldsymbol{u}_m$ spanning Σ such that

$$\begin{aligned}
&\boldsymbol{u}_i \in E_c = \langle \boldsymbol{e}_1, \dots, \boldsymbol{e}_c \rangle && \text{for } 1 \leq i \leq \alpha, \\
&\boldsymbol{u}_i \in E^{(c)} = \langle \boldsymbol{e}_{c+1}, \dots, \boldsymbol{e}_{N-c} \rangle && \text{for } \alpha + 2 \leq i \leq \alpha + \ell, \\
&\boldsymbol{u}_i \in E^{\mathrm{op}}_c = \langle \boldsymbol{e}_{N-c+1}, \dots, \boldsymbol{e}_N \rangle && \text{for } \alpha + \ell + 2 \leq i \leq m.
\end{aligned}$$

In other words, Σ can be represented as the rowspace of a matrix with the following shape (in the sense that all the entries outside the horizontal arrows are zero):

$$\begin{array}{c}
\begin{pmatrix}
\longleftarrow \boldsymbol{u}_1 \longrightarrow & & \\
\vdots & & \\
\longleftarrow \boldsymbol{u}_\alpha \longrightarrow & & \\
\multicolumn{3}{c}{\longleftarrow\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\; \boldsymbol{u}_{\alpha+1} \;-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!\longrightarrow} \\
& \longleftarrow \boldsymbol{u}_{\alpha+2} \longrightarrow & \\
& \vdots & \\
& \longleftarrow \boldsymbol{u}_{\alpha+\ell} \longrightarrow & \\
\multicolumn{3}{c}{\longleftarrow\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\; \boldsymbol{u}_{\alpha+\ell+1} \;-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!\longrightarrow} \\
& & \longleftarrow \boldsymbol{u}_{\alpha+\ell+2} \longrightarrow \\
& & \vdots \\
& & \longleftarrow \boldsymbol{u}_m \longrightarrow
\end{pmatrix} \\
\underbrace{\hphantom{\longleftarrow \boldsymbol{u}_\alpha \longrightarrow}}_{c} \quad \underbrace{\hphantom{\longleftarrow \boldsymbol{u}_{\alpha+2} \longrightarrow}}_{2\ell} \quad \underbrace{\hphantom{\longleftarrow \boldsymbol{u}_{\alpha+\ell+2} \longrightarrow}}_{c}
\end{array}.$$

Furthermore, since $\dim(\Sigma \cap E_{N-c}) = \alpha + \ell$, we can assume without loss of generality that $\boldsymbol{u}_{\alpha+1} \in E_{N-c}$. The matrix with rowspace Σ then has the following shape:

$$\begin{array}{c}
\begin{pmatrix}
\longleftarrow \boldsymbol{u}_1 \longrightarrow & & & \\
\vdots & & & \\
\longleftarrow \boldsymbol{u}_\alpha \longrightarrow & & & \\
\longleftarrow\!\!\!\!\!-\!\!\!\!-\!\!\!\!- & \boldsymbol{u}_{\alpha+1} & \longrightarrow & \\
 & \longleftarrow \boldsymbol{u}_{\alpha+2} \longrightarrow & & \\
 & \vdots & & \\
 & \longleftarrow \boldsymbol{u}_{\alpha+\ell} \longrightarrow & & \\
\longleftarrow\!\!\!\!\!-\!\!\!\!-\!\!\!\!- & \boldsymbol{u}_{\alpha+\ell+1} & & -\!\!\!\!-\!\!\!\!\longrightarrow \\
 & & & \longleftarrow \boldsymbol{u}_{\alpha+\ell+2} \longrightarrow \\
 & & & \vdots \\
 & & & \longleftarrow \boldsymbol{u}_m \longrightarrow
\end{pmatrix} \\
\underbrace{\hphantom{xxxxxxxx}}_{c} \quad \underbrace{\hphantom{xxxxxxxx}}_{2\ell} \quad \underbrace{\hphantom{xxxxxxxx}}_{c}
\end{array}.$$

We shall now consider two cases, corresponding to whether or not $\boldsymbol{u}_{\alpha+1}$ is contained in E^{op}_{N-c}:

Case 1: $\boldsymbol{u}_{\alpha+1} \in E^{\mathrm{op}}_{N-c}$. The matrix with rowspace Σ then has the following shape:

$$\begin{array}{c}
\begin{pmatrix}
\longleftarrow \boldsymbol{u}_1 \longrightarrow & & & \\
\vdots & & & \\
\longleftarrow \boldsymbol{u}_\alpha \longrightarrow & & & \\
 & \longleftarrow \boldsymbol{u}_{\alpha+1} \longrightarrow & & \\
 & \longleftarrow \boldsymbol{u}_{\alpha+2} \longrightarrow & & \\
 & \vdots & & \\
 & \longleftarrow \boldsymbol{u}_{\alpha+\ell} \longrightarrow & & \\
\longleftarrow\!\!\!\!\!-\!\!\!\!-\!\!\!\!- & \boldsymbol{u}_{\alpha+\ell+1} & & -\!\!\!\!-\!\!\!\!\longrightarrow \\
 & & & \longleftarrow \boldsymbol{u}_{\alpha+\ell+2} \longrightarrow \\
 & & & \vdots \\
 & & & \longleftarrow \boldsymbol{u}_m \longrightarrow
\end{pmatrix} \\
\underbrace{\hphantom{xxxxxxxx}}_{c} \quad \underbrace{\hphantom{xxxxxxxx}}_{2\ell} \quad \underbrace{\hphantom{xxxxxxxx}}_{c}
\end{array}.$$

Note that for $1 \leq \beta \leq \alpha + \ell$, we have $u_{\beta,j} = 0$ for any $N + 1 - c \leq j \leq N$. Thus for $1 \leq \beta \leq \alpha + \ell$ we have

$$u_{\alpha+\ell+1,1} \cdot u_{\beta,N} + \cdots + u_{\alpha+\ell+1,c} \cdot u_{\beta,N+1-c} = 0,$$

where $u_{i,j}$ is the j-th coordinate of $\boldsymbol{u}_i$.

Because Σ is isotropic, we have $\omega(\boldsymbol{u}_{\alpha+\ell+1}, \boldsymbol{u}_\beta) = 0$ for all $1 \le \beta \le m$. In particular, for $\alpha + \ell + 2 \le \beta \le m$, we then have

$$u_{\alpha+\ell+1,1} \cdot u_{\beta,N} + \cdots + u_{\alpha+\ell+1,c} \cdot u_{\beta,N+1-c} = 0.$$

Finally, let $\boldsymbol{v}$ be the orthogonal projection of $\boldsymbol{u}_{\alpha+\ell+1}$ onto $E^{(c)}$. The span of $\boldsymbol{u}_{\alpha+1}, \ldots, \boldsymbol{u}_{\alpha+\ell+1}$ is a *maximal* isotropic subspace of $E^{(c)}$, so $\boldsymbol{v}$ must be contained in that span. In particular, $\boldsymbol{v}$ is itself an isotropic vector. Thus

$$u_{\alpha+\ell+1,1} \cdot u_{\alpha+\ell+1,N} + \cdots + u_{\alpha+\ell+1,c} \cdot u_{\alpha+\ell+1,N+1-c} = 0.$$

It follows that $f_c(\boldsymbol{w}) = 0$ for any vector $\boldsymbol{w}$ in Σ.

Case 2: $\boldsymbol{u}_{\alpha+1} \notin E^{\mathrm{op}}_{N-c}$. Since $\dim(\Sigma \cap E^{\mathrm{op}}_{N-c}) = m - (\alpha + 1)$, we may assume $\boldsymbol{u}_{\alpha+\ell+1} \in E^{\mathrm{op}}_{N-c}$, after possibly adding a linear combination of $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{\alpha+1}$. Hence there exists a matrix with rowspace Σ of the shape

$$\left(\begin{array}{ccc}
\longleftarrow \boldsymbol{u}_1 \longrightarrow & & \\
\vdots & & \\
\longleftarrow \boldsymbol{u}_\alpha \longrightarrow & & \\
\multicolumn{2}{c}{\longleftarrow\!\!-\!\!-\!\!-\!\!-\; \boldsymbol{u}_{\alpha+1} \;-\!\!-\!\!-\!\!-\!\!\longrightarrow} & \\
 & \longleftarrow \boldsymbol{u}_{\alpha+2} \longrightarrow & \\
 & \vdots & \\
 & \longleftarrow \boldsymbol{u}_{\alpha+\ell} \longrightarrow & \\
 & \multicolumn{2}{c}{\longleftarrow\!\!-\!\!-\!\!-\!\!-\; \boldsymbol{u}_{\alpha+\ell+1} \;-\!\!-\!\!-\!\!-\!\!\longrightarrow} \\
 & & \longleftarrow \boldsymbol{u}_{\alpha+\ell+2} \longrightarrow \\
 & & \vdots \\
 & & \longleftarrow \boldsymbol{u}_m \longrightarrow
\end{array}\right).$$
$$\underbrace{\hphantom{\longleftarrow \boldsymbol{u}_\alpha \longrightarrow}}_{c} \quad \underbrace{\hphantom{\longleftarrow \boldsymbol{u}_{\alpha+2} \longrightarrow}}_{2\ell} \quad \underbrace{\hphantom{\longleftarrow \boldsymbol{u}_{\alpha+\ell+2} \longrightarrow}}_{c}$$

Let $\rho : \mathbb{C}^N \to \mathbb{C}^N$ be the orthogonal projection on to $E^{(c)}$. Notice that $\rho(\boldsymbol{u}_i) = \boldsymbol{0}$ for $i \le \alpha$ and $i \ge \alpha + \ell + 2$. Define $\boldsymbol{v}_i = \rho(\boldsymbol{u}_{\alpha+i})$ for $i \in [1, \ell + 1]$. Note that $\boldsymbol{v}_i = \boldsymbol{u}_{\alpha+i}$ for $i \in [2, \ell]$, but

$$\boldsymbol{v}_1 = (0, \ldots, 0, u_{\alpha+1,c+1}, \ldots, u_{\alpha+1,N-c}, 0, \ldots, 0),$$
$$\boldsymbol{v}_{\ell+1} = (0, \ldots, 0, u_{\alpha+\ell+1,c+1}, \ldots, u_{\alpha+\ell+1,N-c}, 0, \ldots, 0),$$

where $u_{i,j}$ is the j-th coordinate of $\boldsymbol{u}_i$.

Let

$$\Sigma' = \langle \boldsymbol{v}_2, \ldots, \boldsymbol{v}_\ell \rangle,$$
$$\Sigma_1 = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell \rangle,$$
$$\Sigma_2 = \langle \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{\ell+1} \rangle.$$

Both Σ_1 and Σ_2 are elements of $\mathrm{OG}(\ell, E^{(c)})$, and both contain Σ'. It is a well-known fact that there are exactly two isotropic ℓ-planes containing a given isotropic $(\ell-1)$-plane [Fulton and Harris 1991, §23.3]. In particular, there are exactly two elements of $\mathrm{OG}(\ell, E^{(c)})$ that contain Σ', so it remains to verify that Σ_1 and Σ_2 are indeed the same element. It is here that the types of P and T become relevant.

Namely, let

$$P^{(c)} := \{p - c : p \in P \cap [c+1, N-c]\},$$
$$T^{(c)} := \{t - c : t \in T \cap [c+1, N-c]\}$$

be Schubert symbols for $\mathrm{OG}(\ell, E^{(c)}) \cong \mathrm{OG}(\ell, 2\ell)$. Note that whenever a subspace Λ is contained in an intersection of Schubert cells X_R° and $X_{S^\vee}^\circ$ in $\mathrm{OG}(\ell, 2\ell)$, it must be the case that $S \preceq R$ and hence that $\mathfrak{t}(\Lambda) = \mathfrak{t}(R) = \mathfrak{t}(S)$. In particular, $\Sigma_1 \in X_{P^{(c)}}^\circ \subset \mathrm{OG}(\ell, 2\ell)$ by the definition of $X_{P^{(c)}}^\circ$, so it follows that $\mathfrak{t}(\Sigma_1) = \mathfrak{t}(P^{(c)})$. Similarly, $\Sigma_2 \in X_{T^{(c)},\vee}^\circ \subset \mathrm{OG}(\ell, 2\ell)$ by Observation 3.4, and so $\mathfrak{t}(\Sigma_2) = \mathfrak{t}(T^{(c)})$.

We claim that $\mathfrak{t}(P) = \mathfrak{t}(T)$ if and only if $\mathfrak{t}(P^{(c)}) \neq \mathfrak{t}(T^{(c)})$. To see why, note that

$$\#([1, n+1] \setminus P) = \#([1, \ell] \setminus P^{(c)}) + (c - \alpha),$$
$$\#([1, n+1] \setminus T) = \#([1, \ell] \setminus T^{(c)}) + (c - (\alpha + 1)).$$

Thus $\mathfrak{t}(P) + \mathfrak{t}(T) \equiv \mathfrak{t}(P^{(c)}) + \mathfrak{t}(T^{(c)}) + 1 \pmod 2$.

Since we know that $\mathfrak{t}(P) \neq \mathfrak{t}(T)$, we can conclude that $\Sigma_1 = \Sigma_2$. It follows that $\boldsymbol{v}_{\ell+1} \in \Sigma_1$. But Σ_1 is isotropic, and therefore $\omega(\boldsymbol{v}_1, \boldsymbol{v}_{\ell+1}) = 0$. Thus, for all $\boldsymbol{w} \in \Sigma$, the polynomial $x_{c+1}x_{N-c} + \cdots + x_{n+1}x_{n+2}$ vanishes, and hence $f_c(\boldsymbol{w}) = 0$. □

Having addressed exceptional cuts, we can now prove a generalization of Lemma 2.8 for a Grassmannian X of type B, C, or D:

Proposition 5.2. *Given Schubert symbols $T \preceq P$ in $\Omega(X)$, where X is a Grassmannian of type B, C, or D, the projected Richardson variety $\psi(\pi^{-1}(Y_{P,T}))$ is contained in $Z_{P,T}$.*

Proof. Fix $\Sigma \in Y_{P,T}$ and $\boldsymbol{w} = (w_1, \dots, w_N) \in \Sigma$. Suppose $c \in \mathfrak{Q}_{P,T}$. We will first show that $f_c(\boldsymbol{w}) = 0$.

If c (or $N-c$) is a visible cut, then there exists $j \in [1, m]$ such that $p_j \le c < t_{j+1}$ (or $p_j \le N - c < t_{j+1}$).

Let $W_1 = \langle \boldsymbol{e}_1, \dots, \boldsymbol{e}_{p_j} \rangle$ and $W_2 = \langle \boldsymbol{e}_{p_j+1}, \dots, \boldsymbol{e}_N \rangle$. Since $\mathbb{C}^N = W_1 \oplus W_2$, there exists a unique decomposition $\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2$ with $\boldsymbol{w}_i \in W_i$. Namely, we have $\boldsymbol{w}_1 = (w_1, \dots, w_{p_j}, 0, \dots, 0)$ and $\boldsymbol{w}_2 = (0, \dots, 0, w_{p_j+1}, \dots, w_N)$. Note that, since $w_i = 0$ for $p_j < i < t_{j+1}$, we have $\boldsymbol{w}_2 = (0, \dots, 0, w_{t_{j+1}}, \dots, w_N)$, and hence $f_c(\boldsymbol{w}) = (\boldsymbol{w}_1, \boldsymbol{w}_2)$.

Since $\Sigma \in Y_{P,T}$, $\dim(\Sigma \cap W_1) \geq j$ and $\dim(\Sigma \cap W_2) \geq m-j$. We can therefore write $\Sigma = (\Sigma \cap W_1) \oplus (\Sigma \cap W_2)$, and decompose $\boldsymbol{w}$ as the sum of vectors in these subspaces. Since there is only one such decomposition, we must have $\boldsymbol{w}_1 \in \Sigma \cap W_1$ and $\boldsymbol{w}_2 \in \Sigma \cap W_2$. Since Σ isotropic, $f_c(\boldsymbol{w}) = (\boldsymbol{w}_1, \boldsymbol{w}_2) = 0$.

If $c = n+1$ (implying we are working in type B or D), then $f_c = f_{n+1}$, the inherent quadratic equation, which vanishes on all isotropic vectors.

Lastly, if c is exceptional, then $f_c(\boldsymbol{w}) = 0$ by Proposition 5.1.

We have shown that the projected Richardson variety $\psi(\pi^{-1}(Y_{P,T}))$ lies in the zero set of the polynomial f_c for any element $c \in \mathscr{Q}_{P,T}$. In particular, it satisfies all the quadratic equations defining $Z_{P,T}$.

We must now show that $\psi(\pi^{-1}(Y_{P,T}))$ satisfies the linear equation $x_c = 0$ for any $c \in \mathscr{L}_{P,T}$. Suppose $c \in [1, N]$ is a zero column of $D(P,T)$, and let Σ continue to denote an arbitrary element of $Y_{P,T}$. Let $j = \max\{i \in [1,m] : p_i < c\}$, and note that $t_{j+1} > c$. This time let $W_1 = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{c-1} \rangle$ and $W_2 = \langle \boldsymbol{e}_{c+1}, \ldots, \boldsymbol{e}_N \rangle$. Thus, $\dim(\Sigma \cap W_1) \geq j$ and $\dim(\Sigma \cap W_2) \geq m-j$, so $\Sigma \subset W_1 \oplus W_2$, which is the hyperplane defined by $x_c = 0$.

Finally, suppose column $d \in [1, N]$ contains a lone star. In other words, $d = p_j$ or $d = t_j$ for some j, and $d-1$ and d are both cuts in $D(P,T)$. Let $c = \min(d, N+1-d)$, which is in $[1, \lceil N/2 \rceil]$. Both $c-1$ and c are also cuts in $D(P,T)$, so the set $\psi(\pi^{-1}(Y_{P,T}))$ must lie in the zero set of the polynomial $f_c - f_{c-1} = x_c x_{N+1-c} = x_d x_{N+1-d}$. However, since $\psi(\pi^{-1}(Y_{P,T}))$ is irreducible (by [Richardson 1992]), it must lie in $x_d = 0$ or $x_{N+1-d} = 0$ (or both).

Consider any subspace Λ in the dense subset $X_P^\circ(E_\bullet) \cap X_{T^\vee}^\circ(E_\bullet^{\mathrm{op}}) \subset Y_{P,T}$. Since $d \in P \cup T$, it is impossible that $x_d = 0$ on all vectors in Λ. Therefore, the equation $x_{N+1-d} = 0$ must be satisfied by $\psi(\pi^{-1}(Y_{P,T}))$. □

6. Result 3: $Z_{P,T} \subset \psi(\pi^{-1}(Y_{P,T}))$

Let $X := \mathrm{IG}_\omega(m, \mathbb{C}^N)$ be a Grassmannian of type B, C, or D. Given Schubert symbols P and T in $\Omega(X)$, we write $P \to T$ whenever

(i) $T \preceq P$;

(ii) $p_i \leq t_{i+1}$ for all i, unless $p_i = n+2$ and $t_{i+1} = n+1$ in type D; and

(iii) if $p_i = t_{i+1}$, then p_i is *not* a cut in $D(P,T)$.

When $p_i > t_{i+1}$, we say $D(P,T)$ has a 2×2 *square*. Thus, the second condition says that $D(P,T)$ has no 2×2 squares, except that a single 2×2 square in the central columns is permitted in type D. When $p_i = t_{i+1}$ for some i and $p_i \in \mathscr{C}_{P,T}$, then $(i+1, t_{i+1})$ is a lone star in $D(P,T)$ (in fact, so is (i, p_i)), and hence $N+1-p_i \in \mathscr{L}_{P,T}$. If we are working in type B or C, then $N+1-p_i$ must be a zero column, which yields the equivalent definition of $P \to T$ given in [Buch et al. 2009].

If we are working in type D, then there are other possibilities involving exceptional cuts. For example, if $p_i = t_{i+1} \in \{n+1, n+2\}$, then $n+1$ will be an exceptional cut, causing both (i, p_i) and $(i+1, t_{i+1})$ to be lone stars. It follows that, when $P \to T$, the diagram $D(P, T)$ cannot have exactly three stars in the central columns $n+1$ and $n+2$, which yields the alternative type-D definition of $P \to T$ given in [Buch et al. 2009, §5.2].

The relation $P \to T$ is important because it characterizes precisely when the map $\psi : \pi^{-1}(Y_{P,T}) \to Z_{P,T}$ is a birational isomorphism. In particular, [Buch et al. 2009, Proposition 5.1] says the following:

Proposition 6.1. *Given a Grassmannian X of Lie-type B, C, or D, the map $\psi : \pi^{-1}(Y_{P,T}) \to Z_{P,T}$ is a birational isomorphism if and only if the Schubert symbols P and T satisfy the relation $P \to T$.*

A more detailed proof of Proposition 6.1 can be found in [Ravikumar 2013, §8]. In this section, we prove the following proposition:

Proposition 6.2. *Given Schubert symbols $T \preceq P$ for a Grassmannian X of type B, C, or D, there exists a Schubert symbol $\widetilde{P}$ such that*

(1) $T \preceq \widetilde{P} \preceq P$,

(2) $Z_{\widetilde{P},T} = Z_{P,T}$, *and*

(3) $\widetilde{P} \to T$.

We prove Proposition 6.2 by explicitly constructing the Schubert symbol $\widetilde{P}$. An immediate consequence is that, for any $T \preceq P$ in $\Omega(X)$, the Richardson variety $Y_{P,T}$ must contain a smaller Richardson variety $Y_{\widetilde{P},T}$ such that $\psi : \pi^{-1}(Y_{\widetilde{P},T}) \to Z_{\widetilde{P},T} = Z_{P,T}$ is a birational isomorphism (by Proposition 6.1). It follows that $Z_{P,T} \subset \psi(\pi^{-1}(Y_{P,T}))$. By combining this observation with Proposition 5.2, which states that $Z_{P,T} \supset \psi(\pi^{-1}(Y_{P,T}))$, we have a proof of Theorem 1.1. We mention that Proposition 6.2 is not the only way to prove that $Z_{P,T} \subset \psi(\pi^{-1}(Y_{P,T}))$. For example, instead of "lowering" P we could "raise" T. In fact, the construction we give for $\widetilde{P}$ does precisely that when carried out on the rotated diagram $D(\overline{T}, \overline{P})$.

6A. *Constructing a "smaller" Schubert symbol P'.* Given Schubert symbols $T \preceq P$, define the set $P' = \{p'_1, \dots, p'_m\}$ as follows:

If $p_i < t_{i+1}$, then

$$p'_i = p_i. \tag{$\diamond$}$$

On the other hand, if $p_i \geq t_{i+1}$ and $t_{i+1} - 1 \notin \mathscr{C}_{P,T}$, then

$$p'_i = t_{i+1}. \tag{$\clubsuit$}$$

Finally, if $p_i \geq t_{i+1}$ *and* $t_{i+1} - 1 \in \mathscr{C}_{P,T}$, then

$$p_i' = \max\{c \in [t_i, t_{i+1} - 1] \mid c \notin \mathscr{L}_{P,T}\}. \tag{$\heartsuit$}$$

Note that since t_i cannot be in $\mathscr{L}_{P,T}$, the set $\{c \in [t_i, t_{i+1} - 1] \mid c \notin \mathscr{L}_{P,T}\}$ is nonempty. Thus p_i' is well-defined.

The next property of P' follows from its construction and Corollary 4.13:

Observation 6.3. *For any* $1 \leq i \leq m$, $p_i' \notin \mathscr{L}_{P,T}$.

Proof. If p_i' is defined by case ($\diamondsuit$) or case ($\clubsuit$), then $p_i' \in P \cup T$, so by Corollary 4.13 it is not in $\mathscr{L}_{P,T}$. If it is defined by case ($\heartsuit$), then by its construction it cannot be in $\mathscr{L}_{P,T}$. □

We also have the next observation, which follows directly from the construction of P':

Observation 6.4. *If* p_i' *is produced by case* ($\diamondsuit$) *or case* ($\heartsuit$) *of the previous construction, then* $[p_i', t_{i+1} - 1] \subset \mathscr{C}_{P,T}$. *Furthermore, if* $p_i' < t_{i+1} - 1$, *then* $[p_i' + 1, t_{i+1} - 1] \subset \mathscr{L}_{P,T}$.

Lemma 6.5. *P' is a Schubert symbol.*

Proof. We will first show that $p_i' < p_{i+1}'$ for $1 \leq i \leq m-1$. By our construction of P', we have $p_i' \leq t_{i+1} \leq p_{i+1}'$ for each i. Thus we only need to consider the case that $p_{i+1}' = t_{i+1}$. If $p_{i+1}' = t_{i+1}$, then p_{i+1}' follows case ($\diamondsuit$) or case ($\heartsuit$), so, by Observation 6.4, $t_{i+1} \in \mathscr{C}_{P,T}$. But then $(i+1, t_{i+1})$ is a lone star in $D(P, T)$, so $t_{i+1} - 1 \in \mathscr{C}_{P,T}$ as well (by Corollary 2.7). Thus p_i' follows case ($\diamondsuit$) or case ($\heartsuit$), and we have $p_i' < t_{i+1}$. It follows that $p_i' < p_{i+1}'$ for all $1 \leq i \leq m-1$, and in particular that P' consists of m distinct integers.

We still have to check that P' satisfies the isotropic condition. Suppose on the contrary that $p_i' + p_j' = N + 1$ for some i and j in $[1, m]$. By Observation 6.4, if p_i' is not a cut in $D(P, T)$, then $p_i' = t_{i+1}$. Similarly, if p_j' is not a cut in $D(P, T)$, then $p_j' = t_{j+1}$.

We therefore have three possible cases to consider, each of which results in a contradiction:

Case 1: *Both* p_i' *and* p_j' *are cuts in* $D(P, T)$. In this case $N - p_j' = p_i' - 1$ is also a cut, and by Lemma 4.14 either p_i' or p_j' is in $\mathscr{L}_{P,T}$. But neither p_i' nor p_j' can be in $\mathscr{L}_{P,T}$, by Observation 6.3.

Case 2: *Exactly one of* $\{p_i', p_j'\}$ *is a cut in* $D(P, T)$. We will assume without loss of generality that p_j' is a cut in $D(P, T)$. We then have $p_i' - 1 \in \mathscr{C}_{P,T}$, as in the previous case, implying that p_i' cannot follow case ($\clubsuit$). However, since $p_i' \notin \mathscr{C}_{P,T}$, Observation 6.4 implies that p_i' *does* follow case ($\clubsuit$), a contradiction.

Case 3: *Neither p'_i nor p'_j is a cut in $D(P,T)$*. In this case, $p'_i = t_{i+1}$ and $p'_j = t_{j+1}$. But $t_{i+1} + t_{j+1} \neq N+1$, since T is a Schubert symbol, so once again we arrive at a contradiction. □

6B. ***Types B and C*.** If X is a Grassmannian of type B or C, we set $\widetilde{P} := P'$ in order to prove Proposition 6.2. The type-D case of Proposition 6.2 will be addressed in Section 6C.

Since $T \preceq P$ if and only if $T \leq P$, it is clear by our construction that $t_i \leq \tilde{p}_i \leq p_i$ for $1 \leq i \leq m$, and hence $T \preceq \widetilde{P} \preceq P$. Thus $\widetilde{P}$ satisfies Proposition 6.2(1). We will need the following lemma to verify the remaining conditions:

Lemma 6.6. *The diagrams $D(P,T)$ and $D(\widetilde{P},T)$ have the same cuts.*

Proof. Given an integer $c \in [1,N]$, if $p_i \leq c < t_{i+1}$, then $\tilde{p}_i = p_i \leq c < t_{i+1}$. Therefore $\mathscr{C}_{P,T} \subset \mathscr{C}_{\widetilde{P},T}$. On the other hand, if $\tilde{p}_i < t_{i+1}$, then $\tilde{p}_i$ is defined by ($\diamondsuit$) or ($\heartsuit$). We then have $[\tilde{p}_i, t_{i+1}-1] \subset \mathscr{C}_{P,T}$, by Observation 6.4. Therefore $\mathscr{C}_{P,T} \supset \mathscr{C}_{\widetilde{P},T}$. □

Proof of Proposition 6.2. We show that $\widetilde{P}$ satisfies conditions (2) and (3).

$\widetilde{P}$ satisfies (2) ($Z_{P,T} = Z_{\widetilde{P},T}$): By Lemma 6.6, the diagrams $D(P,T)$ and $D(\widetilde{P},T)$ have the same cuts. We still have to prove that $\mathscr{L}_{P,T} = \mathscr{L}_{\widetilde{P},T}$. Suppose $c \in \mathscr{L}_{P,T}$. Then, by Proposition 2.6, either $p_i = t_i = N+1-c$ for some i or c is a zero column of $D(P,T)$. If $p_i = t_i = N+1-c$ for some i, then $p_i < t_{i+1}$, so $\tilde{p}_i$ must follow case ($\diamondsuit$) of the construction. Thus $\tilde{p}_i = t_i = N+1-c$, so $c \in \mathscr{L}_{\widetilde{P},T}$. On the other hand, if c is a zero column in $D(P,T)$, then $p_i < c < t_{i+1}$ for some i. Once again we have $\tilde{p}_i = p_i$, so c is a zero column in $D(\widetilde{P},T)$, and therefore an element of $\mathscr{L}_{\widetilde{P},T}$.

Now suppose $c \in \mathscr{L}_{\widetilde{P},T}$. If $\tilde{p}_i < c < t_{i+1}$ for some i, then $\tilde{p}_i$ satisfies ($\diamondsuit$) or ($\heartsuit$). Either way, we must have $c \in \mathscr{L}_{P,T}$. On the other hand, if $(i, N+1-c)$ is a lone star in $D(\widetilde{P},T)$ for some i, then $\tilde{p}_i = t_i = N+1-c$. Thus, $\tilde{p}_i \in \mathscr{C}_{\widetilde{P},T} = \mathscr{C}_{P,T}$. In this case (i, t_i) is a lone star in $D(P,T)$, and $c \in \mathscr{L}_{P,T}$.

$\widetilde{P}$ satisfies (3) ($\widetilde{P} \to T$): Since $\tilde{p}_i \leq t_{i+1}$ for $1 \leq i \leq m-1$, the diagram $D(\widetilde{P},T)$ has no 2×2 squares. If $\tilde{p}_i = t_{i+1}$ for some i, then $t_{i+1}-1$ is not a cut in $D(P,T)$. By Lemma 6.6 it is also not a cut in $D(\widetilde{P},T)$, so column $N+1-\tilde{p}_i$ cannot be a zero column in $D(\widetilde{P},T)$. □

6C. ***Type D*.** Let $X := \mathrm{OG}(m, 2n+2)$ be a type-D Grassmannian, let $N := 2n+2$, and let $T \preceq P$ be Schubert symbols in $\Omega(X)$. Unfortunately the Schubert symbol P' constructed in Section 6A fails to satisfy the conditions of Proposition 6.2, as the following example illustrates:

Example 6.7. Consider $\mathrm{OG}(2,6)$, and let $T=\{1,3\}$ and $P=\{5,6\}$. The projected Richardson variety $Z_{P,T}$ is the quadric hypersurface of $\mathbb{P}^5$ consisting of all isotropic lines in $\mathbb{C}^6$. Since $P'=\{3,6\}$ and the variety $Z_{P',T}$ satisfies the additional linear equation $x_4=0$, we have $Z_{P,T}\neq Z_{P',T}$. Furthermore, $P'\not\to T$, since $p_1=t_2=3$ is a cut in $D(P',T)$ — the exceptional center cut

$$\begin{pmatrix} * & * & * & 0 & 0 & 0 \\ 0 & 0 & * & * & * & * \end{pmatrix}.$$

We therefore define the set $\widetilde{P}=\{\tilde{p}_1,\ldots,\tilde{p}_m\}$ as follows:

If $p_i<t_{i+1}$, then

$$\tilde{p}_i=p_i. \tag{$\star$}$$

On the other hand, if $p_i\geq t_{i+1}$ and $t_{i+1}-1\notin\mathscr{C}$, then

$$\tilde{p}_i=\begin{cases} t_{i+1} & \text{if } t_{i+1}\notin\{n+1,n+2\}, \\ N+1-t_{i+1} & \text{if } t_{i+1}\in\{n+1,n+2\}. \end{cases} \tag{$\dagger$}$$

Finally, if $p_i\geq t_{i+1}$ *and* $t_{i+1}-1\in\mathscr{C}$, then

$$\tilde{p}_i=\max\{c\in[t_i,t_{i+1}-1]\mid c\notin\mathscr{L}\}. \tag{$\ddagger$}$$

Note that since t_i cannot be in $\mathscr{L}$, the set $\{c\in[t_i,t_{i+1}-1]\mid c\notin\mathscr{L}\}$ is nonempty. Thus $\tilde{p}_i$ is well-defined.

Recall that ι is the permutation of $\{1,\ldots,2n+2\}$ that interchanges $n+1$ and $n+2$ and leaves all other numbers fixed. We make the following observation:

Observation 6.8. *$\widetilde{P}$ is equal to $\iota P'$ or P'. Moreover, $\widetilde{P}=\iota P'$ if and only if there exists an element $\tilde{p}_i$ defined by case ($\dagger$) and $t_{i+1}\in\{n+1,n+2\}$.*

Given a type-D Schubert symbol R, the set ιR is also a Schubert symbol, since the isotropic condition is preserved. We therefore have the following corollary:

Corollary 6.9. *$\widetilde{P}$ is a Schubert symbol.*

The following observations are exact restatements of Observations 6.3 and 6.4 for type D. We give a brief proof of the first, whereas the second follows directly from the construction of $\widetilde{P}$.

Observation 6.10. *For any $1\leq i\leq m$, $\tilde{p}_i\notin\mathscr{L}_{P,T}$.*

Proof. If $\tilde{p}_i$ is defined by case ($\star$), or if it is defined by case ($\dagger$) and $t_{i+1}\notin\{n+1,n+2\}$, then $\tilde{p}_i\in P\cup T$. By Corollary 4.13, it is not in $\mathscr{L}_{P,T}$. If it is defined by case ($\ddagger$), then by its construction it cannot be in $\mathscr{L}_{P,T}$.

Finally, suppose $\tilde{p}_i$ is defined by case ($\dagger$) and $t_{i+1}\in\{n+1,n+2\}$. If $\tilde{p}_i=N+1-t_{i+1}\in\mathscr{L}_{P,T}$, then $t_{i+1}-1\in\mathscr{C}_{P,T}$, contradicting the assumption that $\tilde{p}_i$ is defined by case ($\dagger$). □

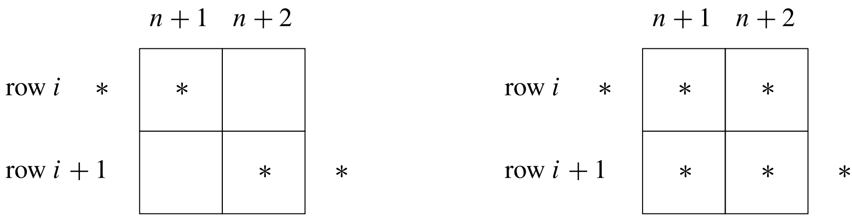


Figure 2. Center columns of $D(\widetilde{P}, T)$ in Lemma 6.12.

Observation 6.11. *If $\tilde{p}_i$ is produced by case ($\star$) or case ($\ddagger$) of the previous construction, then $[\tilde{p}_i, t_{i+1} - 1] \subset \mathscr{C}_{P,T}$. Furthermore, if $\tilde{p}_i < t_{i+1} - 1$, then $[\tilde{p}_i + 1, t_{i+1} - 1] \subset \mathscr{L}_{P,T}$.*

We will now proceed to prove that $T \preceq P$, $\widetilde{P} \to T$, $Z_{P,T} = Z_{\widetilde{P},T}$, and $\widetilde{P} \preceq P$, in that order. We begin with an important lemma:

Lemma 6.12. *If $n + 1 \in [T] \cap [\widetilde{P}]$, then $\mathfrak{t}(T) = \mathfrak{t}(\widetilde{P})$.*

Proof. Suppose $\tilde{p}_i \in \{n + 1, n + 2\}$ for some i and $t_j \in \{n + 1, n + 2\}$ for some j.

Since $\tilde{p}_k \geq t_k > n+2$ for all $k > j$, we have $i \leq j$. Furthermore, $\tilde{p}_{k-1} \leq t_k < n+1$ for all $k < j$, so $i \geq j - 1$. Thus we either have $i = j$ or $i + 1 = j$.

Case 1: $i = j$. We claim that $t_i = \tilde{p}_i \in \{n+1, n+2\}$. Suppose on the contrary that $t_i = n + 1$ and $\tilde{p}_i = n + 2$. Since $t_{i+1} \neq n + 1$, $\tilde{p}_i$ is *not* defined by ($\dagger$). Therefore, by Observation 6.11, $\tilde{p}_i \in \mathscr{C}_{P,T}$. It follows that both n and $n + 2$ are in $\mathscr{Q}_{P,T}$, so, by Lemma 4.14, either $n + 1$ or $n + 2$ is in $\mathscr{L}_{P,T}$. But neither t_i nor $\tilde{p}_i$ can be in $\mathscr{L}_{P,T}$, by Corollary 4.13 and Observation 6.10. To avoid this contradiction, we must have $t_i = \tilde{p}_i \in \{n + 1, n + 2\}$. Thus, $n + 1$ is a visible cut in $D(\widetilde{P}, T)$, and $\#([1, n + 1] \cap T) = \#([1, n + 1] \cap \widetilde{P})$. In other words, $\mathfrak{t}(T) = \mathfrak{t}(\widetilde{P})$.

Case 2: $i + 1 = j$. If $\tilde{p}_i = t_{i+1}$, then $\tilde{p}_i$ is defined by ($\dagger$) and $t_{i+1} \notin \{n + 1, n + 2\}$. Hence $\tilde{p}_i \notin \{n + 1, n + 2\}$, a contradiction. It follows that $\tilde{p}_i \neq t_{i+1}$. If $\tilde{p}_i = n + 1$ and $t_{i+1} = n + 2$, then $n + 1$ is a visible cut in $D(\widetilde{P}, T)$, as seen on the left side of Figure 2. On the other hand, if $\tilde{p}_i = n + 2$ and $t_{i+1} = n + 1$, as seen on the right side, then $\#([1, n + 1] \cap T) = \#([1, n + 1] \cap \widetilde{P}) + 2$, so once again $\mathfrak{t}(T) = \mathfrak{t}(\widetilde{P})$. □

Lemma 6.12 says that $\{\mathfrak{t}(\widetilde{P}), \mathfrak{t}(T)\} \neq \{0, 1\}$. Since we also know that $T \leq \widetilde{P}$ by construction, we have the following immediate corollaries:

Corollary 6.13. $T \preceq \widetilde{P}$.

Proof. The condition that $\{\mathfrak{t}(\widetilde{P}), \mathfrak{t}(T)\} = \{0, 1\}$ is necessary for $T \not\preceq \widetilde{P}$. □

Corollary 6.13 says that $\widetilde{P}$ satisfies the first half of condition (1) of Proposition 6.2. We prove the second half in Proposition 6.21.

Corollary 6.14. *The diagram $D(\widetilde{P}, T)$ has no exceptional cuts, except possibly the center cut $n + 1$.*

Proof. The condition that $\{\mathfrak{t}(\widetilde{P}), \mathfrak{t}(T)\} = \{0, 1\}$ is necessary for the existence of an exceptional cut other than the center cut. $\square$

In fact, any noncentral exceptional cut in $D(P, T)$ becomes a visible cut in $D(\widetilde{P}, T)$. However, we will only need this fact in the special case that the exceptional cut is less than $n + 1$:

Lemma 6.15. *If $1 \le c \le n$ is an exceptional cut in $D(P, T)$, then c is a visible cut in $D(\widetilde{P}, T)$.*

Proof. Since c is an exceptional cut in $D(P, T)$, we have that $\#([1, c] \cap T) = \#([1, c] \cap P) + 1$. In other words, there exists exactly one integer $i \in [1, m]$ such that $t_i \le c < p_i$.

Since row i is the only row crossing from column c to column $c + 1$, we must have $c + 1 \le t_{i+1}$. Furthermore, since $n + 1 \in [T]$, we must have $t_{i+1} \le n + 2$. We will show that $\tilde{p}_i \le c$, and hence that c is a visible cut in $D(\widetilde{P}, T)$. We divide the rest of our argument into two cases:

Case 1: $c + 1 = t_{i+1}$. Since $c \in \mathscr{C}_{P,T}$, $\tilde{p}_i$ is defined by (‡), it follows that $\tilde{p}_i \le c$.

Case 2: $c + 1 < t_{i+1} \le n + 2$. We claim that $p_i \ge t_{i+1}$. To see why, suppose for the sake of contradiction that $p_i < t_{i+1}$. Then p_i is a visible cut in $D(P, T)$, and $\#([1, p_i] \cap T) = \#([1, p_i] \cap P)$. If $p_i < n+1$, then since $[p_i + 1, n + 1] \subset [P] \cap [T]$ and $\mathfrak{t}(T) \ne \mathfrak{t}(P)$ (which follow from the fact that c is an exceptional cut), we have $T \not\le P$, a contradiction. If $p_i = n + 1$, then $\#([1, n + 1] \cap T) = \#([1, n + 1] \cap P)$, so $\mathfrak{t}(P) = \mathfrak{t}(T)$, contradicting the assumption that c is an exceptional cut. Finally, if $p_i > n + 1$, then $t_{i+1} > n + 2$, contradicting the assumption that $n + 1 \in [T]$. It follows that $p_i \ge t_{i+1}$, as claimed.

Now note that any $d \in [c + 1, t_{i+1} - 1]$ must be an exceptional cut in $D(P, T)$, since $\#([1, d] \cap T) = \#([1, d] \cap P) + 1$ (in particular row i is the only row crossing from column d to column $d + 1$). Furthermore, for each $d \in [c + 1, t_{i+1} - 1]$, we have $N + 1 - d \in T$, since $d \notin T$ and $[c + 1, n + 1] \in [T]$. It follows that column $N + 1 - d$ contains a lone star in $D(P, T)$, and hence that $d \in \mathscr{L}_{P,T}$. Since $[c + 1, t_{i+1} - 1] \subset \mathscr{L}_{P,T}$, it follows that $\tilde{p}_i$ is defined by (‡) and that $\tilde{p}_i \le c$. $\square$

The following proposition will also be needed:

Proposition 6.16. $\mathscr{C}_{P,T} \cup \{n + 1\} = \mathscr{C}_{\widetilde{P},T} \cup \{n + 1\}$.

Proof. By Corollary 6.14, there are no exceptional cuts in $D(\widetilde{P}, T)$. If c is a visible cut in $D(\widetilde{P}, T)$, then $\tilde{p}_i \le c < t_{i+1}$ for some $1 \le i \le m - 1$. If $\tilde{p}_i$ is defined, by ($\star$) or (‡), then, by Observation 6.11, $c \in \mathscr{C}_{P,T}$. Otherwise $\tilde{p}_i$ is defined by (†) and $\tilde{p}_i = c = n + 1$. Since all visible cuts in $\mathscr{C}_{\widetilde{P},T}$ are contained in $\mathscr{C}_{P,T} \cup \{n + 1\}$, the same is true for apparent cuts. It follows that $\mathscr{C}_{\widetilde{P},T} \subset \mathscr{C}_{P,T} \cup \{n + 1\}$.

On the other hand, if $p_i < t_{i+1}$ then $\tilde{p}_i = p_i < t_{i+1}$. Thus any visible cut $c \in \mathscr{C}_{P,T}$ is also contained in $\mathscr{C}_{\widetilde{P},T}$. If $c < n+1$ is an exceptional cut in $D(P,T)$, then, by Lemma 6.15, $c \in \mathscr{C}_{\widetilde{P},T}$. It follows that $\mathscr{C}_{P,T} \subset \mathscr{C}_{\widetilde{P},T} \cup \{n+1\}$. ☐

We can now prove that $\widetilde{P}$ satisfies condition (3) of Proposition 6.2.

Proposition 6.17. $\widetilde{P} \to T$.

Proof. By Corollary 6.13, $T \preceq \widetilde{P}$. The construction of $\widetilde{P}$ ensures that $\tilde{p}_i$ is never greater than t_{i+1} except possibly when $t_{i+1} = n+1$ and $\tilde{p}_i = n+2$. Furthermore, if $\tilde{p}_i = t_{i+1}$, then $t_{i+1} - 1$ is not a cut in $D(P,T)$. Therefore t_{i+1} cannot be a cut in $D(P,T)$ either, since that would make $(i+1, t_{i+1})$ a lone star in $D(P,T)$ and hence make $t_{i+1} - 1$ a cut in $D(P,T)$, by Proposition 4.10. It follows that t_{i+1} is not a cut in $D(\widetilde{P},T)$, by Proposition 6.16. ☐

We will now show that $Z_{P,T} = Z_{\widetilde{P},T}$ by examining the diagrams $D(P,T)$ and $D(\widetilde{P},T)$. Recall that the quadratic equations defining $Z_{P,T}$ are entirely determined by $\mathscr{C}_{P,T}$, the set of cuts in the $D(P,T)$.

Notice that $\mathscr{C}_{P,T}$ is not equal to $\mathscr{C}_{\widetilde{P},T}$ in general, because our construction of $\widetilde{P}$ adds the center cut $n+1$ whenever $p_i > t_{i+1} = n+2$. By Proposition 6.16, that is the only change in the cut set, and the addition of the center cut does not alter the equations in these cases.

In fact, by Proposition 6.16, $Z_{P,T}$ and $Z_{\widetilde{P},T}$ satisfy the same quadratic equations: namely, $\{f_c \mid c \in (\mathscr{C}_{P,T} \cap [1,n]) \cup \{n+1\}\}$. The following proposition shows that they satisfy the same linear equations as well:

Proposition 6.18. $\mathscr{L}_{P,T} = \mathscr{L}_{\widetilde{P},T}$.

Proof. Suppose $c \in \mathscr{L}_{\widetilde{P},T}$. If c is a zero column in $D(\widetilde{P},T)$, then $\tilde{p}_i < c < t_{i+1}$ for some i. Whether $\tilde{p}_i$ satisfies ($\star$) or (‡), we then have $c \in \mathscr{L}_{P,T}$.

If c is not a zero column in the diagram, then column $N+1-c$ contains a lone star in $D(\widetilde{P},T)$. If neither $N+1-c$ nor $N-c$ are exceptional cuts in $D(\widetilde{P},T)$ then we must have $t_i = \tilde{p}_i = N+1-c$ for some i, by Proposition 2.6. Thus $t_i \in \mathscr{C}_{\widetilde{P},T}$, and hence, by Proposition 6.16, $t_i \in \mathscr{C}_{P,T}$. It follows that (i, t_i) is a lone star in $D(P,T)$ and $c \in \mathscr{L}_{P,T}$. On the other hand, if either $N+1-c$ or $N-c$ is an exceptional cut in $D(P,T)$ then, by Corollary 6.14, $n+1$ must be that exceptional cut. It follows that $c \in \{n+1, n+2\}$. Since $c \in \mathscr{L}_{\widetilde{P},T}$, we have that n and $n+2$ are in $\mathscr{C}_{\widetilde{P},T}$, and hence in $\mathscr{C}_{P,T}$ by Proposition 6.16. By Lemma 4.14, either $n+1$ or $n+2$ (that is, either c or $N+1-c$) must be in $\mathscr{L}_{P,T}$. But column $N+1-c$ contains a lone star in $D(\widetilde{P},T)$, and hence $N+1-c \in T \cup \widetilde{P}$. Thus by Corollary 4.13 and Observation 6.10, it is impossible for $N+1-c$ to be in $\mathscr{L}_{P,T}$. Therefore $c \in \mathscr{L}_{P,T}$. It follows that $\mathscr{L}_{\widetilde{P},T} \subset \mathscr{L}_{P,T}$.

On the other hand, if $c \in \mathscr{L}_{P,T}$, then $c-1$ and c are both cuts in $D(P,T)$. By Proposition 6.16, they are both cuts in $D(\widetilde{P},T)$ as well, and therefore either c or

$N+1-c$ is in $\mathcal{L}_{\widetilde{P},T}$ by Lemma 4.14. If $N+1-c \in \mathcal{L}_{\widetilde{P},T}$, then we have shown that $N+1-c \in \mathcal{L}_{P,T}$ as well. In that case, by Corollary 4.13, $D(P,T)$ does not have a lone star in column $N+1-c$, so c must be a zero column in $D(P,T)$. In other words $p_i < c < t_{i+1}$ for some i. But then $\tilde{p}_i = p_i$, so c is a zero column in $D(\widetilde{P},T)$ as well. □

We have condition (2) of Proposition 6.2 as an immediate corollary:

Corollary 6.19. $Z_{P,T} = Z_{\widetilde{P},T}$.

Finally, we prove that $\widetilde{P} \preceq P$, and hence that the Richardson variety $Y_{\widetilde{P},T}$ is indeed contained in $Y_{P,T}$. Our proof will require the following somewhat technical lemma:

Lemma 6.20. *Suppose $\mathfrak{t}(T) \neq \mathfrak{t}(P)$, and that there exists an integer $1 \le c \le n$ such that $c \notin [T]$ and $[c+1, n+1] \subset [P] \cap [T] \cap [\widetilde{P}]$. Then $c \notin [P] \cap [\widetilde{P}]$.*

Proof. Suppose for the sake of contradiction that $c \in [P] \cap [\widetilde{P}]$. We divide our argument into four cases.

Case 1: *$c \in \widetilde{P}$ and $N+1-c \in P$.* In this case, $\tilde{p}_i = c$ for some $i \in [1,m]$ and $p_j = N+1-c$ for some $j \in [1,m]$.

Since $c \notin T$, $\tilde{p}_i$ is defined by ($\star$) or (‡). Hence $\tilde{p}_i \in \mathcal{C}_{P,T}$ by Observation 6.11. Since $p_j - 1 = N - \tilde{p}_i$ is also a cut in $D(P,T)$, it follows that (j, p_j) is a lone star in $D(P,T)$, and therefore that $\tilde{p}_i \in \mathcal{L}_{P,T}$, contradicting Observation 6.10.

Case 2: *$c \in P$ and $N+1-c \in \widetilde{P}$.* In this case, $p_i = c$ for some $i \in [1,m]$ and $\tilde{p}_j = N+1-c$ for some $j \in [1,m]$.

Since $N+1-c \notin T$, $\tilde{p}_j$ is defined by ($\star$) or (‡). Hence $\tilde{p}_j \in \mathcal{C}_{P,T}$ by Observation 6.11. Since $p_i - 1 = N - \tilde{p}_j$ is also a cut in $D(P,T)$, it follows that (i, p_i) is a lone star in $D(P,T)$, and therefore that $\tilde{p}_j \in \mathcal{L}_{P,T}$, contradicting Observation 6.10.

Case 3: *$c \in \widetilde{P} \cap P$.* In this case, $p_i = c$ for some i.

Let $\ell := n+1-c$. Since $[c+1, n+1] \subset [P]$, Lemma 4.6 tells us that $[p_{i+1}, p_{i+\ell}] \subset [c+1, N-c]$. We will show that $\tilde{p}_{i+1} < c+1$ and that $\tilde{p}_{i+\ell+1} > N-c$, contradicting the assumption that $[c+1, n+1] \subset \widetilde{P}$ and hence the assumption that $\#([c+1, N-c] \cap \widetilde{P}) = \ell$ (by Lemma 4.6), since in this case there can be at most $\ell - 1$ elements of $\widetilde{P}$ contained in the interval $[c+1, N-c]$.

We will first show that $\tilde{p}_{i+1} < c+1$. Since $c \in \widetilde{P}$, it must be the case that $\tilde{p}_j = c$ for some $j \ge i$. We will show that $j = i+1$.

Note that if $p_i = \tilde{p}_i < t_{i+1}$ then $\#([1,c] \cap T) = \#([1,c] \cap P)$, and hence $T \not\prec P$ (since $\mathfrak{t}(T) \neq \mathfrak{t}(P)$), contradicting the assumption that $T \prec P$.

Also note that if $p_i = \tilde{p}_i = t_{i+1}$, then $t_{i+1} = c$, contradicting the assumption that $c \notin [T]$.

Therefore, $\tilde{p}_j = c$ for some $j > i$. Furthermore, $\tilde{p}_j$ is defined by (‡) since $c < t_{j+1}$, and hence $\tilde{p}_j \in \mathscr{C}_{P,T}$.

However, neither c nor $N-c$ is a visible cut in $D(P,T)$, since that would imply $T \not\preceq P$ by Corollary 4.8. It follows that c is an exceptional cut in $D(P,T)$.

Since c is an exceptional cut, and since $p_i = c$, row $i+1$ must be the only row crossing from column c to column $c+1$, by Corollary 4.8. In other words $t_{i+1} < c < p_{i+1}$ and $t_{i+2} > c$. Thus $j = i+1$ and $\tilde{p}_{i+1} = \tilde{p}_j = c < c+1$.

It remains to show that $\tilde{p}_{i+\ell+1} > N-c$.

Since $[c+1, n+1] \subset [T]$, and since $t_{i+1} < c < t_{i+2}$, Lemma 4.6 tells us that $[t_{i+2}, t_{i+\ell+1}] \subset [c+1, N-c]$. Furthermore, row $i+\ell+1$ must cross from column $N-c$ to column $N-c+1$, since $N-c+1 \leq p_{i+\ell+1}$.

In fact, $t_{i+\ell+1} < N-c+1 < p_{i+\ell+1}$, since $p_{i+\ell+1} \neq N-c+1$ (due to the fact that $c \in P$). Therefore $N-c+1$ is not a visible cut in $D(P,T)$. It is not even an apparent cut, since $t_i < c = p_i$. Finally, $N+1-c$ is not an exceptional cut, since $c \notin [T]$. Thus $N+1-c \notin \mathscr{C}_{P,T}$.

Note that $t_{i+\ell+2} > N-c+1$. Thus, if $\tilde{p}_{i+\ell+1}$ is defined by (†), then $\tilde{p}_{i+\ell+1} > N-c+1$. Furthermore, if $\tilde{p}_{i+\ell+1}$ is defined by ($\star$) or (‡), then since $N+1-c \notin \mathscr{C}_{P,T}$ and $t_{i+\ell+2} > N-c+1$, we must again have $\tilde{p}_{i+\ell+1} > N-c+1$.

Case 4: $N+1-c \in \widetilde{P} \cap P$. As in the previous case, let $\ell := n+1-c$. We have $p_j = N+1-c$ for some j. Since $[c+1, n+1] \subset [P]$, we have $[p_{j-\ell}, p_{j-1}] \subset [c+1, N-c]$ by Corollary 4.8. We will show that $\tilde{p}_j > N-c$ and $\tilde{p}_{j-\ell} < c+1$, contradicting the assumption $[c+1, n+1] \subset [\widetilde{P}]$.

We will first show that $\tilde{p}_j > N-c$. Suppose for the sake of contradiction that $\tilde{p}_k = N+1-c$ for some $k > j$. Then, since $N+1-c \notin T$, $\tilde{p}_k$ is defined by ($\star$) or (‡), and is therefore a cut in $D(P,T)$ by Observation 6.11. Since $c \notin [T]$, it follows that $\tilde{p}_k$ is not an exceptional cut. Since $t_k \leq \tilde{p}_k = p_j < p_k$, it follows that $\tilde{p}_k$ is not a visible cut either.

But $\tilde{p}_k$ is a cut, and hence $c-1$ must be a visible cut in $D(P,T)$. Since $c \notin T$, c must be a zero column in $D(P,T)$. But then c is a visible cut, so by Corollary 4.8 $T \not\preceq P$, a contradiction. It follows that $k = j$, and hence that $\tilde{p}_j = N+1-c > N-c$.

It remains to show that $\tilde{p}_{j-\ell} < c+1$. Since $p_j = \tilde{p}_j = N+1-c$, and since $N+1-c \notin T$, it must be the case that $\tilde{p}_j$ is defined by ($\star$), and therefore that p_j is a visible cut in $D(P,T)$. It follows that row j is the only row crossing from column $N-c$ to column $N-c+1$, and hence that c and $N-c$ are exceptional cuts in $D(P,T)$ by Corollary 4.8.

By Lemma 6.15, c is a visible cut in $D(\widetilde{P}, T)$. It follows that $\tilde{p}_{j-\ell} \leq c$, since $t_{j-\ell} \leq c$. □

We can now prove that $\widetilde{P}$ satisfies the latter half of condition (1) of Proposition 6.2.

Proposition 6.21. $\widetilde{P} \preceq P$.

Proof. Note that $\widetilde{P} \le P$ by construction. Assuming $\widetilde{P} < P$, suppose $\widetilde{P} \not\prec P$ for the sake of contradiction. It follows that $\mathfrak{t}(\widetilde{P}) \neq \mathfrak{t}(P)$, and that there exists an integer $c \in [1, n]$ such that $[c+1, n+1] \subset [\widetilde{P}] \cap [P]$ and $\#([1, c] \cap \widetilde{P}) = \#([1, c] \cap P)$.

Case 1: $n+1 \notin [T]$. Since $n+1 \in [\widetilde{P}]$, we have $\tilde{p}_i \in \{n+1, n+2\}$ for some i. Since $n+1 \notin [T]$, $\tilde{p}_i$ must be defined by ($\star$) or (‡). By Observation 6.11, $\tilde{p}_i \in \mathscr{C}_{P,T}$. If $\tilde{p}_i = n+1$, this means $n+1 \in \mathscr{C}_{P,T}$. If $\tilde{p}_i = n+2$, then also we have $n+1 \in \mathscr{C}_{P,T}$. To see why, note that both n and $n+1$ are in $\mathscr{Q}_{P,T}$, so by Lemma 4.14 either $n+1$ or $n+2$ is in $\mathscr{L}_{P,T}$.

If $n+1$ is an exceptional center cut in $D(P, T)$, then $p_j = n+2 < t_{j+1}$ for some j, since $n+1 \notin [T]$. In this case, $\tilde{p}_j = p_j = n+2$, and hence $n+1$ is a visible cut in $D(P, \widetilde{P})$. On the other hand, if $n+1$ is a visible cut in $D(P, T)$, then it must be a visible cut in $D(P, \widetilde{P})$ as well. It follows that $\#([1, n+1] \cap P) = \#([1, n+1] \cap \widetilde{P})$. In other words $\mathfrak{t}(P) = \mathfrak{t}(\widetilde{P})$, a contradiction.

Case 2: $n+1 \in [T]$. Since $n+1 \in [T] \cap [\widetilde{P}]$, Lemma 6.12 implies $\mathfrak{t}(T) = \mathfrak{t}(\widetilde{P})$. Since we are assuming $\mathfrak{t}(\widetilde{P}) \neq \mathfrak{t}(P)$, this means $\mathfrak{t}(T) \neq \mathfrak{t}(P)$. We can therefore invoke Lemma 6.20 $n-c$ times to deduce that $[c+1, n+1] \subset [T]$.

We divide the remainder of the proof into three subcases, depending on the number of rows crossing from column c to column $c+1$ of $D(P, T)$.

Case 2a: $\#([1, c] \cap T) = \#([1, c] \cap P)$. Since $\mathfrak{t}(T) \neq \mathfrak{t}(P)$ and $[c+1, n+1] \subset [T] \cap [P]$, we have $T \not\preceq P$, a contradiction.

Case 2b: $\#([1, c] \cap T) = \#([1, c] \cap P) + 1$. In other words there is exactly one integer i such that $t_i \le c < p_i$. Furthermore, c is an exceptional cut in $D(P, T)$, since $[c+1, n+1] \subset [T]$ and $\mathfrak{t}(T) \neq \mathfrak{t}(P)$. Therefore, by Lemma 6.15, c is a visible cut in $D(\widetilde{P}, T)$. This means $\tilde{p}_i \le c < t_{i+1}$. But $c < p_i$, so we have $\tilde{p}_i \le c < p_i$, contradicting the assumption that $\#([1, c] \cap \widetilde{P}) = \#([1, c] \cap P)$.

Case 2c: $\#([1, c] \cap T) \ge \#([1, c] \cap P) + 2$. Let i be the smallest integer such that $t_i \le c < p_i$. Note that $t_{i+1} \le c < p_{i+1}$. Since $\tilde{p}_i \le t_{i+1} \le c$, it follows that $\#([1, c] \cap \widetilde{P}) > \#([1, c] \cap P)$, again contradicting the assumption that $\#([1, c] \cap \widetilde{P}) = \#([1, c] \cap P)$. $\square$

Combining Propositions 6.17 and 6.21, and Corollaries 6.13 and 6.19, we see that the Schubert symbol $\widetilde{P}$ satisfies all three conditions of Proposition 6.2. By the discussion at the beginning of Section 6, this completes the proof of Theorem 1.1.

7. The Grothendieck ring

In this section we summarize some facts about K-theory which will be used in Sections 8 and 9. Further details can be found in [Fulton 1998] or [Brion 2005].

Given an algebraic variety X, let $K^0(X)$ denote the Grothendieck ring of algebraic vector bundles on X. Let $K_0(X)$ denote the Grothendieck group of coherent

$\mathcal{O}_X$-modules, which is a module over $K^0(X)$. Both the ring structure of $K^0(X)$ and the module structure of $K_0(X)$ are defined by tensor products. A closed subvariety $Z \subset X$ has a *Grothendieck class* $[\mathcal{O}_Z] \in K_0(X)$ defined by its structure sheaf. If X is nonsingular, the map $K^0(X) \to K_0(X)$ sending a vector bundle to its sheaf of sections is an isomorphism, and we write $K(X) := K_0(X) \cong K^0(X)$, which we refer to as the *Grothendieck ring of* X.

A morphism of varieties $f : X \to Y$ defines a pullback ring homomorphism $f^* : K^0(Y) \to K^0(X)$ by pullback of vector bundles. If f is proper, then there exists a pushforward group homomorphism $f_* : K_0(X) \to K_0(Y)$. Both these maps are functorial with respect to composition of morphisms. The *projection formula* says that f_* is a K^0-module homomorphism, in the sense that

$$f_*(f^*\mathcal{A} \cdot \mathcal{B}) = \mathcal{A} \cdot f_*\mathcal{B}, \tag{3}$$

where $\mathcal{A} \in K^0(Y)$ and $\mathcal{B} \in K_0(X)$. If X is a complete variety, then the sheaf Euler characteristic map $\chi_X : K(X) \to K(\text{point}) = \mathbb{Z}$ is defined to be the pushforward along the morphism $X \to \text{point}$.

We need the following well-known fact (see, e.g., [Buch and Ravikumar 2012, Lemma 2.2]).

Lemma 7.1. *Let X be a nonsingular variety, and let Y and Z be closed varieties of X with Cohen–Macaulay singularities. Assume that each component of $Y \cap Z$ has dimension* $\dim(Y) + \dim(Z) - \dim(X)$. *Then $Y \cap Z$ is Cohen–Macaulay and* $[\mathcal{O}_Y] \cdot [\mathcal{O}_Z] = [\mathcal{O}_{Y\cap Z}]$ *in $K(X)$.*

Finally we recall some facts about the K-theory of the projective space $\mathbb{P}^{N-1}$. Let $h \in K(\mathbb{P}^{N-1})$ be the class of a hyperplane. Then h^j is the class of a codimension-j linear subvariety, $2h - h^2$ is the class of a quadric hypersurface, and $K(\mathbb{P}^{N-1}) = \mathbb{Z}[h]/(h^N)$. The sheaf Euler characteristic $\chi_{\mathbb{P}^{N-1}} : K(\mathbb{P}^{N-1}) \to \mathbb{Z}$ is determined by $h^j \mapsto 1$ for $1 \le j < N$.

8. Computing triple intersection numbers

Let $X := \mathrm{IG}_\omega(m, \mathbb{C}^N)$ be a Grassmannian of type B, C, or D. In this section, we calculate the K-theoretic Pieri-type triple intersection numbers

$$\chi_X([\mathcal{O}_{X_P}] \cdot [\mathcal{O}_{X^T}] \cdot [\mathcal{O}_{X_{(r)}}]), \tag{4}$$

where $T \preceq P$ are Schubert symbols in $\Omega(X)$ and $X_{(r)} \subset X$ is a special Schubert variety, which we will define shortly for each type of Grassmannian. Our technique relies on the projection formula to move our calculation to the K-theory of projective space. A similar technique was used for the type-A Grassmannian in [Buch and Ravikumar 2012].

Recall the projections π and ψ from $\mathrm{IF} := \mathrm{IF}_\omega(1, m, \mathbb{C}^N)$ to X and $Z := \mathrm{IG}_\omega(1, \mathbb{C}^N)$ respectively. Since the K-theoretic pushforward is functorial with respect to composition, the following diagram commutes:

$$\begin{tikzcd} K(\mathrm{IF}) \arrow[r, "\psi_*"] \arrow[d, "\pi_*"'] \arrow[dr, "\chi_{\mathrm{IF}}"] & K(Z) \arrow[d, "\chi_Z"] \\ K(X) \arrow[r, "\chi_X"'] & \mathbb{Z} \end{tikzcd}$$

Lemma 8.1. *Let $X := \mathrm{IG}_\omega(m, \mathbb{C}^N)$ be a Grassmannian of type B, C, or D. Suppose there exists a Schubert variety $W \subset Z := \mathrm{IG}_\omega(1, \mathbb{C}^N)$ such that $\pi(\psi^{-1}(W)) = X_{(r)}$. We then have*

$$\chi_X([\mathcal{O}_{X_P}] \cdot [\mathcal{O}_{X^T}] \cdot [\mathcal{O}_{X_{(r)}}]) = \chi_Z([\mathcal{O}_{Z_{P,T}}] \cdot [\mathcal{O}_W]) \tag{5}$$

for any Schubert symbols $T \preceq P$.

Proof. Since $\pi^{-1}(X_P)$ and $\pi^{-1}(X^T)$ are opposite Schubert varieties in IF, it follows that $\pi^{-1}(Y_{P,T})$ is a Richardson variety in IF (see, e.g., [Brion 2005]). By [Knutson et al. 2014, Theorem 4.5] or [Billey and Coskun 2012, Theorem 3.3], the projection $\psi : \pi^{-1}(Y_{P,T}) \to Z_{P,T}$ is cohomologically trivial, in the sense that $\psi_*[\mathcal{O}_{\pi^{-1}(Y_{P,T})}] = [\mathcal{O}_{Z_{P,T}}]$. Since π is flat, it follows that

$$\psi_* \pi^* [\mathcal{O}_{Y_{P,T}}] = [\mathcal{O}_{Z_{P,T}}] \in K(Z).$$

Similarly, $\pi : \psi^{-1}(W) \to X_{(r)}$ is cohomologically trivial and π is flat, so we have

$$\pi_* \psi^* [\mathcal{O}_W] = [\mathcal{O}_{X_{(r)}}] \in K(X).$$

It is known that all Schubert varieties have rational singularities [Mehta and Srinivas 1987]. Therefore by Lemma 7.1 and two applications of the projection formula, we have

$$\begin{aligned} \chi_X([\mathcal{O}_{X_P}] \cdot [\mathcal{O}_{X^T}] \cdot [\mathcal{O}_{X_{(r)}}]) &= \chi_X([\mathcal{O}_{Y_{P,T}}] \cdot \pi_* \psi^* [\mathcal{O}_W]) \\ &= \chi_{\mathrm{IF}}(\pi^* [\mathcal{O}_{Y_{P,T}}] \cdot \psi^* [\mathcal{O}_W]) \\ &= \chi_Z(\psi_* \pi^* [\mathcal{O}_{Y_{P,T}}] \cdot [\mathcal{O}_W]) \\ &= \chi_Z([\mathcal{O}_{Z_{P,T}}] \cdot [\mathcal{O}_W]). \end{aligned}$$

□

By Lemma 8.1, the projected Richardson variety $Z_{P,T}$ assumes a crucial role in the calculation of (4). Let q be the number of quadratic equations and let l be the number of linear equations defining $Z_{P,T}$ as a complete intersection in $\mathbb{P}^{N-1}$ (from Sections 2 and 4 we know that $q = \#\{c \in \mathcal{Q}_{P,T} : c > 0 \text{ and } c - 1 \notin \mathcal{Q}_{P,T}\}$

and $l = \#\mathscr{L}_{P,T}$). We now calculate (4) when X has Lie-type C, B, or D, in that order, finally presenting a unified treatment in Corollary 8.7.

8A. ***Type C*.** Let $X = \mathrm{SG}(m, 2n)$. Note that $\mathrm{SG}(1, 2n)$, the image of ψ, is equal to $\mathbb{P}^{2n-1}$. The codimension-r special Schubert variety $X_{(r)}$ is defined for $1 \le r \le 2n-m$ by

$$X_{(r)} := \{\Sigma \in X : \dim(\Sigma \cap E_{2n-m-r+1}) \ge 1\}$$

In other words, $X_{(r)} = \pi(\psi^{-1}(W))$, where $W := \mathbb{P}(E_{2n-m-r+1})$. Note that W is a linear subvariety of $\mathbb{P}^{2n-1}$ and therefore a Schubert variety.

Proposition 8.2. *Let $X := \mathrm{SG}(m, 2n)$. Given $T \preceq P$ in $\Omega(X)$ and $1 \le r \le 2n-m$, we have*

$$\chi_X([\mathbb{O}_{X_P}] \cdot [\mathbb{O}_{X^T}] \cdot [\mathbb{O}_{X_{(r)}}]) = \sum_{j=0}^{2n-m-r-l-q} \binom{q}{j} (-1)^j (2)^{q-j},$$

where we define $\binom{q}{j}$ to be zero for $j > q$.

Proof. The triple intersection number is equal to $\chi_{\mathbb{P}^{2n-1}}([\mathbb{O}_{Z_{P,T}}] \cdot [\mathbb{O}_W])$, by Lemma 8.1. Since $Z_{P,T}$ is a complete intersection defined by l linear and q quadratic polynomials, we have $[\mathbb{O}_{Z_{P,T}}] = h^l h^q (2-h)^q$. Since $[\mathbb{O}_W] = h^{m-1+r}$, it follows that

$$\begin{aligned}[\mathbb{O}_{Z_{P,T}}] \cdot [\mathbb{O}_W] &= h^{m+r+l+q-1}(2-h)^q \\ &= \sum_{j=0}^{2n-m-r-l-q} h^{m+r+l+q-1} \binom{q}{j} (-h)^j (2)^{q-j},\end{aligned}$$

where we define $\binom{q}{j}$ to be zero for $j > q$. Taking sheaf Euler characteristic yields the desired triple intersection formula. $\square$

8B. ***Type B*.** Let $X := \mathrm{OG}(m, 2n+1)$. Let $Q := \mathrm{OG}(1, 2n+1)$ denote the quadric hypersurface of isotropic lines in $\mathbb{P}^{2n}$, with inclusion $\iota : Q \hookrightarrow \mathbb{P}^{2n}$.

We describe the Schubert varieties relative to $E_\bullet$ for the odd-dimensional quadric Q. For $0 \le j \le 2n-1$ there is exactly one codimension-j Schubert variety $Q_{(j)} \subset Q$, defined by

$$Q_{(j)} = \begin{cases} \mathbb{P}(E_{2n+1-j}) \cap Q & \text{if } 0 \le j \le n-1, \\ \mathbb{P}(E_{2n-j}) & \text{if } n \le j \le 2n-1. \end{cases}$$

The Schubert varieties $Q_{(j)}$ have a straightforward Bruhat ordering:

$$\underset{\mathbb{P}(E_1)}{\overset{\shortparallel}{Q_{(2n-1)}}} \subset \underset{\mathbb{P}(E_2)}{Q_{(2n-2)}} \subset \cdots \subset \underset{\mathbb{P}(E_n)}{Q_{(n)}} \subset \underset{\mathbb{P}(E_{n+2}) \cap Q}{Q_{(n-1)}} \subset \cdots \subset \underset{\mathbb{P}(E_{2n}) \cap Q}{Q_{(1)}} \subset Q_{(0)} = Q.$$

We mention some facts about the Schubert classes in $K(Q)$ (see [Buch and Samuel 2014] for details). For $0 \le j \le n-1$, we have $[\mathcal{O}_{Q_{(j)}}] = \iota^*(h^j)$. Pushforwards of Schubert classes are given by

$$\iota_*[\mathcal{O}_{Q_{(j)}}] = \begin{cases} h^j(2h-h^2) & \text{for } 0 \le j \le n-1, \\ h^{j+1} & \text{for } n \le j \le 2n-1. \end{cases}$$

Returning to the type-B Grassmannian X, the codimension-r special Schubert variety $X_{(r)}$ is defined by

$$X_{(r)} = \{\Sigma \in X : \mathbb{P}(\Sigma) \cap Q_{(m+1-r)} \neq \varnothing\}$$

for $1 \le r \le 2n-m$. In other words, $X_{(r)} = \pi(\psi^{-1}(Q_{(m+1-r)}))$.

We now rewrite the type-B triple intersection number as the sheaf Euler characteristic of a $K(\mathbb{P}^{2n})$ class. The final step of computing Euler characteristic is exactly the same as in Proposition 8.2, and is postponed to the unified formula in Corollary 8.7.

Proposition 8.3. *Let $X := \mathrm{OG}(m, 2n+1)$. For $T \preceq P$ in $\Omega(X)$ and $1 \le r \le 2n-m$, we have*

$$\chi_X([\mathcal{O}_{X_P}] \cdot [\mathcal{O}_{X^T}] \cdot [\mathcal{O}_{X_{(r)}}]) \\ = \begin{cases} \chi_{\mathbb{P}^{2n}}(h^{m+r+l+q-1}(2-h)^q) & \text{if } r \le n-m, \\ \chi_{\mathbb{P}^{2n}}(h^{m+r+l+q-1}(2-h)^{q-1}) & \text{if } r > n-m \text{ and } q > 0, \\ \chi_{\mathbb{P}^{2n}}(h^{m+r+l-1}) = 0 & \text{if } r > n-m \text{ and } q = 0. \end{cases}$$

Proof. By Lemma 8.1, we must simplify $\chi_Q([\mathcal{O}_{Z_{P,T}}] \cdot [\mathcal{O}_{Q_{(m-1+r)}}])$. In certain situations we can use the projection formula along ι to do this.

Situation 1: Suppose $r \le n-m$. In this case, $m-1+r \le n-1$. It follows that the inclusion $\iota(Q_{(m-1+r)})$ is a complete intersection in $\mathbb{P}^{2n}$ cut out by $m-1+r$ linear equations and the single quadratic equation defining Q. Thus, $[\mathcal{O}_{Q_{(m-1+r)}}] = \iota^*(h^{m-1+r})$. Using the projection formula, we have

$$\begin{aligned} \chi_Q([\mathcal{O}_{Z_{P,T}}] \cdot [\mathcal{O}_{Q_{(m-1+r)}}]) &= \chi_Q([\mathcal{O}_{Z_{P,T}}] \cdot \iota^*(h^{m-1+r})) \\ &= \chi_{\mathbb{P}^{2n}}([\mathcal{O}_{Z_{P,T}}] \cdot h^{m-1+r}) \\ &= \chi_{\mathbb{P}^{2n}}(h^{l+q+m+r-1}(2-h)^q). \end{aligned}$$

Situation 2: Suppose q, the number of quadratic equations defining $Z_{P,T}$, is greater than zero. By ignoring one of the quadratic equations defining $Z_{P,T}$, we define a larger subvariety $Z' \subset \mathbb{P}^{2n}$ such that $Z_{P,T} = Z' \cap Q$. It follows that

$[\mathscr{O}_{Z_{P,T}}] = \iota^*[\mathscr{O}_{Z'}]$, so by the projection formula we have

$$
\begin{aligned}
\chi_Q([\mathscr{O}_{Z_{P,T}}]\cdot[\mathscr{O}_{Q_{(m-1+r)}}])& \\
&= \chi_Q(\iota^*[\mathscr{O}_{Z'}]\cdot[\mathscr{O}_{Q_{(m-1+r)}}]) \\
&= \chi_{\mathbb{P}^{2n}}([\mathscr{O}_{Z'}]\cdot\iota_*[\mathscr{O}_{Q_{(m-1+r)}}]) \\
&= \begin{cases} \chi_{\mathbb{P}^{2n}}(h^l h^{q-1}(2-h)^{q-1}\cdot h^{m-1+r}(2h-h^2)) & \text{if } r \le n-m, \\ \chi_{\mathbb{P}^{2n}}(h^l h^{q-1}(2-h)^{q-1}\cdot h^{m+r}) & \text{if } r > n-m, \end{cases} \\
&= \begin{cases} \chi_{\mathbb{P}^{2n}}(h^{l+q+m+r-1}(2-h)^{q}) & \text{if } r \le n-m, \\ \chi_{\mathbb{P}^{2n}}(h^{l+q+m+r-1}(2-h)^{q-1}) & \text{if } r > n-m. \end{cases}
\end{aligned}
$$

Note that these situations are not mutually exclusive, and that when $r \le n-m$ and $q > 0$ both methods agree.

Situation 3: Suppose $r > n-m$ and $q = 0$. In this case, both $Z_{P,T}$ and $Q_{(m-1+r)}$ can be thought of as linear subvarieties of $\mathbb{P}^{2n}$ that are contained in Q. Note that $\iota_*[\mathscr{O}_{Z_{P,T}}] = h^l$ and $\iota_*[\mathscr{O}_{Q_{(m-1+r)}}] = h^{m+r}$, but that

$$\iota_*([\mathscr{O}_{Z_{P,T}}]\cdot[\mathscr{O}_{Q_{(m-1+r)}}]) = h^{m+r+l-1},$$

since one of these linear equations is redundant in the intersection of generic translates of $Z_{P,T}$ and $Q_{(m-1+r)}$. However, the integer $m+r+l-1$ is at least $2n+1$, and therefore the Grothendieck class $h^{m+r+l-1} \in K(\mathbb{P}^{2n})$ vanishes. $\square$

8C. ***Type D.*** Let $X := \mathrm{OG}(m, 2n+2)$. Let $Q := \mathrm{OG}(1, 2n+1)$ denote the quadric hypersurface of isotropic lines in $\mathbb{P}^{2n+1}$ with inclusion $\iota : Q \hookrightarrow \mathbb{P}^{2n+1}$. We describe the Schubert varieties relative to $E_\bullet$ for the even-dimensional quadric Q. Let $\widetilde{E}_{n+1} = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n, \boldsymbol{e}_{n+2}\rangle$.

The quadric Q has *two* Schubert varieties of codimension n, defined by

$$Q_{(n)} := \mathbb{P}(E_{n+1}) \quad \text{and} \quad \widetilde{Q}_{(n)} := \mathbb{P}(\widetilde{E}_{n+1}).$$

For $j \ne n$, there is a single codimension-j Schubert variety, defined by

$$Q_{(j)} = \begin{cases} \mathbb{P}(E_{2n+2-j}) \cap Q & \text{if } 0 \le j \le n-1, \\ \mathbb{P}(E_{2n+1-j}) & \text{if } n+1 \le j \le 2n. \end{cases}$$

The Schubert varieties in Q have the following Bruhat order:

$$
\begin{tikzcd}[column sep=small]
\mathbb{P}(E_1) \arrow[d, equal] & \mathbb{P}(E_2) \arrow[d, equal] & \mathbb{P}(E_n) \arrow[d, equal] & & Q_{(n)} = \mathbb{P}(E_{n+1}) & & & \mathbb{P}(E_{2n+1}) \cap Q \arrow[d, equal] & \\
Q_{(2n)} \arrow[r, phantom, "\subset"] & Q_{(2n-1)} \arrow[r, phantom, "\subset \cdots \subset"] & Q_{(n+1)} \arrow[urr, phantom, "\subset"] \arrow[drr, phantom, "\subset"] & & & Q_{(n-1)} \arrow[ul, phantom, "\supset"] \arrow[dl, phantom, "\supset"] \arrow[r, phantom, "\subset \cdots \subset"] & & Q_{(1)} \arrow[d, equal] \arrow[r, phantom, "\subset"] & Q_{(0)} = Q \\
& & & & \mathbb{P}(\widetilde{E}_{n+1}) = \widetilde{Q}_{(n)} & & & \mathbb{P}(E_{n+3}) \cap Q &
\end{tikzcd}
$$

We mention some facts about the Grothendieck ring $K(Q)$, which can be found in [Buch and Samuel 2014, pp. 17–18]. As in type B, we have $[\mathbb{O}_{Q_{(j)}}] = \iota^*(h^j)$ for $0 \le j \le n-1$. Pushforwards of Schubert classes are the same as in type B, the only addition being that $\iota_*[\mathbb{O}_{Q_{(n)}}] = \iota_*[\mathbb{O}_{\tilde{Q}_{(n)}}] = h^{n+1}$. The products of codimension-n classes with one another depend on the parity of n, in the sense that

$$[\mathbb{O}_{Q_{(n)}}]^2 = [\mathbb{O}_{\tilde{Q}_{(n)}}]^2 = \begin{cases} [\mathbb{O}_{Q_{(2n)}}] = [\mathbb{O}_{\text{point}}] & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd,} \end{cases}$$

whereas

$$[\mathbb{O}_{Q_{(n)}}] \cdot [\mathbb{O}_{\tilde{Q}_{(n)}}] = \begin{cases} 0 & \text{if } n \text{ is even,} \\ [\mathbb{O}_{Q_{(2n)}}] = [\mathbb{O}_{\text{point}}] & \text{if } n \text{ is odd.} \end{cases}$$

The maximal even orthogonal Grassmannian $\mathrm{OG}(n+1, 2n+2)$ has two connected components. For any $\Sigma \in \mathrm{OG}(n+1, 2n+2)$, let $\mathfrak{t}(\Sigma) \in \{0, 1\}$ to be the codimension mod 2 of $\Sigma \cap E_{n+1}$ in E_{n+1}. If $L \subset Q$ and $L' \subset Q$ are linear subvarieties of codimension n in Q, then the affine cones $\Lambda(L)$ and $\Lambda(L')$ are elements of $\mathrm{OG}(n+1, 2n+2)$. The $K(Q)$ classes $[\mathbb{O}_L]$ and $[\mathbb{O}_{L'}]$ are equal if and only if $\Lambda(L)$ and $\Lambda(L')$ are in the same $\mathrm{SO}(2n+2)$-orbit, which is the case if and only if $\mathfrak{t}(\Lambda(L)) = \mathfrak{t}(\Lambda(L'))$. For any codimension-$n$ linear subvariety $L \subset Q$, we let $\mathfrak{t}([\mathbb{O}_L]) := \mathfrak{t}(\Lambda(L))$. It is easy to check that $\mathfrak{t}([\mathbb{O}_{Q_{(n)}}]) = 0$ and $\mathfrak{t}([\mathbb{O}_{\tilde{Q}_{(n)}}]) = 1$. Hence, for $\mathcal{A}$ and $\mathcal{B}$ in $\{[\mathbb{O}_{Q_{(n)}}], [\mathbb{O}_{\tilde{Q}_{(n)}}]\}$, we have

$$\mathcal{A} \cdot \mathcal{B} = ((\mathfrak{t}(\mathcal{A}) + \mathfrak{t}(\mathcal{B}) + n + 1) \pmod 2)[\mathbb{O}_{Q_{(2n)}}],$$

where mod 2 means “remainder mod 2”: the coefficient is 1 if $\mathfrak{t}(\mathcal{A}) + \mathfrak{t}(\mathcal{B}) + n + 1$ is odd, and 0 otherwise.

The following lemma describes $\mathfrak{t}([\mathbb{O}_{Z_{P,T}}])$ whenever $Z_{P,T}$ has codimension n in Q. The lemma is adapted from the definition of the function $h(P, T)$ in [Buch et al. 2009, §5.2].

Lemma 8.4. *If $Z_{P,T}$ is a linear subvariety of codimension n in the quadric Q, then $\mathfrak{t}(Z_{P,T}) \equiv |S| + |S'| + n + 1 \pmod 2$, where*

$$S = \{i \in [1, n+1] : t_j \le i \le p_j \text{ for some } j\},$$
$$S' = \{p \in P : p \ge n+2 \text{ and } 2n+3-p \in S\}.$$

Proof. Let $\Lambda(Z_{P,T}) \subset \mathbb{C}^{2n+2}$ be the affine cone over $Z_{P,T} \subset \mathbb{P}^{2n+1}$. Note that $|S|$ is the number of $c \in [1, n+1]$ such that c is not a zero column in $D(P, T)$, and that $|S'|$ is the number of nonzero columns $c \in [1, n+1]$ such that column $N+1-c$ of $D(P, T)$ contains a lone star. It follows that $|S| - |S'|$ is the number of $c \in [1, n+1]$ such that $\boldsymbol{e}_c \in \Lambda(Z_{P,T})$, and hence that $n+1-(|S|-|S'|)$ is the codimension of $\Lambda(Z_{P,T}) \cap E_{n+1}$ in E_{n+1}. $\square$

We give an example in which $\mathfrak{t}([\mathcal{O}_{Z_{P,T}}])$ affects a product in $K(Q)$ (which is in fact a triple intersection number).

Example 8.5. Consider $\mathrm{OG}(2,8)$, and let $P=\{1,4\}$ and $T=\{1,2\}$. In this case $\mathfrak{t}([\mathcal{O}_{Z_{P,T}}])=0$. It follows that $[\mathcal{O}_{Z_{P,T}}]\cdot[\mathcal{O}_{Q_{(n)}}]=(0+0+3+1 \bmod 2)[\mathcal{O}_{Q_{(2n)}}]=0$ and $[\mathcal{O}_{Z_{P,T}}]\cdot[\mathcal{O}_{\widetilde{Q}_{(n)}}]=(0+1+3+1 \bmod 2)[\mathcal{O}_{Q_{(2n)}}]=[\mathcal{O}_{Q_{(2n)}}]$.

Returning to the even orthogonal Grassmannian X, the codimension-r special Schubert variety $X_{(r)}$ is defined by $X_{(r)}:=\{\Sigma\in X:\mathbb{P}(\Sigma)\cap Q_{(m-1+r)}\neq\varnothing\}$ for $1\le r\le 2n+1-m$. As in the quadric case, there is an additional special Schubert variety[3] of codimension $k:=n-m+1$, defined by $\widetilde{X}_{(k)}:=\{\Sigma\in X:\mathbb{P}(\Sigma)\cap\widetilde{Q}_{(n)}\neq\varnothing\}$. Thus $X_{(r)}=\pi(\psi^{-1}(Q_{(m-1+r)}))$ for $1\le r\le 2n+1-m$ and $\widetilde{X}_{(k)}=\pi(\psi^{-1}(\widetilde{Q}_{(n)}))$.

Consider the triple intersection number corresponding to X_P, X^T, and a codimension-r special Schubert variety. By Lemma 8.1, this equals $\chi_Q([\mathcal{O}_{Z_{P,T}}]\cdot\mathcal{A})$, where $\mathcal{A}$ is the corresponding codimension-$(m-1+r)$ special Schubert class in $K(Q)$. We now translate this expression to the sheaf Euler characteristic of a $K(\mathbb{P}^{2n+1})$ class. The final step of computing Euler characteristic is postponed to the unified formula in Corollary 8.7.

Proposition 8.6. *Let $X:=\mathrm{OG}(m,2n+2)$, and let $\mathcal{A}$ be a codimension-$(m-1+r)$ Schubert class in $K(Q)$. Define $\delta\in\{0,1\}$ by*

$$\delta\equiv\begin{cases}\mathfrak{t}(\mathcal{A})+|S|+|S'| \pmod 2 & \text{if } r=k, q=0, \text{ and } l=n+1,\\ 1 \pmod 2 & \text{otherwise.}\end{cases}$$

We then have

$$\chi_Q([\mathcal{O}_{Z_{P,T}}]\cdot\mathcal{A})=\begin{cases}\chi_{\mathbb{P}^{2n+1}}(h^{m+r+l+q-1}(2-h)^q) & \text{if } r<k,\\ \chi_{\mathbb{P}^{2n+1}}(h^{m+r+l+q-1}(2-h)^{q-1}) & \text{if } r\ge k \text{ and } q>0,\\ \chi_{\mathbb{P}^{2n+1}}(\delta\cdot h^{m+r+l-1}) & \text{if } r\ge k \text{ and } q=0.\end{cases}$$

Proof. The proof is exactly as in type B, except in the case $r\ge k$ and $q=0$. In this case, $Z_{P,T}$ is a linear subvariety of Q of codimension at least n. If $r>k$ or if $Z_{P,T}$ has codimension greater than n, then $[\mathcal{O}_{Z_{P,T}}]\cdot\mathcal{A}=0$ (in this case $h^{m+r+l-1}\in K(\mathbb{P}^{2n+1})$ is also zero, since $r>n-m+1$ or $l>n+1$). We can therefore assume $[\mathcal{O}_{Z_{P,T}}]$ and $\mathcal{A}$ are both in $\{[\mathcal{O}_{Q_{(n)}}],[\mathcal{O}_{\widetilde{Q}_{(n)}}]\}$. By Lemma 8.4, it follows that

$$\begin{aligned}\chi_Q([\mathcal{O}_{Z_{P,T}}]\cdot\mathcal{A})&=\mathfrak{t}([\mathcal{O}_{Z_{P,T}}])+\mathfrak{t}(\mathcal{A})+n+1 \pmod 2\\&=|S|+|S'|+\mathfrak{t}(\mathcal{A}) \pmod 2.\end{aligned}$$

This number agrees with $\chi_{\mathbb{P}^{2n+1}}(\delta\cdot h^{m+r+l-1})$, since $m+r+l-1=2n+1$. □

[3] We note that our definition of the codimension-k special Schubert varieties differs slightly from the definition in [Buch et al. 2009, §3.2], in which $X_{(k)}$ and $\widetilde{X}_{(k)}$ are switched when n is odd.

8D. ***A general formula.*** Propositions 8.2, 8.3, and 8.6 are summarized concisely in the following formulation of the triple intersection number, which holds for isotropic Grassmannians of all types:

Corollary 8.7. *Let $X := \mathrm{IG}_\omega(m, N)$ be an isotropic Grassmannian, where $N = 2n$ in type C, $N = 2n+1$ in type B, and $N = 2n+2$ in type D. Let $k = n - m$ in types B and C and let $k = n - m + 1$ in type D. Given $1 \le r \le n + k$, suppose $\mathcal{A} \in K(\mathrm{IG}_\omega(1, N))$ is a Schubert class of codimension $(m - 1 + r)$, so that $\pi_* \psi^* \mathcal{A} \in K(X)$ is a special Schubert class of codimension r. Given $T \preceq P$ in $\Omega(X)$, let l and q be the numbers of linear and quadratic equations defining $Z_{P,T}$ as a complete intersection, and let S and S' be the sets defined in Lemma 8.4. Define the integers l', q', η, and δ as follows*:

$$q' = \begin{cases} q-1 & \text{if } X \text{ is orthogonal and } q > 0,\\ q & \text{otherwise,} \end{cases}$$

$$l' = \begin{cases} l+m+r & \text{if } X \text{ is orthogonal, } q > 0, \text{ and } r \ge k,\\ l+m+r-1 & \text{otherwise,} \end{cases}$$

$$\eta = \begin{cases} \mathfrak{t}(\mathcal{A}) + |S| + |S'| & \text{if } X \text{ is type } D,\ q = 0, \text{ and } r = k,\\ 1 & \text{otherwise,} \end{cases}$$

$$\delta = \begin{cases} 0 & \text{if } \eta \text{ is even},\\ 1 & \text{if } \eta \text{ is odd.} \end{cases}$$

We then have

$$\begin{aligned}\chi_X([\mathcal{O}_{X_P}] \cdot [\mathcal{O}_{X^T}] \cdot \pi_* \psi^* \mathcal{A}) &= \chi_{\mathbb{P}^{N-1}}(\delta h^{l'} (2h - h^2)^{q'})\\ &= \delta \cdot \sum_{j=0}^{N-1-l'-q'} \binom{q'}{j} (-1)^j (2)^{q'-j},\end{aligned}$$

where we define $\binom{q'}{j}$ to be zero for $j > q'$.

9. Computing Pieri coefficients

Let X be an isotropic Grassmannian of type B, C, or D. Given Schubert symbols P and Q and a special Schubert class $[\mathcal{O}_{X_{(r)}}]$, the K-theoretic structure constant $\mathcal{N}^{Q}_{P,r}(X)$ is the coefficient of $[\mathcal{O}_{X_Q}]$ in the Pieri product $[\mathcal{O}_{X_P}] \cdot [\mathcal{O}_{X_{(r)}}]$. In this section, we compute $\mathcal{N}^{Q}_{P,r}(X)$ as an alternating sum of triple intersection numbers $\chi_X([\mathcal{O}_{X_P}] \cdot [\mathcal{O}_{X^T}] \cdot [\mathcal{O}_{X_{(r)}}])$, where T ranges over a certain subset of Schubert symbols (if X is type D, the special Schubert class $[\mathcal{O}_{\widetilde{X}_{(k)}}]$ can be substituted in order to calculate the additional Pieri coefficients $\widetilde{\mathcal{N}}^{Q}_{P,r}(X)$).

Given Schubert symbols Q and P, the Möbius function $\mu(Q,P)$ is defined by

$$\mu(Q,P)=\begin{cases}1 & \text{if } Q=P,\\ -\sum_{Q\preceq T\prec P}\mu(Q,T) & \text{if } Q\prec P,\\ 0 & \text{otherwise.}\end{cases}$$

For each $Q\in\Omega(X)$, we define a class $\mathbb{O}_Q^*\in K(X)$ by

$$\mathbb{O}_Q^*:=\sum_{T\in\Omega(X)}\mu(Q,T)[\mathbb{O}_{X^T}].$$

Lemma 9.1. *The class $\mathbb{O}_Q^*$ is the K-theoretic dual to $\mathbb{O}_Q$, in the sense that*

$$\chi_X([\mathbb{O}_{X_P}]\cdot\mathbb{O}_Q^*)=\delta_{P,Q}.$$

Proof. Given Schubert symbols $T\preceq P$, the Richardson variety $Y_{P,T}$ is rational [Deodhar 1985] with rational singularities [Brion 2002]. By [Griffiths and Harris 1978, p. 494] it follows that

$$\chi_X([\mathbb{O}_{X_P}]\cdot[\mathbb{O}_{X^T}])=\begin{cases}1 & \text{if } T\preceq P,\\ 0 & \text{otherwise.}\end{cases}$$

If $Q\prec P$, then we have

$$\begin{aligned}\chi_X([\mathbb{O}_{X_P}]\cdot\mathbb{O}_Q^*)&=\sum_{T\in\Omega(X)}\mu(Q,T)\chi_X([\mathbb{O}_{X_P}]\cdot[\mathbb{O}_{X^T}])\\ &=\sum_{T\preceq P}\mu(Q,T)\\ &=\sum_{Q\preceq T\prec P}\mu(Q,T)+\mu(Q,P)=0.\end{aligned}$$

If $Q\not\preceq P$, then for any $T\succeq Q$ we have $T\not\preceq P$. Thus, $[\mathbb{O}_{X_P}]\cdot[\mathbb{O}_{X^T}]=0$ for every Schubert class $[\mathbb{O}_{X^T}]$ that has nonzero coefficient in $\mathbb{O}_Q^*$. Finally, if $P=Q$ then $\chi_X([\mathbb{O}_{X_Q}]\cdot\mathbb{O}_Q^*)=\chi_X([\mathbb{O}_{X_Q}]\cdot[\mathbb{O}_{X^Q}])=1$. □

Since $\mu(Q,T)=0$ for $Q\not\preceq T$ and $[\mathbb{O}_{X_P}]\cdot[\mathbb{O}_{X^T}]=0$ for $T\not\preceq P$, we have the following corollary:

Corollary 9.2. $$\mathcal{N}_{P,r}^{Q}(X)=\sum_{Q\preceq T\preceq P}\mu(Q,T)\chi_X([\mathbb{O}_{X_P}]\cdot[\mathbb{O}_{X^T}]\cdot[\mathbb{O}_{X_{(r)}}]).$$

It is known that $\mu(Q,T)\in\{0,(-1)^{|Q|-|T|}\}$ for any Schubert symbols Q and T [Björner and Brenti 2005, Corollary 2.7.10]. In [Ravikumar 2013, Appendix A] a conjectured criterion is stated for when $\mu(Q,T)$ vanishes. We hope Corollary 9.2 will lead to a Pieri formula for $\mathcal{N}_{P,r}^{Q}(X)$ with manifestly alternating signs, in the sense that $(-1)^{|Q|-|P|-r}\mathcal{N}_{P,r}^{Q}(X)=1$ (see [Brion 2002] for a proof of this fact).

9A. ***A global rule.*** We briefly describe how to determine K-theoretic dual classes, and hence Pieri coefficients, without a "local" rule, but rather using the global data of the entire Bruhat order. This method requires us to invert an $L \times L$ matrix, where L is the number of Schubert symbols in $\Omega(X)$, and allows for relatively efficient computation of $\mathcal{N}_{P,r}^{Q}(X)$.

Let $\{P_1, \dots, P_L\}$ be the set of Schubert symbols for X. Let $\mathcal{O}_i := [\mathcal{O}_{X_{P_i}}]$ and $\mathcal{O}^i := [\mathcal{O}_{X^{P_i}}]$. The sets $\{\mathcal{O}_1, \dots, \mathcal{O}_L\}$ and $\{\mathcal{O}^1, \dots, \mathcal{O}^L\}$ are both additive bases for $K(X)$.

We will use the following four $L \times L$ matrices:

(1) Let $\boldsymbol{M} := (m_{ij})$ be the *intersection matrix* for X, where

$$m_{ij} = \begin{cases} 1 & \text{if } P_j \preceq P_i, \\ 0 & \text{otherwise.} \end{cases}$$

(2) Let $\boldsymbol{C}_{(r)} := (c_{ij})$ be the *Pieri coefficient matrix* for X, where

$$\mathcal{O}_i \cdot \mathcal{O}_{(r)} = c_{ij} \mathcal{O}_j.$$

(3) Let $\boldsymbol{T}_{(r)} := (t_{ij})$ be the *triple intersection matrix* for X, where

$$t_{ij} = \chi_X(\mathcal{O}_i \cdot \mathcal{O}^j \cdot \mathcal{O}_{(r)}).$$

(4) Let $\boldsymbol{D} := \boldsymbol{M}^{-1}$ be the *matrix of duals* on X.

Let $\boldsymbol{d}^j$ denote the j-th column vector of $\boldsymbol{D}$, and let $\mathcal{O}_{\boldsymbol{d}^j} := \sum_{k=1}^{L} d_{kj} \mathcal{O}^k$.

Observation 9.3. *The element $\mathcal{O}_{\boldsymbol{d}^j}$ is dual to $\mathcal{O}_j$ in the sense that*

$$\chi_X(\mathcal{O}_i \cdot \mathcal{O}_{\boldsymbol{d}^j}) = \delta_{i,j}.$$

Proof. $\chi_X(\mathcal{O}_i \cdot \mathcal{O}_{\boldsymbol{d}^j}) = \boldsymbol{m}_i \cdot \boldsymbol{d}^j$, where $\boldsymbol{m}_i$ is the i-th row of $\boldsymbol{M}$. □

Observation 9.4. *The matrix $\boldsymbol{D}$ transforms triple intersection numbers into Pieri coefficients, via the relation*

$$\boldsymbol{T}_{(r)} \cdot \boldsymbol{D} = \boldsymbol{C}_{(r)}.$$

Proof. $$\sum_{k=1}^{L} t_{ik} d_{kj} = \sum_{k=1}^{L} \chi_X(d_{kj} \mathcal{O}_i \cdot \mathcal{O}^k \cdot \mathcal{O}_{(r)}) = \chi_X(\mathcal{O}_i \cdot \mathcal{O}_{\boldsymbol{d}^j} \cdot \mathcal{O}_{(r)}) = c_{ij}.$$ □

Acknowledgments

The results of this paper are part of the author's dissertation [Ravikumar 2013]. The author wishes to sincerely thank Anders Buch for his encouragement and motivation over the years.

References

[Billey and Coskun 2012] S. Billey and I. Coskun, "Singularities of generalized Richardson varieties", *Comm. Algebra* **40**:4 (2012), 1466–1495. MR 2912998 Zbl 1274.14059

[Björner and Brenti 2005] A. Björner and F. Brenti, *Combinatorics of Coxeter groups*, Graduate Texts in Mathematics **231**, Springer, New York, 2005. MR 2006d:05001 Zbl 1110.05001

[Borel 1956] A. Borel, "Groupes linéaires algébriques", *Ann. of Math.* (2) **64** (1956), 20–82. MR 19,1195h Zbl 0070.26104

[Brion 2002] M. Brion, "Positivity in the Grothendieck group of complex flag varieties", *J. Algebra* **258**:1 (2002), 137–159. MR 2003m:14017 Zbl 1052.14054

[Brion 2005] M. Brion, "Lectures on the geometry of flag varieties", pp. 33–85 in *Topics in cohomological studies of algebraic varieties*, edited by P. Pragacz, Birkhäuser, Basel, 2005. MR 2006f:14058

[Buch and Ravikumar 2012] A. S. Buch and V. Ravikumar, "Pieri rules for the K-theory of cominuscule Grassmannians", *J. Reine Angew. Math.* **668** (2012), 109–132. MR 2948873 Zbl 1298.14059

[Buch and Samuel 2014] A. S. Buch and M. J. Samuel, "K-Theory of minuscule varieties", *J. Reine Angew. Math.* (online publication July 2014).

[Buch et al. 2009] A. S. Buch, A. Kresch, and H. Tamvakis, "Quantum Pieri rules for isotropic Grassmannians", *Invent. Math.* **178**:2 (2009), 345–405. MR 2011a:14112 Zbl 1193.14071

[Buch et al. 2013] A. S. Buch, A. Kresch, and H. Tamvakis, "A Giambelli formula for even orthogonal Grassmannians", *J. Reine Angew. Math.* (online publication September 2013).

[Deodhar 1985] V. V. Deodhar, "On some geometric aspects of Bruhat orderings, I: A finer decomposition of Bruhat cells", *Invent. Math.* **79**:3 (1985), 499–511. MR 86f:20045 Zbl 0563.14023

[Fulton 1998] W. Fulton, *Intersection theory*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **2**, Springer, Berlin, 1998. MR 99d:14003 Zbl 0885.14002

[Fulton and Harris 1991] W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics **129**, Springer, New York, 1991. MR 93a:20069 Zbl 0744.22001

[Griffiths and Harris 1978] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Wiley, New York, 1978. Reprinted 1994. MR 95d:14001 Zbl 0836.14001

[He and Lam 2011] X. He and T. Lam, "Projected Richardson varieties and affine Schubert varieties", preprint, 2011. arXiv 1106.2586v1

[Kleiman 1974] S. L. Kleiman, "The transversality of a general translate", *Compositio Math.* **28** (1974), 287–297. MR 50 #13063 Zbl 0288.14014

[Knutson et al. 2014] A. Knutson, T. Lam, and D. E. Speyer, "Projections of Richardson varieties", *J. Reine Angew. Math.* **687** (2014), 133–157. MR 3176610 Zbl 06296354

[Mehta and Srinivas 1987] V. B. Mehta and V. Srinivas, "Normality of Schubert varieties", *Amer. J. Math.* **109**:5 (1987), 987–989. MR 88j:14063 Zbl 0651.14030

[Ravikumar 2013] V. Ravikumar, *Triple Intersection Formulas for Isotropic Grassmannians*, Ph.D. thesis, Rutgers University, 2013, http://rucore.libraries.rutgers.edu/rutgers-lib/41905/.

[Richardson 1992] R. W. Richardson, "Intersections of double cosets in algebraic groups", *Indag. Math. (N.S.)* **3**:1 (1992), 69–77. MR 93b:20081 Zbl 0833.22001

[Springer 2009] T. A. Springer, *Linear algebraic groups*, 2nd ed., Birkhäuser, Boston, 2009. MR 2009i:20089 Zbl 1202.20048

Communicated by Ravi Vakil
Received 2014-06-25 Revised 2014-08-07 Accepted 2015-03-07

vijayr@cmi.ac.in

Mathematics Department, Chennai Mathematical Institute, H1, SIPCOT IT Park, Siruseri, Kelambakkam 603103, India

ALGEBRA AND NUMBER THEORY 9:3 (2015)
dx.doi.org/10.2140/ant.2015.9.725

On the basepoint-free theorem for log canonical threefolds over the algebraic closure of a finite field

Diletta Martinelli, Yusuke Nakamura and Jakub Witaszek

We prove the basepoint-free theorem for big line bundles on a three-dimensional log canonical projective pair defined over the algebraic closure of a finite field. This theorem is not valid for any other algebraically closed field.

1. Introduction

A line bundle L is called semiample if some positive tensor power $L^{\otimes r}$ is generated by global sections. Semiample line bundles play an important role in algebraic geometry, because they determine morphisms of a variety into projective spaces. Therefore, one would like to find necessary and sufficient conditions for semiampleness. A semiample line bundle is necessarily nef, but the converse is false in general. However, if we assume that L is the canonical bundle and is nef, then the abundance conjecture [Kollár and Mori 1998, Conjecture 3.12] states that L must be semiample. Furthermore, the basepoint-free theorem [Kollár and Mori 1998, Theorem 3.3] asserts that a nef line bundle L on a Kawamata log terminal projective pair (X, Δ) defined over an algebraically closed field of characteristic zero is semiample when $L - (K_X + \Delta)$ is nef and big.

In positive characteristic, questions regarding semiampleness are more difficult, due to the absence of a proof of the resolution of singularities for varieties of dimension greater than three and the failure of the Kawamata–Viehweg vanishing theorem. As such, the basepoint-free theorem remains still unsolved in general. However, many partial results for threefolds may be obtained by reductions to the two-dimensional cases.

The basepoint-free theorem in positive characteristic is known for big line bundles L when (X, Δ) is a three-dimensional Kawamata log terminal projective pair defined over an algebraically closed field of characteristic larger than five (see [Birkar 2013; Xu 2013]). Over $\overline{\mathbb{F}}_p$, the algebraic closure of a finite field, there is a stronger result,

MSC2010: primary 14E30; secondary 14C20.
Keywords: basepoint-free theorem, semiample line bundles, positive characteristic, finite fields.

due to Keel [1999], who proved the basepoint-free theorem for big line bundles L when (X, Δ) is a three-dimensional projective log pair defined over $\overline{\mathbb{F}}_p$ with all coefficients of Δ less than one.

In this paper, we generalize Keel's result to the cases where the coefficients of Δ may be equal to one. Our main theorem is the following:

Theorem 1.1. *Let (X, Δ) be a three-dimensional projective log pair defined over $\overline{\mathbb{F}}_p$. Assume that one of the following conditions holds*:

(1) (X, Δ) *is log canonical.*

(2) *All the coefficients of Δ are at most one and each irreducible component of* $\mathrm{Supp}(\lfloor \Delta \rfloor)$ *is normal.*

Let L be a nef and big line bundle on X. If $L - (K_X + \Delta)$ is also nef and big, then L is semiample.

The next corollary follows easily from Theorem 1.1.

Corollary 1.2. *Let (X, Δ) be a three-dimensional log canonical projective pair defined over $\overline{\mathbb{F}}_p$.*

(1) *If $K_X + \Delta$ is nef and big, then $K_X + \Delta$ is semiample.*

(2) *If $-(K_X + \Delta)$ is nef and big, then $-(K_X + \Delta)$ is semiample.*

Remark 1.3. Theorem 1.1 does not hold over fields $k \neq \overline{\mathbb{F}}_p$ even in the two-dimensional case (Example 7.2). Corollary 1.2(2) also does not hold over algebraically closed fields $k \neq \overline{\mathbb{F}}_p$ (Example 7.3).

In Example 7.1, we give a counterexample to Theorem 1.1 if one does not impose any conditions on the effective $\mathbb{Q}$-divisor Δ. It is not clear whether the theorem remains true if we only assume that all the coefficients of Δ are at most one.

We also prove the basepoint-free theorem for normal surfaces defined over $\overline{\mathbb{F}}_p$ without assuming bigness:

Theorem 1.4. *Let X be a normal projective surface defined over $\overline{\mathbb{F}}_p$ and let Δ be an effective $\mathbb{Q}$-divisor. Assume that we have a nef line bundle L on X such that $L - (K_X + \Delta)$ is also nef. Then L is semiample.*

Remark 1.5. It is not true in general that nef line bundles on smooth surfaces over $\overline{\mathbb{F}}_p$ are semiample (see Totaro's example [2009]).

Remark 1.6. Theorem 1.1 and Theorem 1.4 hold if we assume that L is only a $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor. Note that if L and $L - (K_X + \Delta)$ are big and nef, then

$$nL - (K_X + \Delta) = (n-1)L + (L - (K_X + \Delta))$$

is also big and nef for any integer $n \geq 1$.

The paper is organized as follows: in Section 2, we review some definitions and facts from minimal model theory and about the conductor scheme. Further, we list some results from [Keel 1999] and show lemmas necessary for the proof of the main theorem. In Section 3, we prove the basepoint-free theorem for surfaces under weaker assumptions (Theorem 1.4). In Section 4, generalizing the proof of [Keel 1999, Theorem 0.5], we reduce Theorem 1.1 to showing that the line bundle $L|_{\operatorname{Supp}\lfloor\Delta\rfloor}$ is semiample (Theorem 4.1). If $\operatorname{Supp}\lfloor\Delta\rfloor$ is irreducible, we know that $L|_{\operatorname{Supp}\lfloor\Delta\rfloor}$ is semiample by Theorem 1.4. The nonirreducible case is treated in Section 5. In order to generalize Theorem 1.4 to the nonirreducible surfaces, we combine ideas from Fujino [2000] and Tanaka [2014], together with special properties of varieties defined over $\overline{\mathbb{F}}_p$, which are proved in Section 2. In Section 6, we complete the proof of Theorem 1.1 and of Corollary 1.2. In Section 7, we give the counterexamples stated in Remark 1.3.

Notation and conventions. • When we work over a normal variety X, we often identify a line bundle L with the divisor corresponding to L. For example, we use the additive notation $L+A$ for a line bundle L and a divisor A.

• Following the notation of [Keel 1999], for a morphism $f\colon X\to Y$, a line bundle L on Y, and a section $s\in H^0(Y,L)$, we denote by $L|_X$ and $s|_X$ the pullbacks f^*L and f^*s, respectively.

• With the same notation as above, we say that a section $t\in H^0(X,L|_X)$ *descends to* Y if there exists a section $s\in H^0(Y,L)$ such that $f^*s=t$.

• Let X be a reduced scheme of finite type over a field, $X=\bigcup X_i$ the decomposition into irreducible components, and $\overline{X_i}\to X_i$ the normalizations. Then we define the *normalization of* X as the composition $\bigsqcup\overline{X_i}\to\bigsqcup X_i\to X$.

• Let X be a scheme and $F\subset X$ a closed subscheme. Let L be a line bundle on X and $s\in H^0(X,L)$ its section. We say that s *is nowhere-vanishing on* F if $s|_{\{x\}}$ is not zero as an element in the one-dimensional vector space $H^0(\{x\},L|_{\{x\}})$ for any closed point $x\in F$.

• We say that a line bundle L on X is *semiample* when the linear system $|mL|$ is basepoint-free for a sufficiently large and divisible positive integer m. When L is semiample, the surjective map $f\colon X\to Y$ defined by $|mL|$ satisfies $f_*\mathcal{O}_X=\mathcal{O}_Y$ for a sufficiently large and divisible positive integer m. We call f *the map associated to* L.

2. Preliminaries

2A. *Log pairs.* A *log pair* (X,Δ) is a normal variety X and an effective $\mathbb{Q}$-divisor Δ such that $K_X+\Delta$ is $\mathbb{Q}$-Cartier.

For a proper birational morphism $f\colon X' \to X$ from a normal variety X', we write

$$K_{X'} + \sum_i a_i E_i = f^*(K_X + \Delta),$$

where the E_i are prime divisors. We say that the pair (X, Δ) is *log canonical* if $a_i \leq 1$ for any proper birational morphism f. Further, we say that the pair (X, Δ) is *Kawamata log terminal* if $a_i < 1$ for any proper birational morphism f.

2B. ***Conductor schemes.*** Let X be a reduced scheme of finite type over a field and $\overline{X} \to X$ its normalization. We identify the sheaf of rings $\mathcal{O}_X$ as a subring of $\mathcal{O}_{\overline{X}}$. Let $\mathcal{I} \subset \mathcal{O}_X$ be the maximal ideal sheaf satisfying $\mathcal{I}\mathcal{O}_{\overline{X}} \subset \mathcal{O}_X$. The *conductor* of X is the subscheme $\mathcal{D} \subset X$ defined by $\mathcal{I}$. By abuse of notation, the subscheme $\mathcal{C} \subset \overline{X}$ defined by $\mathcal{I}\mathcal{O}_{\overline{X}}$ will also be called the conductor.

The notion of conductor is important to descend sections, because of the following remark:

Remark 2.1. Let $\mathcal{C} \subset \overline{X}$, $\mathcal{D} \subset X$ be conductors and let L be a line bundle on X:

$$\begin{tikzcd} \mathcal{C} \arrow[r, hook] \arrow[d] & \overline{X} \arrow[d] \\ \mathcal{D} \arrow[r, hook] & X \end{tikzcd}$$

By definition of the conductor, we have the following exact sequence

$$0 \longrightarrow H^0(X, L) \longrightarrow H^0(\overline{X}, L|_{\overline{X}}) \oplus H^0(\mathcal{D}, L|_{\mathcal{D}}) \longrightarrow H^0(\mathcal{C}, L|_{\mathcal{C}}),$$

where the second map is defined by $t \mapsto (t|_{\overline{X}}, t|_{\mathcal{D}})$ and the third map is defined by $(t, u) \mapsto t|_{\mathcal{C}} - u|_{\mathcal{C}}$. Therefore, a section $s \in H^0(\overline{X}, L|_{\overline{X}})$ descends to X if and only if $s|_{\mathcal{C}}$ descends to $\mathcal{D}$.

2C. ***Adjunction formula.*** Let (X, Δ) be a log pair and S the union of the supports of some of the divisors with coefficient one in Δ. Let $p\colon \overline{S} \to S$ be the normalization of S. Then there exists an effective $\mathbb{Q}$-divisor $\Delta_{\overline{S}}$ on $\overline{S}$ such that

$$K_{\overline{S}} + \Delta_{\overline{S}} = (K_X + \Delta)|_{\overline{S}}$$

holds (see for instance [Kollár 2013, Definition 4.2]).

We denote by C the possibly nonreduced divisor on $\overline{S}$ corresponding to the codimension-one part of $\mathcal{C}$, where $\mathcal{C} \subset \overline{S}$ is the conductor of S.

When X is $\mathbb{Q}$-factorial, it follows that $C \leq \Delta_{\overline{S}}$ by [Keel 1999, Theorem 5.3]. In this paper, we use the following proposition, which only states $\operatorname{Supp}(C) \subset \operatorname{Supp}(\lfloor \Delta_{\overline{S}} \rfloor)$, but is valid even for a non-$\mathbb{Q}$-factorial variety X.

Proposition 2.2. *Let (X, Δ) be a log pair, and let S be the union of the supports of some of the divisors with coefficient one in Δ. Let $p \colon \overline{S} \to S$ be the normalization of S, and let $\Delta_{\overline{S}}$ be an effective $\mathbb{Q}$-divisor on $\overline{S}$ defined by the adjunction as above. Further, we denote by C the (possibly nonreduced) divisor on $\overline{S}$ corresponding to the codimension-one part of $\mathscr{C}$, where $\mathscr{C} \subset \overline{S}$ is the conductor of S. Then the following hold:*

(1) $\mathrm{Supp}(C) \subset \mathrm{Supp}(\lfloor \Delta_{\overline{S}} \rfloor)$.

(2) *Let $D_1, \dots, D_c$ be prime divisors with coefficient greater than or equal to one in Δ, and let $T = \bigcup_{1 \le i \le c} \mathrm{Supp}(D_i)$. Assume that each D_i satisfies $\mathrm{Supp}(D_i) \not\subset S$. Then, the codimension-one part of $p^{-1}(S \cap T)$ is contained in $\mathrm{Supp}(\lfloor \Delta_{\overline{S}} \rfloor)$.*

Proof. First we prove (1). Let $V \subset \overline{S}$ be a codimension-one subvariety such that $V \subset \mathscr{C}$. It is sufficient to show $\mathrm{coeff}_V \Delta_{\overline{S}} \ge 1$. When (X, Δ) is not log canonical at the generic point $\eta_{p(V)}$ of $p(V)$, we have $\mathrm{coeff}_V \Delta_{\overline{S}} > 1$ (see [Kollár 2013, Proposition 4.5(2)]). Hence, we may assume that (X, Δ) is log canonical at $\eta_{p(V)}$. In this case, S has a node at $\eta_{p(V)}$ and $\mathrm{coeff}_V \Delta_{\overline{S}} = 1$ (see the proof of [Kollár 2013, Proposition 4.5(6)]).

Next, we prove (2). Let $V \subset \overline{S}$ be a codimension-one subvariety such that $V \subset p^{-1}(S \cap T)$. It is sufficient to show $\mathrm{coeff}_V \Delta_{\overline{S}} \ge 1$. Since the problem is local around V, we may assume that $p(V) \subset \mathrm{Supp}(D_i)$ for all i. If $\mathrm{coeff}_{D_i} \Delta > 1$ for some i, then (X, Δ) is not log canonical at the generic point $\eta_{p(V)}$ of $p(V)$. In this case, we have $\mathrm{coeff}_V \Delta_{\overline{S}} > 1$ as above. Hence, we may assume that $\mathrm{coeff}_{D_i} \Delta = 1$ for all i. Note that $S \cap T$ is contained in the conductor of the normalization of $S \cup T$. Therefore, we conclude the proof by applying (1) to $S \cup T$. □

2D. ***Some properties of varieties over $\overline{\mathbb{F}}_p$.*** The following fact distinguishes $\overline{\mathbb{F}}_p$ from other fields of positive characteristic. For the proof, see for instance [Keel 1999, Lemma 2.16].

Proposition 2.3. *The Picard scheme $\mathrm{Pic}^0 X$ is a torsion group when X is a projective scheme defined over $\overline{\mathbb{F}}_p$. In particular, any numerically trivial Cartier divisor is $\mathbb{Q}$-linearly trivial.*

We need the following lemmas in Section 5:

Lemma 2.4. *Let X be a proper scheme over $\overline{\mathbb{F}}_p$. Let $s_1, s_2 \in H^0(X, \mathcal{O}_X)$ be sections of the structure sheaf. Assume that s_1 and s_2 are nowhere-vanishing on X. Then there exists $n \ge 1$ such that $s_1^n = s_2^n$ in $H^0(X, \mathcal{O}_X)$.*

Proof. Without loss of generality we may assume that X is connected. Set $A := H^0(X, \mathcal{O}_X)$. It is a finite-dimensional vector space over $\overline{\mathbb{F}}_p$, because X is proper.

Since X is connected, A has a unique maximal ideal $\mathfrak{m}$, and it follows that $A/\mathfrak{m} \cong H^0(X^{\mathrm{red}}, \mathcal{O}_{X^{\mathrm{red}}}) \cong \overline{\mathbb{F}}_p$.

Let a_i be the element of A corresponding to s_i, and $\overline{a_i}$ the image of a_i in $\overline{\mathbb{F}}_p$. Since s_i is nowhere-vanishing on X, the element $\overline{a_i} \in \overline{\mathbb{F}}_p$ is not zero. Hence, there exists $e \geq 1$ for which $\overline{a_1}^{p^e-1} = \overline{a_2}^{p^e-1} = 1$. Take $r \geq 1$ such that $\mathfrak{m}^{p^r} = 0$. Then we have

$$a_1^{p^r(p^e-1)} - a_2^{p^r(p^e-1)} = \big(a_1^{p^e-1} - a_2^{p^e-1}\big)^{p^r} \in \mathfrak{m}^{p^r} = 0.$$

Therefore, it is sufficient to set $n = p^r(p^e-1)$. □

Lemma 2.5. *Let X be a one-dimensional reduced scheme of finite type over $\overline{\mathbb{F}}_p$, L a line bundle on X, and $p \colon \overline{X} \to X$ the normalization of X. Let $\mathscr{C} \subset \overline{X}$ be the conductor of X, and $s \in H^0(\overline{X}, L|_{\overline{X}})$ a section nowhere-vanishing on $\mathscr{C}$. Then s^n descends to X for some $n \geq 1$.*

Proof. Let $\mathscr{D} \subset X$ be the conductor. Note that $\mathscr{C}$ and $\mathscr{D}$ are either empty or have dimension zero. By Remark 2.1, it is sufficient to prove that $s^n|_{\mathscr{C}}$ descends to $\mathscr{D}$ for some $n \geq 1$. Let $t \in H^0(\mathscr{D}, L|_{\mathscr{D}})$ be a section nowhere-vanishing on $\mathscr{D}$. Then $t|_{\mathscr{C}}$ is nowhere-vanishing on $\mathscr{C}$. Any line bundle is trivial on a zero-dimensional scheme, and so, by Lemma 2.4, we get $s^n|_{\mathscr{C}} = t^n|_{\mathscr{C}}$ for some $n \geq 1$. In particular, $s^n|_{\mathscr{C}}$ descends to $\mathscr{D}$. □

Lemma 2.6. *Let C be a smooth proper connected curve over $\overline{\mathbb{F}}_p$. Then a finitely generated subgroup of* $\mathrm{Aut}(C)$ *is finite.*

Proof. If $g(C) \geq 2$, then $\mathrm{Aut}(C)$ is finite and the statement is trivial. If $C = \mathbb{P}^1$, then $\mathrm{Aut}(C) \cong \mathrm{PGL}(2, \overline{\mathbb{F}}_p)$. A finitely generated subgroup G of $\mathrm{PGL}(2, \overline{\mathbb{F}}_p)$ is always finite, because G is contained in $\mathrm{PGL}(2, \mathbb{F}_{p^e})$ for some $e \geq 1$. If C is an elliptic curve, then we get $\mathrm{Aut}(C) \cong T \rtimes F$, where T is the group of translations and F is a finite group (see for instance [Silverman 2009, Section X.5]). Note that each element of T has finite order, because C is defined over $\overline{\mathbb{F}}_p$. Hence, a finitely generated subgroup of the abelian group T is always finite, and so a finitely generated subgroup of $\mathrm{Aut}(C)$ is also finite.

For completeness, we note a general fact in group theory: any finitely generated subgroup of $G_1 \rtimes G_2$ is finite, if we assume that any finitely generated subgroup of G_i is finite for each i. □

2E. ***Keel's theorems.*** The following theorem is crucial in reducing problems from threefolds to surfaces:

Theorem 2.7 [Keel 1999, Proposition 1.6]. *Let X be a projective scheme over a field of positive characteristic. Let L be a nef line bundle on X, and let E be an effective Cartier divisor on X such that $L - E$ is ample. Then L is semiample if and only if $L|_{E_{\mathrm{red}}}$ is semiample.*

We note that Cascini, M[c]Kernan, and Mustaţă [Cascini et al. 2014, Theorem 3.2] gave a different proof of Theorem 2.7.

Theorem 2.8 [Artin 1962, Theorem 2.9; Keel 1999, Corollary 0.3]. *Let X be a projective surface over $\overline{\mathbb{F}}_p$, and let L be a nef and big line bundle on X. Then L is semiample.*

Proof. Since by Proposition 2.3 nef line bundles on curves over $\overline{\mathbb{F}}_p$ are semiample, the claim follows from Theorem 2.7. □

We say that a map $f : X \to Y$ is a *finite universal homeomorphism* if it is a finite homeomorphism under any base change. In this case, we have a correspondence, up to taking powers, between the set of sections of a line bundle L on Y and the set of sections of $L|_X$.

Theorem 2.9 [Keel 1999, Lemma 1.4]. *Let $f : X \to Y$ be a finite universal homeomorphism between varieties defined over a field of characteristic $p > 0$, and let L be a line bundle on Y. Then the following hold*:

(1) *For $s \in H^0(X, L|_X)$, the section $s^{p^e} \in H^0(X, L^{\otimes p^e}|_X)$ descends to Y for a sufficiently large integer $e \geq 1$.*

(2) *If $t \in H^0(Y, L)$ satisfies $t|_X = 0$, then $t^{p^e} = 0$ holds for a sufficiently large integer $e \geq 1$.*

In this paper, we frequently use the following theorems:

Theorem 2.10 [Keel 1999, Corollary 2.12]. *Let $X = X_1 \cup X_2$ be a projective scheme over $\overline{\mathbb{F}}_p$, where the X_i are closed subsets. Let L be a nef line bundle on X such that the $L|_{X_i}$ are semiample. Let $g_i : X_i \to Z_i$ be the map associated to $L|_{X_i}$. Assume that all but finitely many fibers of $g_2|_{X_1 \cap X_2}$ are geometrically connected. Then L is semiample.*

Theorem 2.11 [Keel 1999, Corollary 2.14]. *Let X be a reduced projective scheme over $\overline{\mathbb{F}}_p$. Let $p : \overline{X} \to X$ be the normalization of X. Let $D \subset X$ and $C \subset \overline{X}$ be the reductions of the conductors. Let L be a nef line bundle on X such that $L|_{\overline{X}}$ and $L|_D$ are semiample. Let $g : \overline{X} \to Z$ be the map associated to $L|_{\overline{X}}$. Assume that all but finitely many fibers of $g|_C$ are geometrically connected. Then L is semiample.*

3. Basepoint-free theorem for normal surfaces

In this section, we prove Theorem 1.4. The key tool is the following theorem of Tanaka. We say that a $\mathbb{Q}$-divisor B on a variety X is $\mathbb{Q}$-*effective* if $h^0(X, mB) \neq 0$ for some $m \geq 1$. Note that a normal surface over $\overline{\mathbb{F}}_p$ is always $\mathbb{Q}$-factorial (see [Tanaka 2012, Theorem 11.1]).

Theorem 3.1 [Tanaka 2012, Theorem 12.6]. *Let X be a projective normal surface over $\overline{\mathbb{F}}_p$ and let D be a nef divisor. If $qD - K_X$ is $\mathbb{Q}$-effective for some positive rational number $q \in \mathbb{Q}$, then D is semiample.*

We will use the following proposition to reduce the case of hyperelliptic surfaces to abelian surfaces:

Proposition 3.2. *Let $p\colon Y \to X$ be a proper surjection between varieties defined over an algebraically closed field, and let L be a line bundle on X. Assume that X is normal. Then L is semiample if and only if $p^*(L)$ is semiample.*

Proof. See for instance the proof of [Keel 1999, Lemma 2.10]. □

Proof of Theorem 1.4. Recall that we have the nef line bundle L and the $\mathbb{Q}$-divisor $D := L - (K_X + \Delta)$ on the normal surface X over $\overline{\mathbb{F}}_p$.

Claim 3.3. *We can assume that X is smooth.*

Proof. Let $f: Y \to X$ be the minimal resolution of singularities of X. Define Δ_Y so that $K_Y + \Delta_Y = f^*(K_X + \Delta)$. The divisor Δ_Y is an effective $\mathbb{Q}$-divisor by the negativity lemma (see [Kollár and Mori 1998, Corollary 4.3]). Note that f^*L and $f^*D = f^*L - (K_Y + \Delta_Y)$ are nef. By Proposition 3.2 we know that L is semiample if and only if f^*L is semiample. Thus, by replacing X by Y, we may assume that the surface is smooth. □

We extensively use the following lemma:

Lemma 3.4. *If D is $\mathbb{Q}$-effective, then L is semiample.*

Proof. Since D is $\mathbb{Q}$-effective, $L - K_X = D + \Delta$ is also $\mathbb{Q}$-effective, and so L is semiample by Theorem 3.1. □

Claim 3.5. *We can assume that all the following statements are true.*

(1) $L \not\equiv 0$ *and* $D \not\equiv 0$,	(2) $L^2 = 0$,	(3) $D^2 = 0$,
(4) $L \cdot \Delta = 0$,	(5) $L \cdot K_X = 0$,	(6) $(K_X + \Delta) \cdot \Delta = 0$,
(7) $(K_X + \Delta) \cdot K_X = 0$,	(8) $\chi(\mathcal{O}_X) \leq 0$.	

Proof. If $L \equiv 0$, then $L \sim_{\mathbb{Q}} \mathcal{O}_X$ by Proposition 2.3, so L is semiample. Thus, we may assume that $L \not\equiv 0$. Analogously, we may assume that $D \not\equiv 0$.

As L and D are nef, we get $L^2 \geq 0$ and $D^2 \geq 0$. If $L^2 > 0$, then, by Theorem 2.8, the line bundle L is semiample. Thus, we may assume that $L^2 = 0$. If $D^2 > 0$, then D is big, and so $\mathbb{Q}$-effective. In this case L is semiample by Lemma 3.4. Hence, we may assume $D^2 = 0$.

Since $L \not\equiv 0$, we know that there exists a curve C on X satisfying $L \cdot C > 0$. Take an ample divisor A such that $A - C$ is effective. Then $L \cdot A = L \cdot C + L \cdot (A - C) > 0$.

If m is sufficiently large that it satisfies $(K_X - mL) \cdot A < 0$, then $h^2(X, mL) = h^0(X, K_X - mL) = 0$. The Riemann–Roch theorem gives

$$\begin{aligned} h^0(X, mL) &= h^1(X, mL) + \tfrac{1}{2} mL \cdot (mL - K_X) + \chi(\mathbb{O}_X) \\ &= h^1(X, mL) - \tfrac{1}{2} mL \cdot K_X + \chi(\mathbb{O}_X). \end{aligned}$$

As L and D are nef, it follows that

$$0 \le L \cdot D = -L \cdot K_X - L \cdot \Delta.$$

Since Δ is effective and L is nef, we find $0 \le L \cdot D \le -L \cdot K_X$. If $-L \cdot K_X > 0$, then $\kappa(X, L) = 1$ by the calculation of $h^0(X, mL)$ above. A nef line bundle L with $\kappa(X, L) = 1$ is always semiample (see for instance [Fong and McKernan 1992, Theorem 11.3.1]). Thus, we may assume that $L \cdot \Delta = 0$ and $L \cdot K_X = 0$.

As above, $h^2(X, mD) = 0$ holds for sufficiently large m, and so the Riemann–Roch theorem gives

$$\begin{aligned} h^0(X, mD) &= h^1(X, mD) - \tfrac{1}{2} mD \cdot K_X + \chi(\mathbb{O}_X) \\ &= h^1(X, mD) + \tfrac{1}{2} mD \cdot (D - L + \Delta) + \chi(\mathbb{O}_X) \\ &= h^1(X, mD) + \tfrac{1}{2} mD \cdot \Delta + \chi(\mathbb{O}_X) \\ &= h^1(X, mD) - \tfrac{1}{2} m(K_X + \Delta) \cdot \Delta + \chi(\mathbb{O}_X). \end{aligned}$$

If $-(K_X + \Delta) \cdot \Delta > 0$, then D is $\mathbb{Q}$-effective and by Lemma 3.4 the line bundle L is semiample. Since $0 \le D \cdot \Delta = -(K_X + \Delta) \cdot \Delta$ holds by the nefness of D, we may assume $(K_X + \Delta) \cdot \Delta = 0$. Given $D^2 = L^2 = D \cdot L = 0$, it follows that $(K_X + \Delta) \cdot K_X = 0$.

By the Riemann–Roch theorem, we get $h^0(X, mD) = h^1(X, mD) + \chi(\mathbb{O}_X)$. If $\chi(\mathbb{O}_X) > 0$, then D is $\mathbb{Q}$-effective and by Lemma 3.4 the line bundle L is semiample. Hence, we may assume that $\chi(\mathbb{O}_X) \le 0$. □

We divide the proof into cases depending on the Kodaira dimension.

Case 1: Assume $\kappa(X) \ge 0$.

Claim 3.6. *We may assume that K_X is nef.*

Proof. Let $\pi : X \to X_{\min}$ be the minimal model of X. By $\pi_* L$ we denote the pushforward of L as a divisor.

By the assumption $\kappa(X) \ge 0$, we have that K_X is $\mathbb{Q}$-linearly equivalent to an effective $\mathbb{Q}$-divisor containing every π-exceptional curve in its support. Since $L \cdot K_X = 0$ and L is nef, it follows that $L \cdot E = 0$ for every π-exceptional curve E. Hence, we get $L = \pi^* \pi_* L$ by the negativity of the intersection form on the exceptional locus (see [Kollár and Mori 1998, Lemma 3.40]).

Since $L = \pi^*\pi_* L$, it is sufficient to show the semiampleness of $\pi_* L$. Note that $\pi_* L$ and $\pi_* D$ are nef, because L and D are nef. Further, we have $\pi_* D = \pi_* L - (K_{X_{\min}} + \pi_*\Delta)$. Therefore, we can reduce the problem to the case of the minimal model $X_{\min}$. □

In what follows, we assume that X is minimal. We use the classification of minimal surfaces in positive characteristic (see for instance [Mumford 1969; Bombieri and Mumford 1977; 1976; Liedtke 2013]).

Case 1.1: Assume $\kappa(X) = 2$.

We can write $K_X = A + E$ for an ample $\mathbb{Q}$-divisor A and an effective $\mathbb{Q}$-divisor E, because K_X is big. Since L, D are nef and $L \cdot K_X = D \cdot K_X = 0$, it follows that $L \cdot A = D \cdot A = 0$. Thus, $(L - D) \cdot A = (K_X + \Delta) \cdot A = 0$. We get a contradiction

$$0 < A^2 \leq (K_X + \Delta) \cdot A = 0.$$

Hence, there are no line bundles L satisfying the assumptions in Claim 3.5.

Case 1.2: Assume $\kappa(X) = 1$.

In our case, K_X is semiample and it gives an elliptic or quasielliptic fibration $f : X \to B$. Let F be its general fiber. Then $K_X \equiv aF$ holds for some positive rational number a.

Since $D \cdot K_X = 0$, it follows that $D \cdot F = 0$. Therefore, D is f-numerically trivial by the nefness of D. Since D is nef and f-numerically trivial, it satisfies $D \equiv bF$ for some $b \geq 0$, by Lemma 3.7. Hence, D is $\mathbb{Q}$-effective by Proposition 2.3. Therefore, L is semiample by Lemma 3.4.

Lemma 3.7. *Let $f : X \to B$ be a surjective morphism satisfying $f_*(\mathcal{O}_X) = \mathcal{O}_B$ from a smooth projective surface X to a smooth projective curve B. Suppose that L is an f-numerically trivial nef $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor. Then $L \equiv bF$ for some $b \geq 0$, where F denotes a general fiber of f.*

Proof. See for instance [Lehmann 2012, Lemma 2.4]. □

Case 1.3: Assume $\kappa(X) = 0$.

By the classification of minimal surfaces, there are five possibilities: a K3 surface, an Enriques surface, an abelian surface, a hyperelliptic surface, or a quasihyperelliptic surface.

If X is a K3 surface or an Enriques surface, then $\chi(\mathcal{O}_X) = 2$ or $\chi(\mathcal{O}_X) = 1$, respectively, which contradicts Claim 3.5.

If X is an abelian surface, then every nef divisor is numerically equivalent to a semiample divisor (see Proposition 3.10). Therefore, L is semiample by Proposition 2.3.

If X is a hyperelliptic surface, then X is a finite quotient of an abelian surface by a finite group. Therefore, we have a surjective morphism $A \to X$ from an abelian surface A. Since $L|_A$ is a nef line bundle on an abelian surface, it is semiample (see Proposition 3.10). Hence, L is also semiample by Proposition 3.2.

If X is a quasihyperelliptic surface, then X can be written as a finite quotient $E \times C \to X$, where E is an elliptic curve and C is a rational curve with a cusp. Therefore, we have a surjective morphism $X' := E \times \mathbb{P}^1 \to X$. Any divisor on X' is numerically equivalent to $aF_1 + bF_2$ with $a, b \in \mathbb{Q}$, where F_1 is the fiber class of $X' \to E$ and F_2 is the fiber class of $X' \to \mathbb{P}^1$. Hence, any nef divisor on X' is numerically equivalent to a semiample divisor. Thus, we can conclude that L is semiample by Proposition 2.3 and Proposition 3.2.

Case 2: Assume $\kappa(X) = -\infty$.

Since $\chi(\mathcal{O}_X) \leq 0$, the surface X is irrational. Thus, we can assume that $f : X \to B$ is a birationally ruled surface, where B is a curve with $g(B) \geq 1$.

We need the following lemma, which can be found in the proof of [Tanaka 2012, Theorem 12.4].

Lemma 3.8. *Let C be an f-horizontal curve on X such that $D \cdot C = 0$. Then D is $\mathbb{Q}$-effective.*

Proof. Since C is a horizontal curve, it holds that $g(B) \leq h^1(C, \mathcal{O}_C)$. By the Riemann–Roch theorem, we get

$$h^0(X, mD) = h^1(X, mD) + \chi(\mathcal{O}_X) = h^1(X, mD) + 1 - g(B),$$

so it is sufficient to show $h^1(X, mD) \geq h^1(C, \mathcal{O}_C)$ for some $m > 0$.

Since $D \cdot C = 0$, we have $D|_C \equiv 0$. Hence, by Proposition 2.3 we can conclude that $mD|_C$ is trivial for a sufficiently divisible $m > 0$. Therefore, we get an exact sequence

$$0 \longrightarrow \mathcal{O}_X(mD - C) \longrightarrow \mathcal{O}_X(mD) \longrightarrow \mathcal{O}_C \longrightarrow 0.$$

By the same reason as before, $h^2(X, mD - C) = 0$ holds for sufficiently large m. Hence, we get $h^1(X, mD) \geq h^1(C, \mathcal{O}_C)$. □

For any irreducible component C of Δ, it follows that $D \cdot C = 0$, because D is nef and $D \cdot \Delta = 0$. In particular, if Δ has an f-horizontal component, then the lemma above implies that D is $\mathbb{Q}$-effective, and hence L is semiample by Lemma 3.4. Thus, in what follows, we may assume that Δ has only f-vertical components.

Claim 3.9. *Under these assumptions, it follows that $\Delta = 0$, $g(B) = 1$, and X is a minimal ruled surface.*

Proof. Let $\pi : X \to X_{\min}$ be a minimal model of X. We have $K_X \sim \pi^* K_{X_{\min}} + E$, where E is an exceptional divisor. We refer the reader to [Hartshorne 1977, Chapter V, Section 2] for properties of ruled surfaces. It holds that

$$K_{X_{\min}} \equiv -2C_0 + (2g(B) - 2 - e)F$$

for C_0 a normalized section, $e = -C_0^2$, and F a general fiber of $X_{\min} \to B$. Note that $K^2_{X_{\min}} = 8(1 - g(B))$.

Since $(K_X + \Delta) \cdot \Delta = 0$ and $(K_X + \Delta) \cdot K_X = 0$, we get

$$\Delta^2 = -K_X \cdot \Delta = K_X^2.$$

As Δ has only f-vertical components, we have $\pi^* F \cdot \Delta = 0$, and so

$$0 = (K_X + \Delta) \cdot \Delta = -2\pi^* C_0 \cdot \Delta + (E + \Delta) \cdot \Delta.$$

Since $\pi^* C_0 \cdot \Delta \geq 0$, it follows that $E \cdot \Delta \geq -\Delta^2$. Therefore,

$$(E+\Delta)^2 = E^2 + 2E \cdot \Delta + \Delta^2 \geq E^2 - \Delta^2 = E^2 - K_X^2 = -K^2_{X_{\min}} = 8(g(B) - 1) \geq 0.$$

By the Zariski lemma [Liu 2002, Section 9, Theorem 1.23], the intersection form on f-vertical fibers is seminegative-definite with one-dimensional radical equal to the span of a general fiber, so $(E + \Delta)^2 = 0$ and $E + \Delta \equiv \pi^* pF$ for some $p \in \mathbb{Q}$.

Since all the inequalities must be equalities, it follows that $E \cdot \Delta = -\Delta^2$ and $g(B) = 1$. Furthermore, we have $2\pi^* C_0 \cdot \Delta = (E + \Delta) \cdot \Delta$, and thus

$$0 = \pi^* C_0 \cdot \Delta = \pi^* C_0 \cdot (E + \Delta) = p.$$

This implies that $E + \Delta = 0$. Since Δ and E are both effective divisors, we get $\Delta = 0$ and $E = 0$. Hence, X is minimal. □

By this claim, we can assume that X is a minimal ruled surface over an elliptic curve. In this case, it is well-known that $\mathrm{NEF}(X) \subset \mathrm{NE}(X)$ holds (see Proposition 3.13). We can conclude that the nef divisor D is $\mathbb{Q}$-effective and L is semiample by Lemma 3.4. □

For completeness, we prove two propositions which were used in the above proof:

Proposition 3.10. *Let A be an abelian variety defined over an algebraically closed field. Then any nef line bundle on A is numerically equivalent to a semiample line bundle.*

Remark 3.11. Note that any effective divisor on an abelian variety is always semiample (see the proof of Application 1((i)⇒(iii)) in [Mumford 2008, Section 6]).

Proof. Let L be a nef line bundle on A. Define $K(L)$ to be the maximal subscheme of A such that

$$(m^* L - p_1^* L - p_2^* L)|_{K(L) \times A} = \mathcal{O}_{K(L) \times A}$$

as in [Mumford 2008, Section 13], where $m\colon A \times A \to A$ is the multiplication map and p_1 and p_2 are the first and second projections.

By the above remark, we may assume that L is not big, so that $L^g = 0$, where $g = \dim A$. By the Riemann–Roch theorem [loc. cit., Section 16], we have $\chi(L) = 0$. Hence, it follows that $\dim K(L) > 0$ by the vanishing theorem [loc. cit., Section 16].

Set $X := K(L)^0_{\mathrm{red}}$. This is a subabelian variety of A. Thus, there exists a subabelian variety $Y \subset A$ such that the morphism $m\colon X \times Y \to A$, $(x, y) \mapsto x + y$ defined by the group law on A is an isogeny [loc. cit., Section 19, Theorem 1]). Note that $L|_X \in \mathrm{Pic}^0(X)$, because it is invariant under translations by any element of X (see Remark 3.12).

First, we prove $m^*L \equiv p_Y^*(L|_Y)$, where $p_Y\colon X \times Y \to Y$ is the second projection. By definition of $K(L)$, we get $m^*L = p_X^*(L|_X) + p_Y^*(L|_Y)$. Since $L|_X \in \mathrm{Pic}^0(X)$, we have $L|_X \equiv 0$, which proves $m^*L \equiv p_Y^*(L|_Y)$.

Since $\dim Y < \dim A$, we may assume that $L|_Y$ is numerically equivalent to a semiample line bundle by induction on $\dim A$. By Proposition 3.2, in order to complete the proof, it is sufficient to show that $p_Y^*(L|_Y)$ descends to A. This is true, because $\mathrm{Pic}^0(A) \to \mathrm{Pic}^0(X \times Y)$ is surjective [loc. cit., Section 15, Theorem 1]). $\square$

Remark 3.12. Mumford [2008, Section 8] defines $\mathrm{Pic}^0(X)$, for an abelian variety X, to be the subgroup of $\mathrm{Pic}(X)$ consisting of line bundles invariant under translations by any element of X. The existence of the dual abelian variety and the Poincaré line bundle [loc. cit., Section 13] shows that this definition is equivalent to the standard definition of $\mathrm{Pic}^0(X)$ as the identity component of the Picard functor.

Proposition 3.13. *Let X be a minimal ruled surface over an elliptic curve B. Then it follows that* $\mathrm{NEF}(X) \subset \mathrm{NE}(X)$.

Proof. Let $C_0 \subset X$ be a normalized section and F a fiber of $X \to B$. Set $e := -C_0^2$. When $e \geq 0$, we get

$$\mathrm{NEF}(X) = \mathrm{Cone}(F, C_0 + eF),$$

and so nef line bundles are effective.

In what follows, we may assume $e = -1$ by [loc. cit., Chapter V, Theorem 2.15]. We know that

$$\mathrm{NEF}(X) = \overline{\mathrm{NE}}(X) = \mathrm{Cone}(F, 2C_0 - F)$$

by [loc. cit., Chapter V, Proposition 2.21]. Further, there exists a rank-two indecomposable vector bundle E of degree one on C such that $X \cong \mathbb{P}_C(E)$ holds. We denote the projection by $p\colon \mathbb{P}_C(E) \to C$. It is sufficient to show $H^0(X, \mathcal{O}_X(2C_0 - p^*Q)) \neq 0$ for some point $Q \in C$, because then $\overline{\mathrm{NE}}(X) = \mathrm{NE}(X)$. Note that

$$H^0(X, \mathcal{O}_X(2C_0 - p^*Q)) \cong H^0(C, S^2(E) \otimes \mathcal{O}_C(-Q))$$

and $S^2(E)$ has both rank and degree equal to three (see [loc. cit., Chapter II, Example 5.16] and the proof of [loc. cit., Chapter V, Theorem 2.15]). When $S^2(E)$ is indecomposable, we can complete the proof by using the following proposition:

Proposition 3.14 [Atiyah 1957, Lemma 11]. *Let F be an indecomposable vector bundle of rank r and degree d on an elliptic curve. If $r = d$, then F contains a degree-one line bundle as a subbundle.*

When $S^2(E)$ is decomposable, it can be written as $S^2(E) \cong E_1 \oplus E_2$, where E_1 is a line bundle and E_2 is a vector bundle of rank two. If $\deg E_1 \geq 1$, then

$$H^0(C, S^2(E) \otimes \mathcal{O}_C(-Q)) \supset H^0(C, E_1 \otimes \mathcal{O}_C(-Q)) \neq 0$$

for some point $Q \in C$, which finishes the proof in this case. If $\deg E_1 < 1$, then $\deg E_2 \geq 3$, and so $\deg(E_2 \otimes \mathcal{O}_C(-Q)) \geq 1$ for any point $Q \in C$. Therefore,

$$H^0(C, S^2(E) \otimes \mathcal{O}_C(-Q)) \supset H^0(C, E_2 \otimes \mathcal{O}_C(-Q)) \neq 0$$

by the Riemann–Roch theorem. □

4. Reduction to surfaces

The first step in the proof of Theorem 1.1 is to reduce the problem to the case of surfaces.

Theorem 4.1. *Let (X, Δ) be a three-dimensional projective log pair defined over $\overline{\mathbb{F}}_p$, and L a line bundle on X. If we assume that*

- *L and $L - (K_X + \Delta)$ are nef and big,*
- *$L|_{\mathrm{Supp}\lfloor\Delta\rfloor}$ is semiample,*

then L is semiample.

Here, we adopt the convention that, when $\lfloor\Delta\rfloor = 0$, then $L|_{\mathrm{Supp}\lfloor\Delta\rfloor}$ is automatically semiample.

Remark 4.2. Under the assumption $\lfloor\Delta\rfloor = 0$, Theorem 4.1 was proved by Keel [1999, Theorem 0.5].

Proof of Theorem 1.1. Set $S := \lfloor\Delta\rfloor$. Since L is a big line bundle, we can decompose it as $L \sim_{\mathbb{Q}} A + E$, where A is an ample and E is an effective $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor. By Theorem 2.7, it is enough to show that $L|_{E_{\mathrm{red}}}$ is semiample.

We write $E_{\mathrm{red}} = T + \sum_{i=1}^{m} E_i$, where $\mathrm{Supp}(T) \subset \mathrm{Supp}(S)$ and the E_i are prime divisors not contained in $\mathrm{Supp}(S)$. Define $\lambda_i \in \mathbb{Q}$ so that $\Delta + \lambda_i E$ contains E_i with coefficient one. Then, by definition of λ_i, there exists an effective $\mathbb{Q}$-divisor Γ_i such that

$$\Delta + \lambda_i E = E_i + \Gamma_i$$

and $E_i \not\subset \mathrm{Supp}(\Gamma_i)$. Since $E_i \not\subset \mathrm{Supp}(S)$, it follows that $\lambda_i > 0$. By rearranging indices, we may assume without loss of generality that

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m,$$

so we have

$$T + \sum_{1\leq j\leq i-1} E_j \leq \Gamma_i$$

for each i.

We define $U_0 := \mathrm{Supp}(T)$ and $U_i := U_{i-1} \cup E_i$ for $i > 0$. Recall that it is sufficient to show that L restricted to $U_m = \mathrm{Supp}(E_{\mathrm{red}})$ is semiample. We prove it by induction on i.

Observe that $L|_{U_0}$ is semiample, because $U_0 = \mathrm{Supp}(T) \subset \mathrm{Supp}(S)$ and $L|_S$ is semiample by hypothesis. Let us assume that $L|_{U_{i-1}}$ is semiample. In order to prove the semiampleness of $L|_{U_i}$, we first prove the semiampleness of $L|_{E_i}$.

We consider the normalization $p_i\colon \overline{E_i} \to E_i$. By adjunction (see Section 2C), there exists an effective $\mathbb{Q}$-divisor $\Delta_{\overline{E_i}}$ such that

$$(K_X + E_i + \Gamma_i)|_{\overline{E_i}} \sim K_{\overline{E_i}} + \Delta_{\overline{E_i}}.$$

Lemma 4.3. *$L|_{\overline{E_i}}$ is semiample.*

Proof. We define auxiliary divisors D_i by

$$D_i := (1+\lambda_i)L - (K_X + \Delta + \lambda_i E).$$

Observe that

$$D_i = L - (K_X + \Delta) + \lambda_i(L - E) \sim_{\mathbb{Q}} (L - (K_X + \Delta)) + \lambda_i A,$$

and so D_i is ample, because $L - (K_X + \Delta)$ is nef and $\lambda_i A$ is ample. Hence,

$$D_i|_{\overline{E_i}} = (1+\lambda_i)L|_{\overline{E_i}} - (K_{\overline{E_i}} + \Delta_{\overline{E_i}})$$

is nef. Since $(1+\lambda_i)L|_{\overline{E_i}}$ is also nef, the semiampleness of $L|_{\overline{E_i}}$ follows from Theorem 1.4 and Remark 1.6. □

Assume $\kappa(L|_{\overline{E_i}})$ is equal to 0 or 2. Then the assumptions of Theorem 2.11 are satisfied, and so $L|_{E_i}$ is semiample. Using Theorem 2.10 for $X_1 = U_{i-1}$ and $X_2 = E_i$, we get that $L|_{U_i}$ is semiample.

In what follows, we assume $\kappa(L|_{\overline{E_i}}) = 1$.

Lemma 4.4. *Let $\pi_i\colon \overline{E_i} \to Z_i$ be the map associated to the semiample line bundle $L|_{\overline{E_i}}$, and let F be a general fiber of π_i. Further, let $C_i \subset \overline{E_i}$ be the reduction of the conductor of the normalization $p_i\colon \overline{E_i} \to E_i$. Then F and C_i intersect in at most one point.*

Proof. Let D_i be the $\mathbb{Q}$-divisor on $\overline{E_i}$ as in the proof of Lemma 4.3. Then, D_i is ample, so we have $F \cdot D_i|_{\overline{E_i}} > 0$. Since $F \cdot L|_{\overline{E_i}} = 0$, we get

$$F \cdot K_{\overline{E_i}} + F \cdot \Delta_{\overline{E_i}} < 0.$$

Hence

$$F \cdot \Delta_{\overline{E_i}} < -F \cdot K_{\overline{E_i}} = 2 - 2h^1(F, \mathcal{O}_F) \leq 2.$$

By the adjunction formula (Proposition 2.2), the one-dimensional part of C_i is contained in $\operatorname{Supp}(\lfloor \Delta_{\overline{E_i}} \rfloor)$. Hence, we get $\#(F \cap C_i) \leq F \cdot \Delta_{\overline{E_i}} < 2$. □

By this lemma, the assumptions of Theorem 2.11 are satisfied, and so $L|_{E_i}$ is semiample. Let $\rho_i \colon E_i \to Z_i'$ be the map associated to $L|_{E_i}$, and let G be a general fiber of ρ_i. Since π_i is the Stein factorization of $\rho_i \circ p_i$, there exists a finite map $Z_i \to Z_i'$ such that the following diagram commutes [Keel 1999, Definition-Lemma 1.0(4)]:

$$\begin{tikzcd} \overline{E_i} \arrow[r, "p_i"] \arrow[d, "\pi_i"'] & E_i \arrow[d, "\rho_i"] \\ Z_i \arrow[r] & Z_i' \end{tikzcd}$$

We want to apply Theorem 2.10 to $X_1 = U_{i-1}$ and $X_2 = E_i$ to show that $L|_{U_i}$ is semiample. It is sufficient to prove that G intersects $U_{i-1} \cap E_i$ in at most one point.

Recall that

$$T + \sum_{1 \leq j \leq i-1} E_j \leq \Gamma_i, \quad U_{i-1} = \operatorname{Supp}\Big(T + \sum_{1 \leq j \leq i-1} E_j\Big).$$

Hence, the one-dimensional part of $p_i^{-1}(U_{i-1} \cap E_i)$ is contained in $\operatorname{Supp}(\lfloor \Delta_{\overline{E_i}} \rfloor)$ by the adjunction formula (Proposition 2.2). By the proof of Lemma 4.4, we can conclude

$$\begin{aligned} \#((U_{i-1} \cap E_i) \cap G) &= \#\big(p_i(p_i^{-1}(U_{i-1} \cap E_i) \cap F)\big) \\ &\leq \#(p_i^{-1}(U_{i-1} \cap E_i) \cap F) \\ &\leq F \cdot \Delta_{\overline{E_i}} < 2, \end{aligned}$$

which completes the proof. □

5. Semiampleness on nonirreducible surfaces

In this section, we prove Theorem 5.2. Before stating it, we need to introduce some notation. Let S be a pure two-dimensional reduced projective scheme over $\overline{\mathbb{F}}_p$, and let $S = \bigcup_{i=1}^n S_i$ be its irreducible decomposition and $\overline{S} \to S$ its normalization. Let $\mathcal{D} \subset S$ and $\mathcal{C} \subset \overline{S}$ be the conductors of S. Let $\overline{C} \xrightarrow{\text{norm.}} \mathcal{C}_{\text{red}} \longrightarrow \mathcal{C}$ and $\overline{D} \xrightarrow{\text{norm.}} \mathcal{D}_{\text{red}} \longrightarrow \mathcal{D}$

be the compositions of the reduction map and the normalization. Then we have a natural morphism $f: \overline{C} \to \overline{D}$ such that the following diagram commutes:

$$\begin{tikzcd}
\overline{C} \arrow[r, "\text{normalization}"] \arrow[d, "f"'] & \mathscr{C}_{\mathrm{red}} \arrow[r] & \mathscr{C} \arrow[r, hook] \arrow[d] & \overline{S} \arrow[d] \\
\overline{D} \arrow[r, "\text{normalization}"] & \mathscr{D}_{\mathrm{red}} \arrow[r] & \mathscr{D} \arrow[r, hook] & S
\end{tikzcd}$$

Consider the one-dimensional part $\overline{C}^{(1)}$ of $\overline{C}$ and the restriction $f: \overline{C}^{(1)} \to \overline{D}$. We say that S *satisfies the condition* $(\star)$ when the restriction of f to any one-dimensional connected component of $\overline{C}$ is an isomorphism onto its image. Further, we say that S *satisfies the condition* $(\star\star)$ when any fiber of the restriction $f: \overline{C}^{(1)} \to \overline{D}$ has length at most two.

Remark 5.1. If each S_i is normal, then S satisfies the condition $(\star)$. If S is regular or nodal in codimension one, then S satisfies the condition $(\star\star)$. See the proof of Theorem 1.1.

Theorem 5.2. *Let S be a pure two-dimensional reduced projective scheme over $\overline{\mathbb{F}}_p$, and let $S = \bigcup_{i=1}^n S_i$ be its irreducible decomposition. Let L be a nef line bundle on S. Suppose that S satisfies the condition $(\star)$ or $(\star\star)$ defined above and that there exists an effective $\mathbb{Q}$-divisor $\Delta_{\overline{S}}$ on the normalization $\overline{S}$ of S such that*:

- $L|_{\overline{S}} - (K_{\overline{S}} + \Delta_{\overline{S}})$ *is nef.*
- $\mathrm{Supp}(\mathscr{C}^{(1)})$ *is contained in* $\mathrm{Supp}(\lfloor \Delta_{\overline{S}} \rfloor)$, *where* $\mathscr{C}^{(1)} \subset \overline{S}$ *is the one-dimensional part of the conductor scheme of the normalization of S.*

Then L is semiample.

Proof. We use the same notation as above. Let $\nu: \overline{S} := \bigsqcup \overline{S_i} \to S$ be the normalization of S. Set $\Delta_{\overline{S_i}} := \Delta_{\overline{S}}|_{\overline{S_i}}$. We know that the $L|_{\overline{S_i}}$ are semiample from Theorem 1.4. Let $g_i: \overline{S_i} \to Z_i$ be the map associated to $L|_{\overline{S_i}}$. Set $g: \overline{S} \to Z$, where $g := \bigsqcup g_i$ and $Z := \bigsqcup Z_i$. If $\dim Z_i \neq 1$, then g_i satisfies the conditions of Theorem 2.10. Hence, we may assume that $\dim Z_i = 1$ for any i by the inductive argument in the proof of Theorem 4.1.

$$\begin{tikzcd}
\overline{C} \arrow[r, "\text{normalization}"'] \arrow[d, "f_1"'] \arrow[rrrr, bend left, "f_2"] & \mathscr{C}_{\mathrm{red}} \arrow[r] & \mathscr{C} \arrow[r, hook] \arrow[d] & \overline{S} \arrow[r, "g"'] \arrow[d, "\nu"] & Z \\
\overline{D} \arrow[r, "\text{normalization}"] & \mathscr{D}_{\mathrm{red}} \arrow[r] & \mathscr{D} \arrow[r, hook] & S &
\end{tikzcd}$$

By Remark 2.1, it is sufficient to show that, for any point $p \in \overline{S}$, there exist $m \geq 1$ and a section $s \in H^0(\overline{S}, L^{\otimes m}|_{\overline{S}})$ such that $s|_{\mathscr{C}}$ descends to $\mathscr{D}$ and $s|_p \neq 0$. To obtain this, we prove the following claim:

Claim 5.3. *For any finite set $F \subset \overline{S}$ of closed points of $\overline{S}$, we can find $m \geq 1$ and a section $s \in H^0(\overline{S}, L^{\otimes m}|_{\overline{S}})$ such that $s|_{\overline{C}}$ descends to $\overline{D}$ and s is nowhere-vanishing on F.*

First, we assume this claim and complete the proof of Theorem 5.2. Let $F' \subset \mathscr{D}_{\mathrm{red}}$ be the conductor corresponding to the normalization $\overline{D} \to \mathscr{D}_{\mathrm{red}}$. Let F'' be the image of F' in S. Set $F := \nu^{-1}(F'') \cup \{p\}$. Then F is a finite set.

By Claim 5.3, we can take $s \in H^0(\overline{S}, L^{\otimes m}|_{\overline{S}})$ and $s_{\overline{D}} \in H^0(\overline{D}, L^{\otimes m}|_{\overline{D}})$ such that $s|_{\overline{C}} = s_{\overline{D}}|_{\overline{C}}$ and s is nowhere-vanishing on F. By Lemma 2.5, if we replace $s_{\overline{D}}$ by some power of it, then $s_{\overline{D}}$ descends to a section $s_{\mathscr{D}_{\mathrm{red}}}$ on $\mathscr{D}_{\mathrm{red}}$. Since $\mathscr{D}_{\mathrm{red}} \to \mathscr{D}$ is a universal homeomorphism, $s_{\mathscr{D}_{\mathrm{red}}}$ descends to a section $s_{\mathscr{D}}$ on $\mathscr{D}$, if we replace $s_{\overline{D}}$ by some power of it (see Theorem 2.9).

It is sufficient to show that $s|_{\mathscr{C}} = s_{\mathscr{D}}|_{\mathscr{C}}$. By construction, $(s|_{\mathscr{C}})|_{\overline{C}} = (s_{\mathscr{D}}|_{\mathscr{C}})|_{\overline{C}}$ holds. Since $\overline{C} \to \mathscr{C}_{\mathrm{red}}$ is surjective, we get $(s|_{\mathscr{C}})|_{\mathscr{C}_{\mathrm{red}}} = (s_{\mathscr{D}}|_{\mathscr{C}})|_{\mathscr{C}_{\mathrm{red}}}$. As $\mathscr{C}_{\mathrm{red}} \to \mathscr{C}$ is a universal homeomorphism, if we replace s by some power of it, then we get $s|_{\mathscr{C}} = s_{\mathscr{D}}|_{\mathscr{C}}$ (see Theorem 2.9). This completes the proof of Theorem 5.2.

Proof of Claim 5.3. Let f_1 and f_2 be as in the above diagram. For a one-dimensional scheme X, we write $X = X^{(0)} \sqcup X^{(1)}$, where $X^{(i)}$ is the i-dimensional part. Further, we write $\overline{C}^{(1)} = \overline{C}^{\mathrm{h}} \sqcup \overline{C}^{\mathrm{v}}$, where $\overline{C}^{\mathrm{h}}$ is the f_2-horizontal part and $\overline{C}^{\mathrm{v}}$ is the f_2-vertical part.

First, we claim that, for any closed point $p \in Z$, the inverse image of p by $\overline{C}^{\mathrm{h}} \to Z$ has length at most two. This can be proved as follows: by the nefness of $L - (K_{\overline{S_i}} + \Delta_{\overline{S_i}})$, we have

$$0 \leq G_i \cdot (L - (K_{\overline{S_i}} + \Delta_{\overline{S_i}})) = -G_i \cdot (K_{\overline{S_i}} + \Delta_{\overline{S_i}}) \leq 2 - G_i \cdot \Delta_{\overline{S_i}},$$

where G_i is a general fiber of $g_i \colon \overline{S_i} \to Z_i$. Since the one-dimensional part of $\mathscr{C}|_{\overline{S_i}}$ is contained in $\mathrm{Supp}(\lfloor \Delta_{\overline{S_i}} \rfloor)$, we have

$$\#(G_i \cap \mathscr{C}|_{\overline{S_i}}) \leq G_i \cdot \Delta_{\overline{S_i}} \leq 2.$$

Hence, $f_2 \colon \overline{C}^{\mathrm{h}} \to Z$ satisfies the assumption of Lemma 5.4. Further, by conditions ($\star$) and ($\star\star$), $f_1 \colon \overline{C}^{\mathrm{h}} \to D'$ also satisfies the assumption of Lemma 5.4, where we define $D' := f_1(\overline{C}^{\mathrm{h}})$.

$$\begin{tikzcd} \overline{C} = \overline{C}^{\mathrm{h}} \sqcup \overline{C}^{\mathrm{v}} \sqcup \overline{C}^{(0)} \arrow[r, "f_2"] \arrow[d, "f_1"'] & Z \\ \overline{D} = D' \sqcup (\overline{D} \setminus D') & \end{tikzcd} \qquad \begin{tikzcd} \overline{C}^{\mathrm{h}} \arrow[r, "f_2"] \arrow[d, "f_1"'] & Z \\ D' & \end{tikzcd}$$

By Lemma 5.4, we can find sections $s_{\overline{D}} \in H^0(\overline{D}, L^{\otimes m}|_{\overline{D}})$ and $s_Z \in H^0(Z, L^{\otimes m}|_Z)$ such that $s_{\overline{D}}|_{\overline{C}^{\mathrm{h}}} = s_Z|_{\overline{C}^{\mathrm{h}}}$ holds, the section s_Z is nowhere-vanishing on the finite set $g(F) \cup f_2(\overline{C}^{\mathrm{v}} \sqcup \overline{C}^{(0)})$, and the section $s_{\overline{D}}$ is nowhere-vanishing on $\overline{D} \setminus D'$. Since $L|_{\overline{C}^{\mathrm{v}} \sqcup \overline{C}^{(0)}}$ is trivial, we have $s^n_{\overline{D}}|_{\overline{C}^{\mathrm{v}} \sqcup \overline{C}^{(0)}} = s^n_Z|_{\overline{C}^{\mathrm{v}} \sqcup \overline{C}^{(0)}}$ for some $n \geq 1$ by Lemma 2.4. Therefore, we get $s^n_{\overline{D}}|_{\overline{C}} = s^n_Z|_{\overline{C}}$ and this completes the proof of Claim 5.3. □

Finally, we show the next lemma, which was used in the proof of Theorem 5.2.

Lemma 5.4. *Let X, Z_1, Z_2 be disjoint unions of smooth proper curves and $f_1\colon X \to Z_1$, $f_2\colon X \to Z_2$ finite surjective morphisms. Let L_1 and L_2 be line bundles on Z_1 and Z_2, respectively, such that $f_1^*L_1 = f_2^*L_2$. Suppose that $L := f_1^*L_1 = f_2^*L_2$ is semiample. Further, assume that each f_i satisfies either of the following conditions:*

- *The restriction of f_i to any connected component of X is an isomorphism onto its image.*
- *Any fiber of f_i has length at most two.*

Then, for any finite set $F \subset X$ of closed points of X, we can take $m \geq 1$ and a section $s \in H^0(X, L^{\otimes m})$ such that s is nowhere-vanishing on F and s descends to both Z_1 and Z_2.

Proof. First, we prove that there exists a finite group G_i acting on X such that $X \to Z_i$ decomposes into the quotient morphism $X \to X/G_i$ and a universal homeomorphism $X/G_i \to Z_i$:

$$\begin{tikzcd}
& & X \arrow[dll, "f_1"'] \arrow[dl] \arrow[dr] \arrow[drr, "f_2"] & & \\
Z_1 & X/G_1 \arrow[l, "\text{universal homeomorphism}"'] & & X/G_2 \arrow[r, "\text{universal homeomorphism}"'] & Z_2
\end{tikzcd}$$

This is trivial when the restriction of f_i to any connected component of X is an isomorphism. Indeed, it is sufficient to take G_i such that it identifies the components with the same image under f_i. Then $X \to Z_i$ is isomorphic to the quotient morphism $X \to X/G_i$.

For the second case, assume that any fiber of f_i has length at most two. Let Z'_i be a connected component of Z_i. Set $X' = f_i^{-1}(Z'_i)$. There are four possibilities:

(1) X' is connected and $X' \to Z'_i$ is an isomorphism.

(2) X' is connected and $X' \to Z'_i$ is the Frobenius map (this case may only occur for characteristic $p = 2$).

(3) X' is connected and every fiber of $X' \to Z'_i$ has length two. There exists an involution $\iota\colon X' \to X'$ such that $X' \to Z'_i$ is the quotient by ι.

(4) X' has two connected components X'_1 and X'_2. Further, $X'_1 \to Z'_i$ and $X'_2 \to Z'_i$ are isomorphisms. In this case, we have $X'_1 \cong X'_2$.

In the cases (3) and (4), we have a finite group G' acting on X' such that the morphism $X' \to Z'_i$ is isomorphic to the quotient morphism $X' \to X'/G'$.

Hence, we have a finite group G_i acting on X such that the morphism $X \to Z_i$ decomposes as $X \to X/G_i \to Z_i$, where $X \to X/G_i$ is the quotient morphism and $X/G_i \to Z_i$ is a universal homeomorphism (actually, if we restrict it to a connected component, it is either an isomorphism or the Frobenius map).

Note that $L = g^*L$ for any $g \in G_i$. We claim that if $s \in H^0(X, L^{\otimes m})$ is G_i-equivariant, then s^{p^e} descends to Z_i for sufficiently large e. This is because s descends to X/G_i and $X/G_i \to Z_i$ is a universal homeomorphism (see Theorem 2.9).

Let $G := G_1 G_2 \subset \mathrm{Aut}(X)$ be a composition of the groups, and let $S \subset X$ be the G-orbit of the set F. By Lemma 2.6, G is a finite group, and therefore S is a finite set.

Take $m \geq 1$ and a section $s \in H^0(X, L^{\otimes m})$ such that s is nowhere-vanishing on S. Set

$$s^G := \prod_{\sigma \in G} \sigma^* s \in H^0(X, L^{\otimes m|G|}).$$

The section s^G is G_i-invariant for each i and nowhere-vanishing on F. Hence, $(s^G)^{p^e}$ satisfies the statement of the lemma for sufficiently large $e \geq 1$. □

The main issue of this section is related to the following question, discussed by Keel [2003].

Question 5.5. Let L be a line bundle on a variety X and let $p\colon \overline{X} \to X$ be the normalization of X. Assume that p^*L is semiample. What additional assumptions are necessary for L to be semiample?

6. Proof of Theorem 1.1

In this section, we prove Theorem 1.1 using Theorem 4.1 and Theorem 5.2.

Proof of Theorem 1.1. Let $S := \lfloor \Delta \rfloor$. By Theorem 4.1, it is sufficient to show that $L|_{\mathrm{Supp}(S)}$ is semiample. Note that in both case (1) and case (2), all the coefficients of Δ are at most one.

By the adjunction formula (Proposition 2.2), if we define $\Delta_{\overline{S}}$ on $\overline{S}$ so that $(K_X + \Delta)|_{\overline{S}} = K_{\overline{S}} + \Delta_{\overline{S}}$, then $\Delta_{\overline{S}}$ satisfies the conditions in the statement of Theorem 5.2.

In the case (2), that is, the case when each component S_i of S is normal, S clearly satisfies the condition ($\star$). In the case (1), that is, the case when (X, Δ) is log canonical, the surface S is regular or nodal in codimension one (see [Kollár 2013, Corollary 2.32]), and so S satisfies condition ($\star\star$) (see [Kollár 2013, Claim 1.41.1] or [Tanaka 2014, Lemma 3.4, 3.5]).

Thus, we can complete the proof by using Theorem 5.2. □

We easily deduce Corollary 1.2:

Proof of Corollary 1.2. It is enough to take $L = 2(K_X + \Delta)$ and $L = -(K_X + \Delta)$, respectively. □

7. Examples

Theorem 1.1 does not hold if we do not impose any conditions on Δ. It is in fact possible to construct a nef and big line bundle L on a smooth threefold X such that $L - (K_X + \Delta)$ is nef and big for $\Delta \geq 0$, but L is not semiample. We construct such L and Δ in the following way:

Example 7.1. Let L be a nef and big line bundle on a smooth threefold which is not semiample (see an example in [Totaro 2009, Theorem 7.1]). Since L is big, we can write $L = A + E$ for an ample $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor A and an effective $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor E. Take $\Delta = mE$ for $m \in \mathbb{N}$ big enough. Then $mL - (K_X + \Delta)$ is an ample Cartier divisor, and so the pair $L' := mL$ and Δ is an example which we were looking for.

Theorem 1.1 does not hold over algebraically closed fields $k \neq \overline{\mathbb{F}}_p$ even in the two-dimensional case:

Example 7.2 [Tanaka 2012, Example 19.3]. Let $C_0 \subset \mathbb{P}^2$ be an elliptic curve in $\mathbb{P}^2$, and let $p_1, \dots, p_{10} \in C_0$ be ten general points on C_0. Let X be the blowup of $\mathbb{P}^2$ along these ten points, and C the proper transform of C_0. Note that $K_X + C \sim 0$ and $C^2 = -1$.

Take an ample divisor H on X, and set $L := H + aC$, where $a := H \cdot C > 0$. Note that L is a nef and big divisor. Further, (X, C) is log canonical, and $L - (K_X + C)$ is also nef and big. Nevertheless, L is not semiample if the base field is not $\overline{\mathbb{F}}_p$. This is because $L \cdot C = 0$, but the elliptic curve C is not contractible.

Corollary 1.2(2) also does not hold over algebraically closed fields $k \neq \overline{\mathbb{F}}_p$:

Example 7.3 [Gongyo 2012, Example 5.2]. Let S be the blowup of $\mathbb{P}^2$ along nine general points. Note that $-K_S$ is nef but not semiample if the base field is not $\overline{\mathbb{F}}_p$. Take a very ample divisor H on S, and set $X := \mathbb{P}_S(\mathcal{O}_S \oplus \mathcal{O}_S(-H))$. Let E be the tautological section of $\mathcal{O}_S \oplus \mathcal{O}_S(-H)$. Since $E \cong S$, it follows that $-K_E$ is not semiample.

Then, (X, E) is log canonical, and $L := -(K_X + E)$ is nef and big by the nefness of $-K_S$ (for details, see [Gongyo 2012, Example 5.2]). Nevertheless, L is not semiample, because $L|_E = -K_E$ is not semiample.

Acknowledgements

We would like to thank Paolo Cascini and Yoshinori Gongyo for suggesting this problem and for their constant support and guidance.

The project started during the Pragmatic Research Summer School 2013 in Catania and part of the work carried on during the visit of Martinelli and Witaszek to the University of Tokyo. We are grateful to the University of Catania and the organizers of the Pragmatic Research School, as well as to Yujiro Kawamata and Yoshinori Gongyo for the invitation to the University of Tokyo.

We are grateful to Andrea Fanelli, Enrica Floris, Atsushi Ito, and Hiromu Tanaka for useful comments and suggestions.

Martinelli is supported by a Roth studentship. Nakamura is supported by the Grant-in-Aid for Scientific Research (KAKENHI No. 25-3003) and the Program for Leading Graduate Schools, MEXT, Japan. Witaszek was supported by Bonn International Graduate School's Pre-Ph.D. scholarship.

References

[Artin 1962] M. Artin, "Some numerical criteria for contractability of curves on algebraic surfaces", *Amer. J. Math.* **84** (1962), 485–496. MR 26 #3704 Zbl 0105.14404

[Atiyah 1957] M. F. Atiyah, "Vector bundles over an elliptic curve", *Proc. London Math. Soc.* (3) **7** (1957), 414–452. MR 24 #A1274 Zbl 0084.17305

[Birkar 2013] C. Birkar, "Existence of flips and minimal models for 3-folds in char p", preprint, 2013. arXiv 1311.3098v1

[Bombieri and Mumford 1976] E. Bombieri and D. Mumford, "Enriques' classification of surfaces in char p, III", *Invent. Math.* **35** (1976), 197–232. MR 58 #10922b Zbl 0336.14010

[Bombieri and Mumford 1977] E. Bombieri and D. Mumford, "Enriques' classification of surfaces in char p, II", pp. 23–42 in *Complex analysis and algebraic geometry*, edited by W. L. Baily, Jr., Iwanami Shoten, Tokyo, 1977. MR 58 #10922a Zbl 0348.14021

[Cascini et al. 2014] P. Cascini, J. McKernan, and M. Mustaţă, "The augmented base locus in positive characteristic", *Proc. Edinb. Math. Soc.* (2) **57**:1 (2014), 79–87. MR 3165013 Zbl 1290.14006

[Fong and McKernan 1992] L.-Y. Fong and J. McKernan, "Log abundance for surfaces", pp. 127–137 in *Flips and abundance for algebraic threefolds* (Salt Lake City, 1991), edited by J. Kollár, Astérisque **211**, Société Mathématique de France, Paris, 1992. MR 94f:14013 Zbl 0807.14029

[Fujino 2000] O. Fujino, "Abundance theorem for semi log canonical threefolds", *Duke Math. J.* **102**:3 (2000), 513–532. MR 2001c:14032 Zbl 0986.14007

[Gongyo 2012] Y. Gongyo, "On weak Fano varieties with log canonical singularities", *J. Reine Angew. Math.* **665** (2012), 237–252. MR 2908745 Zbl 1243.14018

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York-Heidelberg, 1977. MR 57 #3116 Zbl 0367.14001

[Keel 1999] S. Keel, "Basepoint freeness for nef and big line bundles in positive characteristic", *Ann. of Math.* (2) **149**:1 (1999), 253–286. MR 2000j:14011 Zbl 0954.14004

[Keel 2003] S. Keel, "Polarized pushouts over finite fields", *Comm. Algebra* **31**:8 (2003), 3955–3982. MR 2004h:14010 Zbl 1051.14017

[Kollár 2013] J. Kollár, *Singularities of the minimal model program*, Cambridge Tracts in Mathematics **200**, Cambridge University Press, 2013. MR 3057950 Zbl 1282.14028

[Kollár and Mori 1998] J. Kollár and S. Mori, *Birational geometry of algebraic varieties*, Cambridge Tracts in Mathematics **134**, Cambridge University Press, 1998. MR 2000b:14018 Zbl 0926.14003

[Lehmann 2012] B. Lehmann, "Numerical triviality and pullbacks", preprint, 2012. arXiv 1109.4382v3

[Liedtke 2013] C. Liedtke, "Algebraic surfaces in positive characteristic", pp. 229–292 in *Birational geometry, rational curves, and arithmetic*, edited by F. Bogomolov et al., Springer, New York, 2013. MR 3114931 Zbl 06211443

[Liu 2002] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Graduate Texts in Mathematics **6**, Oxford University Press, 2002. MR 2003g:14001 Zbl 0996.14005

[Mumford 1969] D. Mumford, "Enriques' classification of surfaces in char p, I", pp. 325–339 in *Global Analysis (Papers in Honor of K. Kodaira)*, edited by D. C. Spencer and S. Iyanaga, University of Tokyo Press, 1969. MR 40 #7266 Zbl 0188.53201

[Mumford 2008] D. Mumford, *Abelian varieties*, Tata Institute of Fundamental Research Studies in Mathematics **5**, Hindustan Book Agency, New Delhi, 2008. Corrected reprint of the 1974 edition. MR 2010e:14040 Zbl 1177.14001

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Mathematics **106**, Springer, Dordrecht, 2009. MR 2010i:11005 Zbl 1194.11005

[Tanaka 2012] H. Tanaka, "Minimal models and abundance for positive characteristic log surfaces", preprint, 2012. arXiv 1201.5699v2

[Tanaka 2014] H. Tanaka, "Abundance theorem for semi log canonical surfaces in positive characteristic", preprint, 2014. arXiv 1301.6889v2

[Totaro 2009] B. Totaro, "Moving codimension-one subvarieties over finite fields", *Amer. J. Math.* **131**:6 (2009), 1815–1833. MR 2011c:14069 Zbl 1200.14022

[Xu 2013] C. Xu, "On base point free theorem of threefolds in positive characteristic", preprint, 2013. arXiv 1311.3819v1

Communicated by Shigefumi Mori
Received 2014-09-02 Revised 2015-01-11 Accepted 2015-02-16

d.martinelli12@imperial.ac.uk *Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2AZ, United Kingdom*

nakamura@ms.u-tokyo.ac.jp *Graduate School of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan*

jakub.witaszek.14@ucl.ac.uk *The London School of Geometry and Number Theory, Department of Mathematics, University College London, Gower Street, London WC1E 6BT, United Kingdom*

ALGEBRA AND NUMBER THEORY 9:3 (2015)
dx.doi.org/10.2140/ant.2015.9.749

The torsion group of endotrivial modules

Jon F. Carlson and Jacques Thévenaz

Let G be a finite group and let $T(G)$ be the abelian group of equivalence classes of endotrivial kG-modules, where k is an algebraically closed field of characteristic p. We determine, in terms of the structure of G, the kernel of the restriction map from $T(G)$ to $T(S)$, where S is a Sylow p-subgroup of G, in the case when S is abelian. This provides a classification of all torsion endotrivial kG-modules in that case.

1. Introduction

Endotrivial modules for a finite group G over a field k of prime characteristic p play a significant role in modular representation theory. Among other things, they form an important part of the Picard group of self-equivalences of the stable category stmod(kG) of finitely generated kG-modules modulo projectives. They are modules which have universal deformation rings [de Smit and Lenstra 1997]. The endotrivial modules have been classified in the case that G is a p-group [Carlson and Thévenaz 2004; 2005] and various results have appeared since for some specific families of groups [Carlson et al. 2006; 2009; 2010; 2011; 2013; 2014a; 2014b; Mazza 2007; Mazza and Thévenaz 2007; Navarro and Robinson 2012; Lassueur and Mazza 2015]. Recently, another line of research has developed that is concerned with the classification of all endotrivial modules which are simple [Robinson 2011; Lassueur et al. 2013; Lassueur and Malle 2015].

The main purpose of this paper is to classify all torsion endotrivial modules when a Sylow p-subgroup is abelian. We let G be a finite group and $T(G)$ be the abelian group of equivalence classes of endotrivial kG-modules, where k is an algebraically closed field of characteristic p. Let S be a Sylow p-subgroup of G. We fix the notation

$$K(G) = \mathrm{Ker}\{\mathrm{Res}^G_S : T(G) \longrightarrow T(S)\}.$$

One of the main open questions is to describe this kernel explicitly, and we achieve this goal here in the case that S is abelian. Actually, $K(G)$ is known to be equal to

Carlson is partially supported by a grant from NSF.
MSC2010: 20C20.
Keywords: modular representation theory.

the torsion subgroup of $T(G)$ in most cases. Specifically, this happens whenever S is not cyclic, generalized quaternion, or semidihedral — because, if we exclude these three cases, then $T(S)$ is torsion-free by [Carlson and Thévenaz 2005]. The excluded cases are treated in [Mazza and Thévenaz 2007; Kawata 1993; Carlson et al. 2013].

Let $N = N_G(S)$ denote the normalizer of a Sylow p-subgroup S of G. It is known that the restriction map $\mathrm{Res}^G_N : K(G) \to K(N)$ is injective, induced by the Green correspondence, and the problem is to describe its image. In fact, $K(N)$ consists of all one-dimensional representations of N and the main difficulty is to know which of them lie in the image of Res^G_N; that is, which of them have Green correspondents that are endotrivial. Indeed, if J is the intersection of the kernels of the one-dimensional kN-modules whose Green correspondents are endotrivial, then $K(G)$ is isomorphic to the dual group $(N/J)^*$. So the problem of finding $K(G)$ comes down to the question of what is J.

Another approach was introduced by Balmer [2013], in which he shows that $K(G)$ is isomorphic to the group $A(G)$ of all weak homomorphisms $G \to k^\times$ (defined below in Section 3). Balmer's method was used effectively in [Carlson et al. 2014a] to compute $K(G)$ in some crucial cases for G a special linear group. The method involved the construction of a system of local subgroups $\{\rho^i(Q)\}$ indexed on the collection of nontrivial subgroups Q of the Sylow subgroup S of G and for $i \geq 1$ (see Section 4 for the definition). These subgroups are in the kernels of all weak homomorphisms and we show here that $\rho^i(S) \subseteq N_G(S)$ is in the kernel of all one-dimensional representations of $N_G(S)$ whose Green correspondents are endotrivial modules. That is, $\rho^i(S) \subseteq J$, for J defined as above.

Hence, the question of determining $K(G)$ becomes: is J equal to $\rho^\infty(S)$, the limit of the system? The answer is yes in all examples that we know. The main theorem of this paper says that the answer is yes whenever the Sylow p-subgroup of G is abelian. Indeed, we prove more. We show that $J = \rho^2(S)$, the subgroup of $N = N_G(S)$ generated by the commutator subgroup $[N, N]$, S, and all intersections $N \cap [N_G(Q), N_G(Q)]$ for Q a nontrivial subgroup of S. Thus, $K(G) \cong (N/\rho^2(S))^*$ in the case that S is abelian. Balmer's characterization of $K(G)$ in terms of weak homomorphisms is crucial for the proof, which appears in Section 5.

In Section 6 we show that the main theorem can be used to describe $K(G)$ explicitly when S is cyclic. An extension of the main theorem to the case where the normalizer N of S controls p-fusion is given in Section 7. The paper ends with some examples of simple or almost simple groups where the subgroup $K(G)$ is not trivial.

2. Endotrivial modules and the restriction to the Sylow subgroup

Throughout this paper, k denotes an algebraically closed field of prime characteristic p and G is a finite group. We assume that all modules are finitely generated. If

M and L are kG-modules, the notation $M \cong L \oplus (\mathrm{proj})$ means that M is isomorphic to the direct sum of L with some projective kG-module, which might be zero. We write k for the trivial kG-module. Unless otherwise specified, the symbol $\otimes$ is the tensor product $\otimes_k$ of the underlying vector spaces. In the case that M and L are kG-modules, the tensor product $M \otimes L$ is a kG-module with G acting by the diagonal action on the factors.

We assume that G has order divisible by p and we let S be a Sylow p-subgroup of G. Recall that a kG-module M is endotrivial provided its endomorphism algebra $\mathrm{End}_k(M)$ is isomorphic (as a kG-module) to the direct sum of the trivial module k and a projective kG-module. In other words, M is endotrivial if and only if $\mathrm{Hom}_k(M, M) \cong M^* \otimes M \cong k \oplus (\mathrm{proj})$, where M^* denotes the k-dual of M. Any endotrivial module M splits as the direct sum $M = M_0 \oplus (\mathrm{proj})$ for an indecomposable endotrivial kG-module M_0, which is unique up to isomorphism. We let $T(G)$ be the set of equivalence classes of endotrivial kG-modules for the equivalence relation

$$M \sim L \iff M_0 \cong L_0.$$

Every equivalence class contains a unique indecomposable module up to isomorphism. The tensor product induces an abelian group structure on the set $T(G)$, written additively as $[M]+[L]=[M\otimes L]$. The zero element of $T(G)$ is the class $[k]$ of the trivial module, while the inverse of the class of a module M is the class of the dual module M^*. The group $T(G)$ is known to be a finitely generated abelian group.

Lemma 2.1 [Carlson et al. 2011, Lemma 2.3]. *Let $K(G)$ be the kernel of the restriction map $\mathrm{Res}^G_S : T(G) \longrightarrow T(S)$ to a Sylow p-subgroup S.*

(a) *$K(G)$ is a finite subgroup of $T(G)$.*

(b) *$K(G)$ is the entire torsion subgroup $TT(G)$ of $T(G)$, provided S is not cyclic, generalized quaternion, or semidihedral.*

We say that a kG-module M has trivial Sylow restriction if the restriction of M to a Sylow p-subgroup S has the form $M{\downarrow}^G_S \cong k \oplus (\mathrm{proj})$. Any such module is endotrivial and is the direct sum of an indecomposable trivial source module and a projective module. Thus M has trivial Sylow restriction if and only if its class $[M]$ belongs to $K(G)$.

Proposition 2.2. *Suppose that a finite group H has a nontrivial normal p-subgroup. Then every indecomposable kH-module with trivial Sylow restriction has dimension one.*

Proof. The proof is a straightforward application of the Mackey formula. The details appear in Lemma 2.6 in [Mazza and Thévenaz 2007]. $\square$

In the situation of Proposition 2.2, for any one-dimensional kH-module L we write $\chi_L : H \to k^\times$ for the corresponding group homomorphism (representation).

Our next result is an easy application of the Green correspondence. For details, see Proposition 2.6 in [Carlson et al. 2006].

Proposition 2.3. *Let S be a Sylow p-subgroup of G and let $N = N_G(S)$.*

(a) *The restriction map* $\mathrm{Res}^G_N : T(G) \to T(N)$ *is injective, induced by the Green correspondence.*

(b) *In particular, the restriction map* $\mathrm{Res}^G_N : K(G) \to K(N)$ *is injective.*

We emphasize that, if M and L are kG-modules with trivial Sylow restriction then, by Proposition 2.2, $M{\downarrow}^G_N \cong U \oplus (\mathrm{proj})$ and $L{\downarrow}^G_N \cong V \oplus (\mathrm{proj})$, where U and V are kN-modules of dimension one. Here U is the Green correspondent of M, V is the Green correspondent of L, and we see automatically that $U \otimes V$ is the Green correspondent of $(M \otimes L)_0$, the unique indecomposable nonprojective direct summand of $M \otimes L$.

We know that $K(N_G(S))$ consists exactly of all one-dimensional representations of N. The main problem is to know which of them are in the image of the restriction map from $K(G)$. In other words, given a one-dimensional kN-module U, we need to know when its Green correspondent M is endotrivial.

Another way of viewing the situation is the following:

Proposition 2.4. *Let S denote a Sylow p-subgroup of G and let $N = N_G(S)$. Let $J \subseteq N$ be the intersection of the kernels of all one-dimensional kN-modules U such that the Green correspondent M of U is an endotrivial kG-module. That is, J is the intersection of the kernels of U such that $[U]$ is in the image of the restriction* $\mathrm{Res}^G_N : T(G) \to T(N)$. *Then* $K(G) \cong (N/J)^* \cong \mathrm{Hom}(N/J, k^\times)$.

Proof. The restriction map $\mathrm{Res}^G_N : K(G) \to K(N)$, being injective, gives an isomorphism between $K(G)$ and a subgroup $A \subseteq K(N) \cong \mathrm{Hom}(N, k^\times)$, the group of one-dimensional representations of N. But A is isomorphic to the dual group of N/J, where J is the intersection of the kernels of the elements of A. □

From the proposition, we see that the problem of characterizing the group $K(G)$ is equivalent to finding the group J. The main purpose of this paper is to offer a possible candidate for the subgroup J. In addition, we prove that the candidate is, in fact, equal to J in the case that the Sylow p-subgroup S is abelian.

3. Weak homomorphisms and the kernel of restriction

Balmer [2013] provided a new characterization of the group $K(G)$ in terms of the group of weak S-homomorphisms, which we describe in this section. Note that Balmer's construction is more general than the one that we use here. He defined "weak H-homomorphisms" for any subgroup H containing S. Because, we only deal with the case that $H = S$ in this paper, we call them "weak homomorphisms". Balmer [2015] has expanded his results, but here we only need the formulation in [Balmer 2013].

For notation, recall that ${}^{g}S$ denotes the conjugate subgroup gSg^{-1}, while S^{g} denotes $g^{-1}Sg$.

Definition 3.1. A map $\chi : G \to k^{\times}$ is called a *weak homomorphism* ("weak S-homomorphism" in the language of [Balmer 2013]) if it satisfies the following three conditions:

(a) If $s \in S$, then $\chi(s) = 1$.

(b) If $g \in G$ and $S \cap {}^{g}S = \{1\}$, then $\chi(g) = 1$.

(c) If $a, b \in G$ and if $S \cap {}^{a}S \cap {}^{ab}S \neq \{1\}$, then $\chi(ab) = \chi(a)\chi(b)$.

The product of two weak homomorphisms χ and ψ is defined pointwise by $(\chi\psi)(g) = \chi(g)\psi(g)$ and is again a weak homomorphism. The set $A(G)$ of all weak homomorphisms is an abelian group under this operation.

Theorem 3.2 [Balmer 2013]. *The groups $K(G)$ and $A(G)$ are isomorphic.*

The isomorphism is explicit and is described in detail in [Balmer 2013]. In particular, it is shown that, given a weak homomorphism χ, there is a certain norm-type formula using χ that constructs a homomorphism from the permutation module $k(G/S)$ to itself, whose image is the endotrivial module associated to χ.

It is important, in what follows, to understand Balmer's isomorphism on restriction to a subgroup H having a nontrivial normal p-subgroup. This is our next result. For notation, given a finite group H, we let $\diamondsuit(H)$ be the smallest normal subgroup of H such that $H/\diamondsuit(H)$ is an abelian p'-group. In other words $\diamondsuit(H) = [H, H]S$ is the subgroup of H generated by the commutator subgroup $[H, H]$ and by a Sylow p-subgroup S of H.

Proposition 3.3. *Suppose that a finite group H has a nontrivial normal p-subgroup.*

(a) *Every weak homomorphism $\chi : H \to k^{\times}$ is a group homomorphism.*

(b) *The isomorphism $K(H) \to A(H)$ maps the class of a one-dimensional kH-module M to the corresponding group representation $H \to k^{\times}$.*

(c) *The group $A(H)$ is isomorphic to the group of one-dimensional representations of H, that is, the dual group $(H/\diamondsuit(H))^{*}$.*

Proof. To prove (a), let Q be a nontrivial normal p-subgroup of H. For any $a, b \in G$, the subgroup $S \cap {}^{a}S \cap {}^{ab}S$ is nontrivial because it contains $Q = {}^{a}Q = {}^{ab}Q$. Thus the third condition in Definition 3.1 implies that $\chi(ab) = \chi(a)\chi(b)$.

Statement (b) follows from the construction of the isomorphism given in Section 2.5 of [Balmer 2013]. See also Corollary 5.1 in [Balmer 2013]. The proof of (c) is straightforward, because the image of any group homomorphism $H \to k^\times$ is contained in the subgroup of $k^\times$ consisting of roots of unity, namely the group of all p'-roots of unity since k is algebraically closed and has characteristic p. □

Proposition 3.4. *Let $N = N_G(S)$, where S is the Sylow p-subgroup of G. The restriction map $\operatorname{Res}^G_N : A(G) \to A(N)$ is injective.*

Proof. This follows from Proposition 2.3, Balmer's isomorphism in Theorem 3.2, and the fact that this isomorphism is natural. □

4. A system of local subgroups

In this section, we discuss the properties of a sequence of subgroups $\rho^i(Q) \subseteq N_G(Q)$, where Q is a nontrivial p-subgroup of S and $i \geq 1$. The construction of the sequence was first presented in [Carlson et al. 2014a], though the version here is slightly different. The definition of the subgroups $\rho^i(Q)$ depends not only on $N_G(Q)$, but involves all normalizers of nontrivial subgroups of the Sylow subgroup S of G.

The definition proceeds inductively as follows. We fix a Sylow p-subgroup S of G. For any nontrivial subgroup Q of S, let

$$\rho^1(Q) := \diamondsuit(N_G(Q)).$$

As before, $\diamondsuit(N_G(Q))$ is the product of the commutator subgroup of $N_G(Q)$ and a Sylow p-subgroup of $N_G(Q)$. The original version in [Carlson et al. 2014a] uses simply the commutator subgroup of $N_G(Q)$. For $i > 1$, we let

$$\rho^i(Q) := \langle N_G(Q) \cap \rho^{i-1}(Q') \mid \{1\} \neq Q' \subseteq S \rangle,$$

the subgroup generated by the subgroups $N_G(Q) \cap \rho^{i-1}(Q')$ for all nontrivial subgroups Q' of S. This contains $\rho^{i-1}(Q)$, so we have a nested sequence of subgroups

$$\rho^1(Q) \subseteq \rho^2(Q) \subseteq \rho^3(Q) \subseteq \cdots \subseteq N_G(Q).$$

Since G is finite, the sequence eventually stabilizes, and we let $\rho^\infty(Q)$ be the limit subgroup of the sequence $\{\rho^i(Q) \mid i \geq 1\}$, namely their union.

The definition of the subgroups $\rho^i(Q)$ was originally motivated in [Carlson et al. 2014a] by the following observation:

Proposition 4.1. *Suppose that $\chi : G \to k^\times$ is a weak homomorphism as defined in the last section. If $x \in \rho^i(Q)$ for some $i \geq 1$ and for some nontrivial subgroup Q of S, then $\chi(x) = 1$.*

Proof. In the case that $i = 1$, the statement is a trivial consequence of the definition of a weak homomorphism. That is, $\chi(x) = 1$ for any $x \in Q$ by Definition 3.1(a), and is a homomorphism to an abelian group when restricted to $N_G(Q)$ by Definition 3.1(c). So assume that $i > 1$ and that $\chi(x) = 1$ for all $x \in \rho^{i-1}(Q)$ for all nontrivial subgroups $Q \subseteq S$. Then $\chi(x) = 1$ for all $x \in N_G(Q) \cap \rho^{i-1}(Q')$ for any nontrivial subgroups Q and Q'. Thus, $\chi(x) = 1$ for all $x \in \rho^i(Q)$, and the proposition is proved by induction. □

Suppose that M is a kG-module with trivial Sylow restriction. By Proposition 2.2, for any nontrivial subgroup Q of S there is a one-dimensional $kN_G(Q)$-module L_Q such that $M{\downarrow}^G_{N_G(Q)} \cong L_Q \oplus (\text{proj})$. We write $\chi_Q = \chi_{L_Q}$ for the corresponding group homomorphism $\chi_Q : N_G(Q) \to k^\times$.

The next result encapsulates the main idea of this section. It should be compared with Proposition 4.1.

Proposition 4.2. *Suppose that M is a kG-module with trivial Sylow restriction, let Q be a nontrivial subgroup of S and let $\chi_Q : N_G(Q) \to k^\times$ be defined as above. Then $\rho^\infty(Q)$ is contained in the kernel of χ_Q.*

Proof. We prove, by induction on i, that if $x \in \rho^i(Q)$ then $\chi_Q(x) = 1$. In the case that $i = 1$, the result is a consequence of the fact that $\rho^1(Q) = \diamondsuit(N_G(Q))$ is in the kernel of every one-dimensional character on $N_G(Q)$. Inductively, we assume that the lemma is true for $\rho^j(Q')$ whenever $j < i$ and $1 \neq Q' \subseteq S$. Because $x \in \rho^i(Q)$, we have that $x = x_1 \cdots x_m$ for some m, where $x_t \in N_G(Q) \cap \rho^{i-1}(Q_t)$, for some nontrivial subgroup Q_t of S, for each $t = 1, \dots, m$. For each t we consider the restriction of M to $N_G(Q) \cap N_G(Q_t)$. Using the notation above, this yields

$$M{\downarrow}^G_{N_G(Q)\cap N_G(Q_t)} \cong L_Q{\downarrow}^{N_G(Q)}_{N_G(Q)\cap N_G(Q_t)} \oplus (\text{proj}) \cong L_{Q_t}{\downarrow}^{N_G(Q_t)}_{N_G(Q)\cap N_G(Q_t)} \oplus (\text{proj})\,.$$

The intersection $N_G(Q) \cap N_G(Q_t)$ contains the center of S, which is a nontrivial p-subgroup. Thus, any indecomposable projective module over $N_G(Q) \cap N_G(Q_t)$ has dimension divisible by p. Therefore, by the Krull–Schmidt theorem,

$$L_Q{\downarrow}^{N_G(Q)}_{N_G(Q)\cap N_G(Q_t)} \cong L_{Q_t}{\downarrow}^{N_G(Q_t)}_{N_G(Q)\cap N_G(Q_t)}\,.$$

For all t, it follows by induction that $\chi_Q(x_t) = \chi_{Q_t}(x_t) = 1$, because $x_t \in \rho^{i-1}(Q_t)$. Thus $\chi_Q(x) = 1$, as asserted. □

The main theorem of this section is a special case of the above.

Theorem 4.3. *Suppose that M is an endotrivial kG-module with trivial Sylow restriction, i.e., $M{\downarrow}^G_S \cong k \oplus (\text{proj})$. Then $M{\downarrow}^G_{\rho^\infty(S)} \cong k \oplus (\text{proj})$.*

There are several immediate corollaries.

Corollary 4.4. *Suppose that $\rho^\infty(S) = N_G(S)$. Then the only indecomposable kG-module with trivial Sylow restriction is the trivial module; that is, $K(G) = \{0\}$.*

Proof. If M is a kG-module with trivial Sylow restriction, then our assumption implies that $M{\downarrow}^G_{N_G(S)} \cong k \oplus (\text{proj})$. Because M is indecomposable, it is the Green correspondent of the trivial module $k_{N_G(S)}$. Thus M is the trivial module. □

Corollary 4.5. *Suppose that Q is a normal subgroup of S and that, for some i, $N_G(S) \subseteq \rho^i(Q)$. Then the only indecomposable kG-module with trivial Sylow restriction is the trivial module. In other words, $K(G) = \{0\}$.*

Proof. Under the hypothesis, we have that $N_G(S) \subseteq \rho^{i+1}(S)$, and we are done by the previous corollary. □

Note in the above corollary that Q being normal in S does not assure that $N_G(S)$ is a subgroup of $N_G(Q)$. However, if Q is characteristic in S, then this is a certainty.

Corollary 4.6. *Suppose that Q is a characteristic subgroup of S and that, for some i, $N_G(Q) = \rho^i(Q)$. Then the only indecomposable kG-module with trivial Sylow restriction is the trivial module. In other words, $K(G) = \{0\}$.*

5. Abelian Sylow subgroup

The purpose of this section is to prove our main theorem, which describes $K(G)$ when a Sylow p-subgroup S of G is abelian. Specifically, we show the following:

Theorem 5.1. *Suppose that a Sylow p-subgroup S of G is abelian. Let $N = N_G(S)$.*

(a) *The image of the restriction map $\mathrm{Res}^G_N : A(G) \to A(N)$ consists exactly of all group homomorphisms $N_G(S) \to k^\times$ having $\rho^2(S)$ in their kernel.*

(b) $K(G) \cong A(G) \cong (N_G(S)/\rho^2(S))^*$.

(c) $\rho^2(S) = \rho^\infty(S)$.

In other words, the theorem says that if the Sylow subgroup S of G is abelian, then the subgroup J of Proposition 2.4 is equal to $\rho^2(S)$ and that $K(G) \cong (N/\rho^2(S))^*$.

Note that, by Proposition 4.1, the restriction to $N_G(Q)$ of any weak homomorphism $\phi : G \to k^\times$ is a homomorphism having $\rho^2(Q)$ in its kernel. Our task then is to prove a very strong converse, namely, that any group homomorphism $\chi : N = N_G(S) \to k^\times$ having $\rho^2(S)$ in its kernel is the restriction to N of a weak homomorphism.

We need a couple of preliminary results before beginning the proof. The first observation is essential to our efforts.

Lemma 5.2. *Let S be a Sylow p-subgroup of a finite group H.*

(a) *The inclusion map $N_H(S) \to H$ induces an isomorphism*

$$N_H(S)/N_H(S) \cap \Diamond(H) \cong H/\Diamond(H).$$

(b) *For any group homomorphism $\phi : N_H(S) \to k^\times$ having $N_H(S) \cap \Diamond(H)$ in its kernel, there exists a unique group homomorphism $\psi : H \to k^\times$ whose restriction to $N_H(S)$ is equal to ϕ.*

Proof. By the definition of $\diamondsuit(H)$, the Sylow p-subgroup S is contained in $\diamondsuit(H)$. The Frattini argument yields $N_H(S)\diamondsuit(H) = H$, proving (a). Then (b) follows from (a). $\square$

Lemma 5.3 (Burnside's fusion theorem). *Let S be an abelian Sylow p-subgroup of G and suppose that R is a nontrivial subgroup of $S \cap {}^gS$.*

(a) *There exists $c \in C_G(R)$ and $n \in N_G(S)$ such that $g = cn$.*

(b) *If $g = c'n'$ with $c' \in C_G(R)$ and $n' \in N_G(S)$, then there exists $d \in N_G(S) \cap C_G(R)$ such that $c' = cd$ and $n' = d^{-1}n$.*

Proof. Statement (a) is essentially Burnside's theorem. For the proof, observe that S and ${}^gS$ are Sylow p-subgroups of $C_G(R)$, hence conjugate by an element $c \in C_G(R)$. Then $n = c^{-1}g$ normalizes S. Statement (b) follows by defining $d = c^{-1}c' = nn'^{-1}$. $\square$

The essence of the proof of Theorem 5.1 is the following:

Proposition 5.4. *Suppose that a Sylow p-subgroup S of G is abelian, and let $N = N_G(S)$. Let $\chi : N \to k^\times$ be a homomorphism whose kernel contains $\rho^2(S)$. Then there is a unique weak homomorphism $\theta : G \to k^\times$ whose restriction to N is equal to χ.*

Proof. Let Q be a nontrivial subgroup of S, and let $H = C_G(Q)$, the centralizer of Q in G. Since S is abelian, $S \subseteq H \subseteq N_G(Q)$. Clearly, $\diamondsuit(H) = [H, H]S \subseteq \diamondsuit(N_G(Q))$, hence χ vanishes on $N_H(S) \cap \diamondsuit(H) \subseteq \rho^2(S)$ by assumption. By Lemma 5.2, there exists a unique group homomorphism

$$\psi_Q : H = C_G(Q) \longrightarrow k^\times$$

that coincides with χ on the subgroup $N_H(S) = N_G(S) \cap C_G(Q) = N \cap H$.

We define $\theta : G \to k^\times$ by the following rule. First, we set $\theta(g) = 1$ if $S \cap {}^gS = 1$. If $S \cap {}^gS \neq \{1\}$, we use Lemma 5.3 and write $g = cn$, with $c \in C_G(S \cap {}^gS)$ and $n \in N_G(S)$. Then let

$$\theta(g) = \psi_{S\cap {}^gS}(c)\chi(n).$$

In order to prove that θ is well defined, we must consider another decomposition $g = c'n'$, with $c' \in C_G(S \cap {}^gS)$ and $n' \in N_G(S)$, and show that the algorithm produces the same result for $\theta(g)$. In fact, we prove more: that the algorithm produces the same result even if we replace $S \cap {}^gS$ by a proper nontrivial subgroup. That is, assume that R is any nontrivial subgroup of $S \cap {}^gS$ and write $g = c'n'$, with $c' \in C_G(R)$ and $n' \in N_G(S)$. We claim that

$$\psi_{S\cap {}^gS}(c)\chi(n) = \psi_R(c')\chi(n').$$

In order to prove the claim, we observe that $C_G(S \cap {}^gS) \subseteq C_G(R)$ because $R \subseteq S \cap {}^gS$. So we have two decompositions $g = cn = c'n'$ with $c, c' \in C_G(R)$ and $n, n' \in N_G(S)$. By Lemma 5.3, there exists $d \in N_G(S) \cap C_G(R)$ such that $c' = cd$ and $n' = d^{-1}n$. Thus,

$$\psi_R(c')\chi(n') = \psi_R(cd)\chi(d^{-1}n) = \psi_R(c)\psi_R(d)\chi(d)^{-1}\chi(n) = \psi_R(c)\chi(n)$$

because $d \in N_G(S) \cap C_G(R)$ and we know that ψ_R and χ coincide on $N_G(S) \cap C_G(R)$. We next observe that the uniqueness of $\psi_{S\cap {}^gS} : C_G(S \cap {}^gS) \to k^\times$ implies that it must be equal to the restriction to $C_G(S \cap {}^gS)$ of the group homomorphism $\psi_R : C_G(R) \to k^\times$. Therefore $\psi_R(c) = \psi_{S\cap {}^gS}(c)$, completing the proof of the claim.

Our next task is to prove that θ is a weak homomorphism. If $s \in S$, then $S \cap {}^sS = S$ and we have $C_G(S) \subseteq N_G(S)$. We use the decomposition $s = 1 \cdot s$ with $1 \in C_G(S)$ and $s \in N_G(S)$ and we get $\theta(s) = \chi(s)$. But χ vanishes on S because $S \subseteq \Diamond(N)$. Therefore $\theta(s) = 1$, proving the first condition for a weak homomorphism. The second condition is obvious, since $\theta(g) = 1$ if $S \cap {}^gS = 1$, by definition.

Assume that $a, b \in G$ with $S \cap {}^aS \cap {}^{ab}S \neq \{1\}$. We must show that $\theta(ab) = \theta(a)\theta(b)$. Let $R = S \cap {}^aS \cap {}^{ab}S$, and note that $R \subseteq S \cap {}^aS$. Using what we have proved, we can write

$$\theta(a) = \psi_R(c)\chi(n), \quad \text{where } a = cn,\ c \in C_G(R),\ n \in N_G(S).$$

Next notice that $R \subseteq {}^aS \cap {}^{ab}S$, so that $R^a \subseteq S \cap {}^bS$. We write

$$\theta(b) = \psi_{R^a}(d)\chi(m), \quad \text{where } b = dm,\ d \in C_G(R^a),\ m \in N_G(S).$$

It follows that ${}^ab = {}^ad\,{}^am$, with ${}^ad \in C_G(R)$, ${}^am \in N_G({}^aS)$. Therefore

$$ab = {}^aba = {}^ad\,{}^ama = {}^adcc^{-1}am = {}^adcnm,$$

because $c^{-1}a = n$. But ${}^adc \in C_G(R)$, because $d \in C_G(R^a)$ and $nm \in N_G(S)$. So we obtain

$$\theta(ab) = \psi_R({}^adc)\chi(nm) = \psi_R({}^ad)\psi_R(c)\chi(n)\chi(m) = \psi_R(c)\chi(n)\psi_R({}^ad)\chi(m).$$

Now we claim that $\psi_R({}^ad) = \psi_{R^a}(d)$. Writing $a = cn$, we first find that

$$\psi_R({}^ad) = \psi_R(c\,{}^ndc^{-1}) = \psi_R(c)\psi_R({}^nd)\psi_R(c^{-1}) = \psi_R({}^nd),$$

because $k^\times$ is commutative. Moreover, writing $\mathrm{conj}_n(x) = nxn^{-1}$, we observe that the composite

$$C_G(R^a) = C_G(R^n) \xrightarrow{\mathrm{conj}_n} C_G(R) \xrightarrow{\psi_R} k^\times$$

has a restriction to $N_G(S)\cap C_G(R^n)$ equal to

$$N_G(S)\cap C_G(R^n) \xrightarrow{\text{conj}_n} N_G(S)\cap C_G(R) \xrightarrow{\chi} k^\times .$$

Hence, it is equal to the map

$$N_G(S)\cap C_G(R^n) \xrightarrow{\chi} k^\times \ ,$$

because $n \in N_G(S)$ and $\chi(nxn^{-1}) = \chi(n)\chi(x)\chi(n^{-1}) = \chi(x)$ by commutativity of $k^\times$. It follows that $\psi_R \circ \text{conj}_n : C_G(R^n) \to k^\times$ is the unique extension of $\chi : N_G(S)\cap C_G(R^n) \to k^\times$, hence equal to the homomorphism $\psi_{R^n} : C_G(R^n) \to k^\times$. In other words, $\psi_R({}^n d) = \psi_R(\text{conj}_n(d)) = \psi_{R^n}(d)$. Finally,

$$\psi_R({}^a d) = \psi_R({}^n d) = \psi_{R^n}(d) = \psi_{R^a}(d),$$

as claimed.

Returning to the computation of $\theta(ab)$, we find

$$\theta(ab) = \psi_R(c)\chi(n)\psi_R({}^a d)\chi(m) = \psi_R(c)\chi(n)\psi_{R^a}(d)\chi(m) = \theta(a)\theta(b).$$

This completes the proof of the third condition for a weak homomorphism. Thus $\theta : G \to k^\times$ is a weak homomorphism, and we have proved the proposition. □

Proof of Theorem 5.1. To prove (a), we recall that any weak homomorphism on N is a homomorphism. The image under the restriction map $\text{Res}^G_N : A(G) \to A(N)$ of any element of $A(G)$ must be a homomorphism with $\rho^2(S)$ in its kernel. The previous proposition says that the restriction map must be injective and surjective onto this subset.

Statement (b) follows immediately from (a) and the injectivity of the restriction map (Proposition 3.4). The proof of (c) — that is, the equality $\rho^2(S) = \rho^\infty(S)$ — follows from (a) and the fact that $\rho^\infty(S)$ is in the kernel of any weak homomorphism, by Theorem 4.3. □

All of the experimental evidence suggests that something like Theorem 5.1 should be true in general. Hence, we suggest the following question:

Question 5.5. *Suppose that G is any finite group with Sylow p-subgroup S. Let $N = N_G(S)$, and let $J \subseteq N$ be the intersection of the kernels of all one-dimensional kN-modules U such that the Green correspondent of U is an endotrivial kG-module. Is $J = \rho^\infty(S)$?*

While it might be possible to prove an affirmative answer by means similar to those in the proof of Theorem 5.1, we should point out that at least two difficulties arise. The first is that the fusion theorem of Burnside does not hold in greater generality. It would have to be replaced with something like Alperin's fusion theorem, whose conditions are more complicated. In addition, as we see in Section 7, even a stringent assumption such as control of fusion by the normalizer of the Sylow subgroup S does not readily lead to a generalization. Other assumptions seem to be necessary for the proof.

A second difficulty generalizing Theorem 5.1 is that the equality $\rho^2(S) = \rho^\infty(S)$ does not hold in general. An example is $G_2(5)$ for k a field of characteristic 3. Computer calculations using Magma [Bosma and Cannon 1996] show that $\rho^3(S) = \rho^\infty(S) = N_G(S)$, but $\rho^2(S)$ is a proper subgroup of $N_G(S)$.

6. The cyclic case

If a Sylow p-subgroup S of G is cyclic, then the structure of $K(G)$ is determined in Theorem 3.6 of [Mazza and Thévenaz 2007]. In this section, we show that this result can be recovered using Theorem 5.1. We prove the following:

Theorem 6.1. *Suppose that a Sylow p-subgroup S of G is cyclic. Let Z be the unique subgroup of S of order p. Then $K(G) \cong K(N_G(Z)) \cong \big(N_G(Z)/\diamondsuit(N_G(Z))\big)^*$.*

Proof. For any subgroup Q such that $Z \subseteq Q \subseteq S$, we have that $N_G(Q) \subseteq N_G(Z)$. Hence $\diamondsuit(N_G(Q)) \subseteq \diamondsuit(N_G(Z))$, and

$$\rho^2(S) = N_G(S) \cap \diamondsuit(N_G(Z)).$$

By Lemma 5.2 applied to the subgroup $H = N_G(Z)$, we have an isomorphism

$$N_G(S)/\rho^2(S) = N_G(S)/N_G(S) \cap \diamondsuit(N_G(Z)) \cong N_G(Z)/\diamondsuit(N_G(Z))\,.$$

Since $K(G) \cong \big(N_G(Z)/\rho^2(S)\big)^*$ by Theorem 5.1, we obtain

$$K(G) \cong \big(N_G(Z)/\diamondsuit(N_G(Z))\big)^*,$$

as required. This is also isomorphic to $K(N_G(Z))$, by Proposition 3.3. $\square$

The isomorphism $K(G) \cong K(N_G(Z))$ actually follows directly from the isomorphism $T(G) \cong T(N_G(Z))$, which is a consequence of the fact that $N_G(Z)$ is strongly p-embedded in G (see Lemma 3.5 in [Mazza and Thévenaz 2007]).

7. Control of fusion

Assume that the normalizer $N_G(S)$ of a Sylow p-subgroup S of G controls p-fusion. One may wonder if Theorem 5.1 still holds under this assumption. The analysis of the proof shows that we need more, as follows.

Theorem 7.1. *Suppose that the normalizer $N = N_G(S)$ of a Sylow p-subgroup S of G controls p-fusion. Assume, in addition, that, for any subgroup Q of S, we have $N_H(S)\diamondsuit(H) = H$, where $H = C_G(Q)$ and $N_H(S) = N_G(S) \cap H$. Then the following hold*:

(a) *The image of the restriction map* $\mathrm{Res}^G_{N_G(S)} : A(G) \to A(N_G(S))$ *consists exactly of all group homomorphisms* $N_G(S) \to k^\times$ *having* $\rho^2(S)$ *in their kernel.*

(b) $K(G) \cong A(G) \cong (N_G(S)/\rho^2(S))^*$.

(c) $\rho^2(S) = \rho^\infty(S)$.

Proof. The proof is exactly the same as that of Theorem 5.1, with the following observations. The assumption on each group $H = C_G(Q)$ implies that the conclusions of Lemma 5.2 hold. Thus the use of Lemma 5.2 remains valid. More precisely, for any group homomorphism $\chi : N_G(S) \to k^\times$ vanishing on $N_H(S) \cap \diamondsuit(H)$, there exists a unique group homomorphism $\psi_Q : C_G(S) \to k^\times$ which coincides with χ on the subgroup $N_H(S) = N_G(S) \cap H$. Here S is not necessarily contained in H (while it is when S is abelian), so the Frattini argument cannot be applied as it was in Lemma 5.2. However, our assumption allows us to make the argument work.

On the other hand, the assumption on control of fusion means exactly that the conclusions of Lemma 5.3 hold. Thus, the use of Lemma 5.3 remains valid, and the whole proof goes through. □

8. Examples

If H is a strongly p-embedded subgroup of G, then any one-dimensional representation of H has a Green correspondent which is endotrivial (see Proposition 2.8 and Remark 2.9 in [Carlson et al. 2006]). This fact was used to produce torsion endotrivial modules of dimension greater than one in various cases, in particular for groups of Lie type of rank one in the defining characteristic (Proposition 5.2 in [Carlson et al. 2006]) and for groups with a cyclic Sylow p-subgroup (Lemma 3.5 in [Mazza and Thévenaz 2007]).

However, there are other cases when torsion endotrivial modules of dimension greater than one occur. Several examples are given in [Lassueur and Mazza 2015] for various sporadic groups. The purpose of this section is to provide two explicit examples for classical groups with an abelian Sylow p-subgroup. We first start by an easy case.

Example 8.1 (PSL(2, q) in characteristic 2). Let $G = \mathrm{PSL}(2, q)$ in characteristic 2 and assume that $q \equiv 3$ or 5 modulo 8, so that a Sylow 2-subgroup S of G is a Klein four-group. Then $C_G(S) = S$ has index 3 in $N_G(S)$, hence $\diamondsuit(N_G(S)) = S$. Any subgroup C of order 2 satisfies $N_G(C) = C_G(C)$ and

$$S \subseteq N_G(S) \cap \diamondsuit(N_G(C)) \subseteq N_G(S) \cap N_G(C) = N_G(S) \cap C_G(C) = S.$$

Therefore $N_G(S) \cap \diamondsuit(N_G(C)) = S$, and it follows that $\rho^2(S) = \rho^1(S) = S$. By Theorem 5.1, $K(G)$ is the dual group of $N_G(S)/S$, which is cyclic of order 3. Hence $TT(G) = K(G) \cong \mathbb{Z}/3\mathbb{Z}$.

For our second example, we compute the torsion part of the group of endotrivial modules over the group $G = \mathrm{PSL}(3, q)$ in characteristic 3, in the case that $q \equiv 4$ or 7 modulo 9. In this case, the Sylow 3-subgroup of G is elementary abelian of order 9. The point is to show that the torsion subgroup $TT(G) = K(G)$ of the group of endotrivial module is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$, which has three nontrivial elements.

For notation, we use an overline to indicate the class in G of an element in $H = \mathrm{SL}(3, q)$. Let ζ denote a cubed root of unity in $\mathbb{F}_q$. We fix the following elements of H:

$$a = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta^2 \end{bmatrix}, \quad x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad u = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \zeta & \zeta^2 \\ 1 & \zeta^2 & \zeta \end{bmatrix}, \quad v = \begin{bmatrix} \zeta & 1 & 1 \\ \zeta & \zeta & \zeta^2 \\ 1 & \zeta & 1 \end{bmatrix}.$$

Then it is not difficult to verify the following:

Lemma 8.2. *The subgroup $\langle a, x\rangle$ is a Sylow 3-subgroup of $H = \mathrm{SL}(3, q)$, and hence the Sylow 3-subgroup of G is $S = \langle \bar{a}, \bar{x}\rangle$. In addition, we have the relations:*

(1) $u^{-1}xu = a$,

(2) $u^{-1}au = x^{-1}$,

(3) $v^{-1}xv = a^2v$,

(4) $v^{-1}av = \zeta(ax)^{-1}$.

In particular, the elements $\bar{u}$ and $\bar{v}$ are in the normalizer of S, and each acts on S by exchanging the four maximal subgroups in pairs. The commutator $\sigma = u^{-1}v^{-1}uv$ acts on S by inverting $\bar{a}$ and $\bar{x}$, and hence also inverts every nonidentity element.

Now we consider the centralizers and normalizers. Let $z \in H$ be the scalar matrix with nonzero entries equal to ζ. Then $\langle z\rangle$ is the kernel of the natural homomorphism of H onto G. A general principle here is that, if y is an element of H such that $\bar{y}$ commutes with $\bar{x}$, then, for some j, $a^j y$ is in the centralizer of x, and the same holds with a and x exchanged.

It is easy to see that the centralizer of a in H is the Levi subgroup L of all diagonal matrices of determinant one. Then the normalizer of $\langle a, z\rangle$ is the normalizer of L which is generated by L, x and the element σ, the commutator of u and v. The centralizer of x, which has the same order as that of a, consists of all elements of the form

$$\begin{bmatrix} c & d & e \\ e & c & d \\ d & e & c \end{bmatrix}$$

having determinant one (that is, $c^3 + d^3 + e^3 - 3cde = 1$). Then, the normalizer of $\langle x, z\rangle$ is generated by this centralizer, a and σ.

Thus we can prove the following:

Proposition 8.3. *The Sylow* 3*-subgroup S is self-centralizing. Its normalizer is generated by $\bar{a}$, $\bar{x}$, $\bar{u}$, $\bar{v}$ and $\bar{\sigma}$.*

Proof. We can see that the centralizer of S is generated by the classes $\bar{a}$, $\bar{x}$ and the classes of the intersection of the centralizers of a and x in H. However, this intersection consists only of the elements of $\langle z\rangle$. The elements $\bar{u}$ and $\bar{v}$ permute the maximal subgroups of S transitively. So suppose that $\bar{y}$ is an element of the normalizer of S and y a preimage of $\bar{y}$ in H. By replacing y by its product with a power of u and/or a power of v we may assume that $\bar{y}$ normalizes $\langle\bar{a}\rangle$. By replacing $\bar{y}$ by its product with σ, if necessary, we may assume that $\bar{y}$ centralizes $\bar{a}$. Multiplying y by a power of x, if necessary, we may assume that y has the form

$$\begin{bmatrix} r & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & t \end{bmatrix},$$

where $rst = 1$. Thus we have that

$$yxy^{-1} = \begin{bmatrix} 0 & r^2t & 0 \\ 0 & 0 & rs^2 \\ st^2 & 0 & 0 \end{bmatrix}.$$

The point of this is that $\bar{y}\bar{x}\bar{y}^{-1}$ must be one of the elements $\bar{x}$, $\bar{a}\bar{x}$ or $\bar{a}^2\bar{x}$ and *cannot* be $\bar{x}^2$, $\bar{a}\bar{x}^2$ or $\bar{a}^2\bar{x}^2$. Because y centralizes a, conjugation by $\bar{y}$ is an automorphism of order either one or three on S. Order 3 is not possible. Consequently, $\bar{y}\bar{x}\bar{y}^{-1} = \bar{x}$ and $\bar{y}$ centralizes S. This proves the proposition. □

Now we are ready for the main theorem.

Theorem 8.4. *Assume that $G = \mathrm{PSL}(3, q)$, where q is congruent to* 4 *or* 7 *modulo* 9*.*

(a) *Let S be a Sylow* 3*-subgroup of G. Then $\rho^\infty(S) = \rho^1(S) = [N_G(S), N_G(S)]$.*

(b) $TT(G) = K(G) \cong (\mathbb{Z}/2\mathbb{Z})^2$.

Proof. (a) First, $\rho^1(S) = \diamond(N_G(S)) = [N_G(S), N_G(S)]$ by definition and the fact that $[N_G(S), N_G(S)]$ has index prime to 3 in $N_G(S)$ (namely index 4). We also have that, for any subgroup U of order 3 in S, $N_G(U) \cap N_G(S) \subseteq [N_G(S), N_G(S)]$. Consequently, again from the definition, we have that $\rho^n(S) = [N_G(S), N_G(S)]$ for all n.

(b) Since a Sylow 3-subgroup is abelian of rank two, $TT(G) = K(G)$ is isomorphic to the dual group of $N_G(S)/\rho^2(S)$ by Theorem 5.1. By (a), $\rho^2(S) = [N_G(S), N_G(S)]$ and, by Proposition 8.3, $N_G(S)/[N_G(S), N_G(S)] \cong (C_2)^2$, a Klein four-group, generated by the classes of $\bar{u}$ and $\bar{v}$. Its dual group (in additive notation) is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$. □

Remark 8.5. In the case that $q = 4$, the normalizer of the Sylow subgroup S of G is strongly 3-embedded, and the result of the theorem could be deduced from that fact. In all other cases, $N_G(S)$ is not strongly 3-embedded and it is not strongly 3-embedded in any subgroup of G that properly contains it.

Finally, it is not difficult to perform the computations of the subgroups $\rho^i(Q)$ for all subgroups Q of the Sylow subgroup S of G on a computer using a standard computer algebra system. From this computation, the structure of $K(G) = TT(G)$ can in many cases be deduced using Theorem 5.1 or something similar. Below are a few calculations using Magma [Bosma and Cannon 1996]. In most cases, only a few seconds of computing time was required. The computing time depends on such things as the size of the permutation representation of G and the number of subgroups of S. Here we list only groups where $K(G)$ is not trivial. The results should be compared with those of [Lassueur et al. 2013; Lassueur and Malle 2015; Lassueur and Mazza 2015]. The notation for the groups is the Atlas notation.

TT(G)	Group	Characteristic p
$\mathbb{Z}/2\mathbb{Z}$	M_{23}, Ru	3
	J_2, Suz, 2Suz, 6Suz, Fi_{22}, Fi_{23}	5
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$	M_{11}, M_{22}, M_{23}, HS	3
$\mathbb{Z}/4\mathbb{Z}$	2Ru	3
	Co_3, Sz_{32}	5
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$	$2M_{22}$, $4M_{22}$	3
$\mathbb{Z}/8\mathbb{Z}$	McL	5
$\mathbb{Z}/24\mathbb{Z}$	3McL	5

Acknowledgment

The authors are grateful to Paul Balmer and Caroline Lassueur for several useful discussions.

References

[Balmer 2013] P. Balmer, "Modular representations of finite groups with trivial restriction to Sylow subgroups", *J. Eur. Math. Soc.* (*JEMS*) **15**:6 (2013), 2061–2079. MR 3120735 Zbl 1291.20009

[Balmer 2015] P. Balmer, "Stacks of group representations", *J. Eur. Math. Soc.* (*JEMS*) **17**:1 (2015), 189–228. MR 3312406

[Bosma and Cannon 1996] W. Bosma and J. Cannon, "Handbook on Magma functions", University of Sydney, 1996, http://magma.maths.usyd.edu.au/magma/handbook/.

[Carlson and Thévenaz 2004] J. F. Carlson and J. Thévenaz, "The classification of endo-trivial modules", *Invent. Math.* **158**:2 (2004), 389–411. MR 2005e:20013 Zbl 1110.20004

[Carlson and Thévenaz 2005] J. F. Carlson and J. Thévenaz, "The classification of torsion endo-trivial modules", *Ann. of Math.* (2) **162**:2 (2005), 823–883. MR 2006f:20012 Zbl 1110.20005

[Carlson et al. 2006] J. F. Carlson, N. Mazza, and D. K. Nakano, "Endotrivial modules for finite groups of Lie type", *J. Reine Angew. Math.* **595** (2006), 93–119. MR 2007c:20023 Zbl 1106.20007

[Carlson et al. 2009] J. F. Carlson, N. Mazza, and D. K. Nakano, "Endotrivial modules for the symmetric and alternating groups", *Proc. Edinb. Math. Soc.* (2) **52**:1 (2009), 45–66. MR 2009k:20023 Zbl 1165.20005

[Carlson et al. 2010] J. F. Carlson, D. J. Hemmer, and N. Mazza, "The group of endotrivial modules for the symmetric and alternating groups", *Proc. Edinb. Math. Soc.* (2) **53**:1 (2010), 83–95. MR 2011d:20017 Zbl 1217.20004

[Carlson et al. 2011] J. F. Carlson, N. Mazza, and J. Thévenaz, "Endotrivial modules for p-solvable groups", *Trans. Amer. Math. Soc.* **363**:9 (2011), 4979–4996. MR 2012f:20026 Zbl 1250.20007

[Carlson et al. 2013] J. F. Carlson, N. Mazza, and J. Thévenaz, "Endotrivial modules over groups with quaternion or semi-dihedral Sylow 2-subgroup", *J. Eur. Math. Soc.* (*JEMS*) **15**:1 (2013), 157–177. MR 2998832 Zbl 1276.20007

[Carlson et al. 2014a] J. F. Carlson, N. Mazza, and D. K. Nakano, "Endotrivial modules for the general linear group in a nondefining characteristic", *Math. Z.* **278**:3-4 (2014), 901–925. MR 3278897 Zbl 06384008

[Carlson et al. 2014b] J. F. Carlson, N. Mazza, and J. Thévenaz, "Torsion-free endotrivial modules", *J. Algebra* **398** (2014), 413–433. MR 3123775 Zbl 06334760

[Kawata 1993] S. Kawata, "On Auslander–Reiten components for certain group modules", *Osaka J. Math.* **30**:2 (1993), 137–157. MR 95c:20008 Zbl 0819.20007

[Lassueur and Malle 2015] C. Lassueur and G. Malle, "Simple endo-trivial modules for the linear, unitary and exceptional groups", preprint, 2015. arXiv 1501.00400

[Lassueur and Mazza 2015] C. Lassueur and N. Mazza, "Endotrivial modules for the sporadic simple groups and their covers", *J. Pure Appl. Algebra* (online publication February 2015).

[Lassueur et al. 2013] C. Lassueur, G. Malle, and E. Schulte, "Simple endotrivial modules for quasi-simple groups", *J. Reine Angew. Math.* (online publication November 2013).

[Mazza 2007] N. Mazza, "The group of endotrivial modules in the normal case", *J. Pure Appl. Algebra* **209**:2 (2007), 311–323. MR 2007j:20010 Zbl 1117.20009

[Mazza and Thévenaz 2007] N. Mazza and J. Thévenaz, "Endotrivial modules in the cyclic case", *Arch. Math.* (*Basel*) **89**:6 (2007), 497–503. MR 2009b:20015 Zbl 1155.20009

[Navarro and Robinson 2012] G. Navarro and G. R. Robinson, "On endo-trivial modules for p-solvable groups", *Math. Z.* **270**:3-4 (2012), 983–987. MR 2892933 Zbl 1242.20004

[Robinson 2011] G. R. Robinson, "On simple endotrivial modules", *Bull. Lond. Math. Soc.* **43**:4 (2011), 712–716. MR 2012g:20022 Zbl 1234.20003

[de Smit and Lenstra 1997] B. de Smit and H. W. Lenstra, Jr., "Explicit construction of universal deformation rings", pp. 313–326 in *Modular forms and Fermat's last theorem* (Boston, MA, 1995), edited by G. Cornell et al., Springer, New York, 1997. MR 1638482 Zbl 0907.13010

Communicated by David Benson
Received 2014-11-05 Revised 2015-01-28 Accepted 2015-02-27

jfc@math.uga.edu *Department of Mathematics, University of Georgia, Athens, GA 30602, United States*

jacques.thevenaz@epfl.ch *Section de mathématiques, EPFL, Station 8, CH-1015 Lausanne, Switzerland*

mathematical sciences publishers msp

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 9 No. 3 2015

[Hurwitz monodromy and full number fields](511)
David P. Roberts and Akshay Venkatesh — 511

The characteristic polynomial of the Adams operators on graded connected Hopf algebras — 547
Marcelo Aguiar and Aaron Lauve

Secant spaces and syzygies of special line bundles on curves — 585
Marian Aprodu and Edoardo Sernesi

Complex group algebras of the double covers of the symmetric and alternating groups — 601
Christine Bessenrodt, Hung Ngoc Nguyen, Jørn B. Olsson and Hung P. Tong-Viet

Fano schemes of determinants and permanents — 629
Melody Chan and Nathan Ilten

Triple intersection formulas for isotropic Grassmannians — 681
Vijay Ravikumar

On the basepoint-free theorem for log canonical threefolds over the algebraic closure of a finite field — 725
Diletta Martinelli, Yusuke Nakamura and Jakub Witaszek

The torsion group of endotrivial modules — 749
Jon F. Carlson and Jacques Thévenaz