Geometry & Topology

Volume 25 (2021)

Factorization statistics and bug-eyed configuration spaces

DAN PETERSEN
PHILIP TOSTESON

msp

msp

Geometry & Topology 25 (2021) 3691–3723

Factorization statistics and bug-eyed configuration spaces

DAN PETERSEN
PHILIP TOSTESON

A recent theorem of Hyde proves that the factorization statistics of a random polynomial over a finite field are governed by the action of the symmetric group on the configuration space of n distinct ordered points in $\mathbb{R}^3$. Hyde asked whether this result could be explained geometrically. We give a geometric proof of Hyde's theorem as an instance of the Grothendieck–Lefschetz trace formula applied to an interesting, highly nonseparated algebraic space. An advantage of our method is that it generalizes uniformly to any Weyl group. In the process we study certain non-Hausdorff models for complements of hyperplane arrangements, first introduced by Proudfoot.

11T06, 14F20, 14N20, 55R80; 14A20, 14G15

1 Introduction

1.1 What, why, and how: factorization statistics and configuration spaces

1.1.1 Let $\chi\colon S_n \to \mathbb{C}$ be a class function. We consider χ as a function on the set of integer partitions of n. Let $\mathrm{Pol}_n(\mathbb{F}_q)$ denote the set of monic polynomials of degree n over the finite field $\mathbb{F}_q$. If $f \in \mathrm{Pol}_n(\mathbb{F}_q)$ has the factorization $f = \prod_i f_i$ into irreducible factors over $\mathbb{F}_q$, then the degrees of the f_i form a partition of n, and we let $\chi(f)$ denote the value of χ on the resulting partition. If f is squarefree, we may also define $\chi(f)$ as follows: the Frobenius automorphism Frob_q acts by permuting the roots of f, the resulting permutation is well defined up to conjugation, and we define $\chi(f)$ as the value of χ on this permutation.

1.1.2 We will be interested in sums of the form $\sum_{f \in \mathrm{Pol}_n(\mathbb{F}_q)} \chi(f)$. These sums often have number-theoretic significance: we may for example compute the average number of irreducible factors, the variance, the probability of having more linear factors than quadratic factors, and so on, of a random degree-n polynomial over $\mathbb{F}_q$. In general, we may compute the expected value of any *factorization statistic*, by which we mean any quantity $\chi(f)$ that only depends on the degrees of the irreducible factors of f. The subject of this paper is a recent theorem of Hyde [14] that relates these statistics to point configurations in $\mathbb{R}^3$.

Published: 25 January 2022 DOI: 10.2140/gt.2021.25.3691

1.1.3 To state the theorem we need the following notation. If V is a representation of the finite group G, then we denote by $\operatorname{ch} V$ its character. We denote by $\langle -, - \rangle_G$ the standard inner product of class functions, for which the characters of irreducible complex representations form an orthonormal basis.

1.1.4 Theorem (Hyde) *Let* $\mathrm{PConf}_n(\mathbb{R}^3)$ *be the configuration space of* n *distinct ordered points in* $\mathbb{R}^3$. *The sum of* χ *over* $\mathrm{Pol}_n(\mathbb{F}_q)$ *is given by*

$$\sum_{f \in \mathrm{Pol}_n(\mathbb{F}_q)} \chi(f) = q^n \sum_{i \geq 0} q^{-i} \langle \chi, \operatorname{ch} H^{2i}(\mathrm{PConf}_n(\mathbb{R}^3), \mathbb{Q}) \rangle_{\mathrm{S}_n}.$$

1.1.5 Remark The odd-degree cohomology groups of $\mathrm{PConf}_n(\mathbb{R}^3)$ vanish, which explains why Hyde's theorem only involves the even cohomology.

1.1.6 It is natural to ask:

What is the meaning of this formula?

If you have some experience with étale cohomology, your first thought is perhaps also:

This ***must*** *be an instance of the Grothendieck–Lefschetz trace formula.*

But how? Why should the two spaces Pol_n and $\mathrm{PConf}_n(\mathbb{R}^3)$ have any kind of relationship with each other? And isn't $\mathrm{PConf}_n(\mathbb{R}^3)$ very far from being a complex algebraic variety? Hyde's method of proof does not answer these questions: he calculates explicitly a generating function for all factorization statistics, considered as an element of the ring of symmetric functions, and then compares with the character of $H^*(\mathrm{PConf}_n(\mathbb{R}^d), \mathbb{Q})$, which is well understood.

1.1.7 In this paper we give a fully geometric proof of Hyde's result, and explain why there should be a connection between factorization statistics and point configurations in $\mathbb{R}^3$. In a nutshell, we will consider a certain non-Hausdorff scheme X_n over $\mathbb{Z}$ with an action of S_n, satisfying the following two properties: (i) The analytification $X_n(\mathbb{C})$ has the S_n–equivariant weak homotopy type of $\mathrm{PConf}_n(\mathbb{R}^3)$. (ii) The stack quotient $[X_n/\mathrm{S}_n]$ — which happens to be an algebraic space — can be understood as the badly non-Hausdorff space parametrizing degree-n monic polynomials $p(t)$, together with a choice of factorization of $p(t)$ into irreducibles. In particular we have a bijection of sets $[X_n/\mathrm{S}_n](\mathbb{F}_q) = \mathrm{Pol}_n(\mathbb{F}_q)$. Applying the Grothendieck–Lefschetz trace formula with twisted coefficients to $[X_n/\mathrm{S}_n]$ produces Hyde's formula.

1.1.8 A reader may object to the claim in the preceding paragraph that the space parametrizing a degree-n polynomial with a choice of decomposition into irreducible factors is not Hausdorff. Surely every polynomial has a *unique* factorization into irreducibles? Surely the space of polynomials with a fixed such factorization is just the space of polynomials, ie affine n–space? But this is precisely the crux of the matter — the space of polynomials with a factorization into irreducibles is *not* affine n–space but rather something deceptively similar: it is a *bug-eyed cover* of affine n–space. A bug-eyed cover is a type of geometric peculiarity that occurs in the world of algebraic spaces: it is a space which fails to be Hausdorff not because it has "too many points" but rather because it has "too many tangent directions".

1.1.9 The nonclassical nature of the space $[X_n/\mathrm{S}_n]$, and the relationship to configurations of points in $\mathbb{R}^3$, can be understood geometrically in terms of the theory of foliations. (However, we will not use this perspective in the body of the paper.) Consider a foliated manifold $(M,\mathscr{F})$. There is a natural induced foliation on $\mathrm{PConf}_n(M)$, which descends to the quotient manifold $\mathrm{Conf}_n(M) := \mathrm{PConf}_n(M)/\mathrm{S}_n$. Now let us specialize to the product foliation of $\mathbb{R}^3$ whose leaves are the lines parallel to the z–axis. Then the leaf space of the induced foliation of $\mathrm{Conf}_n(\mathbb{R}^3)$ can be identified with the analytification of $[X_n/\mathrm{S}_n]$. In general the leaf space of a foliation is most naturally thought of as a Lie groupoid, the *holonomy groupoid*. In this case the holonomy groupoid has trivial isotropy groups, but is still not weakly equivalent to a manifold, not even a non-Hausdorff manifold. On the algebraic side this is reflected in the fact that the algebraic space $[X_n/\mathrm{S}_n]$ is not locally separated, which implies that it does not admit a well-defined analytification in the category of complex-analytic spaces. But if we consider $[X_n/\mathrm{S}_n]$ as a stack then we obtain a well-defined analytification as an étale groupoid in analytic spaces, weakly equivalent to the leaf space of the above foliation of $\mathrm{Conf}_n(\mathbb{R}^3)$.

1.1.10 The scheme X_n is a special case of a more general construction: for any arrangement A of hyperplanes in $\mathbb{R}^n$ there exists a scheme X_A with the property that the induced codimension-3 subspace arrangement in $(\mathbb{R}^3)^n$ is a fiber bundle over $X_A(\mathbb{C})$ with fiber $\mathbb{R}^n$ (and in particular they are weakly homotopy equivalent), and similarly the complement of the complexification of the arrangement in $\mathbb{C}^n$ is an $\mathbb{R}^n$–fiber bundle over the set of real points $X_A(\mathbb{R})$. The scheme X_A was introduced by Proudfoot [28] as a non-Hausdorff model for the complement of a complexified real hyperplane arrangement, although Proudfoot only considered the real points of X_A, not its set of complex points.

1.1.11 From our geometric approach, we will also see that Hyde's theorem is a special case of a formula valid for any Weyl group. This more general formula specializes to Hyde's theorem when the Weyl group W is S_n with its action on $\mathbb{A}^n$ given by permuting the coordinates.

1.1.12 Theorem *Let $W \subset \mathrm{GL}(n,\mathbb{Z})$ be a Weyl group acting by reflections on affine n–space. Let $\chi\colon W \to \mathbb{C}$ be a class function. Let M_W be the complement of the associated codimension-3 subspace arrangement in $(\mathbb{R}^3)^n$. For any finite field $\mathbb{F}_q$ we have*

$$\sum_{x \in (\mathbb{A}^n/W)(\mathbb{F}_q)} \chi(x) = q^n \sum_{i \geq 0} q^{-i} \langle \chi, \mathrm{ch}\, H^{2i}(M_W, \mathbb{Q}) \rangle_W .$$

1.1.13 Let us explain the notation $\chi(x)$ in the left-hand side of Theorem 1.1.12. If $x \in (\mathbb{A}^n/W)(\mathbb{F}_q)$, choose an arbitrary element $y \in \mathbb{A}^n(\overline{\mathbb{F}}_q)$ in the fiber of $\mathbb{A}^n \to \mathbb{A}^n/W$ over x. Since Frobenius acts preserving the fiber over x, there exists an element $w \in W$ such that $w \cdot y = \mathrm{Frob}(y)$. But w is not unique in general; the best we can say is that we get a well-defined coset $w \cdot \mathrm{Stab}(y) \subseteq W$, where $\mathrm{Stab}(y)$ is the stabilizer of y. But a stabilizer of a point is precisely what is called a *parabolic* subgroup of a reflection group, and it is a standard fact that any coset of a parabolic subgroup has a distinguished *minimal representative* (which depends on the choice of a system of positive roots). Up to conjugacy the minimal representative does not depend on the choice of y or the choice of system of positive roots, and we define $\chi(x)$ to be the value of χ on this conjugacy class.

1.1.14 For a general Weyl group we do not have an interpretation of Theorem 1.1.12 as concrete as the one in the type A case in terms of factorizations of polynomials. However, in the type B case there is one: we obtain a formula that relates the expected factorizations of monic *even* polynomials of degree $2n$ over finite fields to the action of the hyperoctahedral group on the cohomology of the complement of the type B subspace arrangement in $(\mathbb{R}^3)^n$.

1.2 Structure of the paper

1.2.1 This paper consists of two parts of rather different flavor.

In Section 2 we describe previous work, and how we were led to define the scheme X_n; we explain how Hyde's theorem follows by applying the Grothendieck–Lefschetz trace formula to $[X_n/\mathrm{S}_n]$; and finally we explain how to generalize the scheme X_n to a more

general scheme X_W for a Weyl group W, so that Grothendieck–Lefschetz applied to $[X_W/W]$ produces Theorem 1.1.12. This part of the paper is rather leisurely written, with few or no proofs. In particular we will, in the course of the arguments, use some properties of the schemes X_n and X_W without careful justifications, in order to not disrupt the flow of the text.

The careful justifications are supplied in Section 3, where we study the schemes X_n and X_W in detail. As mentioned previously these schemes are special cases of a more general construction due to Proudfoot, which associates a scheme X_A to a hyperplane arrangement A in $\mathbb{R}^n$. We show that the complement of the induced codimension-3 subspace arrangement in $(\mathbb{R}^3)^n$ is an $\mathbb{R}^n$–bundle over $X_A(\mathbb{C})$, and that the complement of the complexification is similarly fibered over $X_A(\mathbb{R})$; in the real case this is a theorem of Proudfoot. The combinatorial structure of the scheme X_A is independently interesting, and closely tied to the Salvetti complex. Using a theorem of Delucchi [7] and Lofano and Paolini [25], we construct an algebraic cell decomposition (also known as an affine paving) of X_A, which can be used to determine its cohomology as a Galois representation.

Acknowledgements We are grateful to Dustin Clausen, Emanuele Delucchi, Trevor Hyde, Dmitri Panov and David Rydh for comments and suggestions. Will Sawin has informed us that he has obtained a related geometric approach to Hyde's theorem. We also thank Michael Falk for informing us of Proudfoot's work.

Petersen acknowledges support by ERC-2017-STG 759082 and a Wallenberg Academy Fellowship. Tosteson acknowledges support by NSF Grant No. DMS-1903040.

2 The path to Hyde's theorem

2.1 The case of squarefree polynomials

2.1.1 It will be convenient to rewrite the right-hand side of Hyde's Theorem 1.1.4 as follows. Let V_χ be a rational representation of S_n and $\chi = \operatorname{ch} V_\chi$. If we can compute $\sum_{f\in \mathrm{Pol}_n(\mathbb{F}_q)} \chi(f)$ for any V_χ, then we can compute this sum for an arbitrary class function by linearity, so it suffices to prove this case of Hyde's theorem. If Y is any topological space then the finite-sheeted covering $\mathrm{PConf}_n(Y) \to \mathrm{Conf}_n(Y) = \mathrm{PConf}_n(Y)/\mathrm{S}_n$ is classified by a homomorphism $\pi_1(\mathrm{Conf}_n(Y)) \to \mathrm{S}_n$. Via this homomorphism, we

may think of V_χ as defining a local system $\mathbb{V}_\chi$ on $\mathrm{Conf}_n(Y)$ for any Y, and there is an isomorphism

$$H^k(\mathrm{Conf}_n(Y), \mathbb{V}_\chi) \cong H^k(\mathrm{PConf}_n(Y), \mathbb{Q}) \otimes_{S_n} V_\chi$$

for all k. In these terms we may state Hyde's formula equivalently in the form

$$\sum_{f \in \mathrm{Pol}_n(\mathbb{F}_q)} \chi(f) = q^n \sum_{i \geq 0} q^{-i} \dim H^{2i}(\mathrm{Conf}_n(\mathbb{R}^3), \mathbb{V}_\chi). \tag{1}$$

In this form the connection to Grothendieck–Lefschetz will become more transparent.

2.1.2 Before trying to find a geometric interpretation of Hyde's formula (1), it is instructive to first consider a similar but simpler case, namely the case where we restrict f to be *squarefree*. In this setting, the analogue of Hyde's theorem is

$$\sum_{\substack{f \in \mathrm{Pol}_n(\mathbb{F}_q) \\ f \text{ squarefree}}} \chi(f) = q^n \sum_{i \geq 0} (-1)^i q^{-i} \dim H^i(\mathrm{Conf}_n(\mathbb{C}), \mathbb{V}_\chi), \tag{2}$$

which *is* an instance of the Grothendieck–Lefschetz trace formula in a more direct manner. The formula (2) can be found in a paper of Church, Ellenberg and Farb [5]; as they point out, closely related statements were in the literature previously, eg [19]. In fact, Hyde was led to discover his formula by investigating what happens when the "squarefree" condition is dropped from the left-hand side of (2). We give a brief account of the proof of (2) here, and refer to [5] for a more detailed account. We will use the Grothendieck–Lefschetz formula in the form

$$\sum_{x \in X(\mathbb{F}_q)} \mathrm{Tr}(\Phi_q \,|\, \mathcal{F}_{\overline{x}}) = q^{\dim X} \sum_i (-1)^i \,\mathrm{Tr}(\Phi_q \,|\, H^i_{\text{ét}}(X_{\overline{\mathbb{F}}_q}, \mathcal{F})), \tag{3}$$

where X is a smooth finite type scheme over $\mathbb{F}_q$, $\mathcal{F}$ is a locally constant $\mathbb{Q}_\ell$–sheaf on X (where ℓ is a prime invertible in $\mathbb{F}_q$), $\mathcal{F}_{\overline{x}}$ is the stalk of $\mathcal{F}$ at a geometric point $\overline{x}$ over x, and Φ_q is the arithmetic Frobenius (whose eigenvalues are the inverses of those of the geometric Frobenius).

2.1.3 We apply Grothendieck–Lefschetz to the scheme $\mathrm{Conf}_n(\mathbb{A}^1)$ over $\mathbb{Z}$, which parametrizes squarefree monic degree-n polynomials. In the same way that the representation V_χ defined a local system $\mathbb{V}_\chi$ on $\mathrm{Conf}_n(Y)$ for any topological space Y, we obtain a locally constant ℓ–adic sheaf $\mathbb{V}_{\chi,\ell}$ on $\mathrm{Conf}_n(\mathbb{A}^1)$. Now recall that when the polynomial f is squarefree we may also define $\chi(f)$ as follows: the Frobenius automorphism acts by permuting the roots of f, the resulting permutation is well defined up to conjugation, and we define $\chi(f)$ as the value of χ on this conjugacy class. But the monodromy of the covering $\mathrm{PConf}_n(\mathbb{A}^1) \to \mathrm{Conf}_n(\mathbb{A}^1)$ is precisely given by the

action of S_n permuting the roots, which implies that $\chi(f) = \mathrm{Tr}(\Phi_q \,|\, (\mathbb{V}_{\chi,\ell})_{\bar{f}})$, and so the left-hand sides of (2) and (3) agree.

2.1.4 To identify the right-hand sides of (2) and (3) we need to prove that

$$H^i(\mathrm{Conf}_n(\mathbb{C}), \mathbb{V}_\chi) \otimes \mathbb{Q}_\ell \cong H^i_{\text{ét}}(\mathrm{Conf}_n(\mathbb{A}^1)_{\overline{\mathbb{F}}_q}, \mathbb{V}_{\chi,\ell})$$

for all i, or equivalently that there is an S_n–equivariant isomorphism

$$H^i(\mathrm{PConf}_n(\mathbb{C}), \mathbb{Q}_\ell) \cong H^i_{\text{ét}}(\mathrm{PConf}_n(\mathbb{A}^1)_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell);$$

we also need to know that all eigenvalues of Φ_q acting on $H^i_{\text{ét}}(\mathrm{PConf}_n(\mathbb{A}^1)_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell)$ equal q^{-i}. The first part can be deduced in all characteristics from the Artin comparison theorem and the fact that $\mathrm{PConf}_n(\mathbb{A}^1)$ is the complement of a simple normal crossing divisor in a smooth proper scheme over $\mathbb{Z}$ (using the Fulton–MacPherson compactification [11]). However, to understand the Galois action as well, it is more efficient to argue that both Betti and étale cohomology of the complement of an arrangement of hyperplanes can be computed by an identical deletion–restriction scheme; the latter argument implies in addition that the étale cohomology of the complement of any hyperplane arrangement is pure Tate of weight $2i$ in degree i, as needed. This argument is independently due to Kim [18] and Lehrer [24].

2.1.5 As we mentioned in Section 1.1.14, our more general formula Theorem 1.1.12 has an interpretation in the type B case in terms of expected factorizations of even polynomials of degree $2n$ over $\mathbb{F}_q$. There is also a "squarefree" Type B version of the formula, which relates the expected factorizations of squarefree even polynomials of degree $2n$ over $\mathbb{F}_q$ and the action of the hyperoctahedral group on the cohomology of the complement of the type B arrangement in $\mathbb{C}^n$. This formula is due to Jiménez Rolland and Wilson [16], and is proven using the Grothendieck–Lefschetz trace formula in much the same way as in the type A case.

2.1.6 The upshot is that in the squarefree case we have a good geometric understanding of the factorization statistics in terms of ℓ–adic cohomology, and we would like a similar geometric description of the factorization statistics for general polynomials.

2.2 The plot thickens: weighted statistics

2.2.1 We saw that to compute the factorization statistics of squarefree polynomials (2), one uses the cover $\mathrm{PConf}_n(\mathbb{A}^1) \to \mathrm{Conf}_n(\mathbb{A}^1)$. So at first it may seem that if we wish

to compute the factorization statistics for *all* polynomials, we need to replace the cover of the space of squarefree polynomials, $\mathrm{PConf}_n(\mathbb{A}^1) \to \mathrm{Conf}_n(\mathbb{A}^1)$, by the ramified cover of the space of all polynomials, $h\colon (\mathbb{A}^1)^n \to \mathbb{A}^n$. Even though h is not étale, we can still associate a sheaf on $\mathbb{A}^n$ to every irreducible representation V_χ of S_n, namely the sheaf

$$(h_*\mathbb{Q}_\ell)_\chi := \mathrm{Hom}_{\mathrm{S}_n}(V_\chi, h_*\mathbb{Q}_\ell).$$

Here V_χ is considered as a constant sheaf of $\mathbb{Q}_\ell[\mathrm{S}_n]$–modules on $\mathbb{A}^n$ (and should not be confused with the local system $\mathbb{V}_\chi$), so that $(h_*\mathbb{Q}_\ell)_\chi \otimes_{\mathbb{Q}} V_\chi$ is the V_χ–isotypic component of the sheaf $h_*\mathbb{Q}_\ell$ with respect to its canonical S_n–action. Over the locus of squarefree polynomials $(h_*\mathbb{Q}_\ell)_\chi$ restricts to the local system $\mathbb{V}_{\chi,\ell}$ of Section 2.1.3.

2.2.2 This approach to computing factorization statistics was carried out by Gadish [12], resulting in the identity

$$\sum_{f\in \mathrm{Pol}_n(\mathbb{F}_q)} \frac{1}{|\mathrm{Aut}(f)|} \sum_{\sigma\in \mathrm{Aut}(f)} \chi(F_f\circ\sigma) = q^n \cdot \langle \chi, \mathbf{1}\rangle_{\mathrm{S}_n}. \tag{4}$$

Let us explain the notation $\sum_{\sigma\in\mathrm{Aut}(f)} \chi(F_f\circ\sigma)$. Suppose $f\in \mathrm{Pol}_n(\mathbb{F}_q)$ factors as $f=\prod f_i^{a_i}$, where all the f_i are distinct and irreducible, and let us number the roots of f as $z_1,\dots,z_n$. Then we define $\mathrm{Aut}(f) := \prod_i \mathrm{S}_{a_i}$, and the ordering of the roots identifies $\mathrm{Aut}(f)$ with a Young subgroup of S_n. We let $F_f\in \mathrm{S}_n$ denote the permutation representing the action of Frobenius on the roots, ie such that $\mathrm{Frob}_q(z_i) = z_{F_f(i)}$. The permutation F_f is not uniquely determined unless f is squarefree. Nevertheless the sum

$$\sum_{\sigma\in\mathrm{Aut}(f)} \chi(F_f\circ\sigma)$$

is well-defined and independent of our choices. Note that if f is squarefree then this sum is just $\chi(f)$.

2.2.3 Note also that the right-hand side of (4) can trivially be rewritten as

$$\langle \chi, \mathbf{1}\rangle_{\mathrm{S}_n} = \sum_{i\geq 0} (-1)^i \langle \chi, \mathrm{ch}\, H^i_{\text{ét}}((\mathbb{A}^1)^n, \mathbb{Q}_\ell)\rangle_{\mathrm{S}_n},$$

since the cohomology of $(\mathbb{A}^1)^n$ is just a single copy of the trivial representation of S_n in degree 0. Thus, as expected, on the right-hand side of (4) we see the cohomology of $(\mathbb{A}^1)^n$, together with its symmetric group action.

2.2.4 Now, clearly the application of Grothendieck–Lefschetz to the map $(\mathbb{A}^1)^n \to \mathbb{A}^n$ did not result in the formula for the factorization statistics that we wanted. Instead, it gave us a different formula, for *weighted* factorization statistics. Here the quantity associated to a polynomial f involves an averaging procedure, weighted by $|\mathrm{Aut}(f)|^{-1}$, where $\mathrm{Aut}(f)$ is the stabilizer group of the action of S_n on the fiber above f. So the key problem in recovering Hyde's formula from the action of S_n on $(\mathbb{A}^1)^n$ seems to be that S_n does not act freely.

2.2.5 Indeed, the presence of the weight $|\mathrm{Aut}(f)|^{-1}$ in Gadish's formula suggests that it may alternatively be obtained by using the *stack quotient* $[(\mathbb{A}^1)^n/\mathrm{S}_n]$. The Grothendieck–Lefschetz trace formula for a smooth algebraic stack $\mathcal{X}$ takes the form [2, Corollary 6.4.10]

$$\sum_{x\in\mathcal{X}(\mathbb{F}_q)/\simeq} \frac{1}{|\mathrm{Aut}(x)|}\,\mathrm{Tr}(\Phi_q \,|\, \mathcal{F}_{\bar{x}}) = q^{\dim \mathcal{X}} \sum_i (-1)^i \,\mathrm{Tr}(\Phi_q \,|\, H^i_{\text{ét}}(\mathcal{X}_{\overline{\mathbb{F}}_q}, \mathcal{F})), \tag{5}$$

where $\mathrm{Aut}(x)$ denotes the $\mathbb{F}_q$–points of the isotropy group of $\mathcal{X}$ at x. This formula, applied to the cohomology of $[(\mathbb{A}^1)^n/\mathrm{S}_n]$, with coefficients in the local system $\mathbb{V}_\chi$ induced by the étale S_n–cover $(\mathbb{A}^1)^n \to [(\mathbb{A}^1)^n/\mathrm{S}_n]$, reproduces Gadish's formula (4). In fact, Gadish proved a general formula applicable to a ramified G–cover $h\colon X \to X/G$, where G is any finite group, and it can be realized as the Grothendieck–Lefschetz trace formula applied to $[X/G]$. The reason why the two approaches produce the same formula is that the pushforward of $\mathbb{V}_\chi$ along $[X/G] \to X/G$ is precisely the sheaf $(h_*\mathbb{Q}_l)_\chi$ that Gadish uses to obtain his formula (Section 2.2.1).

2.2.6 All of the above suggests that the nonfree action of S_n on $(\mathbb{A}^1)^n$ is the culprit that causes weighted factorization statistics to appear, rather than ordinary factorization statistics, when applying Grothendieck–Lefschetz to $(\mathbb{A}^1)^n \to \mathbb{A}^n$. In other words, the isotropy of the stack $[(\mathbb{A}^1)^n/\mathrm{S}_n]$ is the source of the weights $|\mathrm{Aut}(f)|^{-1}$, and to recover Hyde's formula geometrically we should somehow try to remove the stackiness of $[(\mathbb{A}^1)^n/\mathrm{S}_n]$. We could try to do this by replacing $(\mathbb{A}^1)^n$ by a space X_n that looks somehow "similar" to affine space but which has a free action of S_n. In particular, there should be a bijection of sets

$$[X_n/\mathrm{S}_n](\mathbb{F}_q) \xrightarrow{\sim} \mathrm{Pol}_n(\mathbb{F}_q).$$

This is in contrast to $[(\mathbb{A}^1)^n/\mathrm{S}_n](\mathbb{F}_q)$, which is a groupoid with $\pi_0[(\mathbb{A}^1)^n/\mathrm{S}_n](\mathbb{F}_q) = \mathrm{Pol}_n(\mathbb{F}_q)$ but with nontrivial stabilizers. Moreover, we saw that the right-hand side of Gadish's formula corresponds to the cohomology of $(\mathbb{A}^1)^n$ (Section 2.2.3); if we

want the Grothendieck–Lefschetz trace formula to replicate the right-hand side of Hyde's formula (1), then the space X_n should have the same cohomology groups as $\mathrm{PConf}_n(\mathbb{R}^3)$, and the cohomology groups $H^{2i}_{\text{ét}}$ should be pure Tate of weight $2i$.

2.3 Bug-eyed configuration spaces

2.3.1 The simplest way to modify $(\mathbb{A}^1)^n$ into a space with a free S_n–action is to add extra data at the points stabilized by S_n. Accordingly, we propose the following construction. Let us define a nonseparated scheme X_n as the space parametrizing n not-necessarily-distinct ordered points on the line $\mathbb{A}^1$, together with an auxiliary total ordering on each of the sets of points which happen to coincide. For example, X_2 is the plane with doubled diagonal. In general, X_n is obtained by gluing together $n!$ copies of affine n–space $\mathbb{A}^n$ along open subspaces given by complements of diagonal loci. (A more general construction will be carefully described in Section 3.2.3.) There is a natural action of S_n on X_n, which is now a *free* action: the loci in $\mathbb{A}^n$ which have been "duplicated" are precisely those with nontrivial S_n–stabilizers. Thus the quotient stack $[X_n/\mathrm{S}_n]$ has trivial isotropy groups and is now actually an algebraic space. The algebraic space $[X_n/\mathrm{S}_n]$ is smooth and looks deceptively similar to affine n–space $\mathbb{A}^n$: there is a natural map $p\colon [X_n/\mathrm{S}_n] \to \mathbb{A}^n$ such that for the natural stratification of $[X_n/\mathrm{S}_n]$ with strata indexed by the partition lattice, the restriction of p to each stratum is an isomorphism onto its image. Nevertheless p is not an isomorphism — for one, $[X_n/\mathrm{S}_n]$ is not separated (in fact not even locally separated), since X_n is not separated. This is a geometric phenomenon which is impossible for schemes but possible for algebraic spaces: informally speaking, schemes can only be nonseparated because they have "too many points", whereas algebraic spaces can be nonseparated also because they have "too many tangent directions"; cf [20, Example 1, page 9]. For example, $[X_2/\mathrm{S}_2]$ is the plane with doubled tangent directions along the diagonal. In general $[X_n/\mathrm{S}_n]$ is a "bug-eyed cover" of $\mathbb{A}^n$, in the terminology of Kollár [21].

2.3.2 Our first claim is that if we associate a locally constant ℓ–adic sheaf $\mathbb{V}_{\chi,\ell}$ on $[X_n/\mathrm{S}_n]$ as before, then the left-hand side of the Grothendieck–Lefschetz trace formula applied to $\mathcal{X} = [X_n/\mathrm{S}_n]$ and $\mathcal{F} = \mathbb{V}_{\chi,\ell}$ becomes precisely the left-hand side of (1). The argument is almost exactly the same as in the squarefree case: we observe that since $[X_n/\mathrm{S}_n] \to \mathbb{A}^n$ is a bug-eyed cover, it follows that $[X_n/\mathrm{S}_n](\mathbb{F}_q) \to \mathrm{Pol}_n(\mathbb{F}_q)$ is a bijection. Moreover, the monodromy of $X_n \to [X_n/\mathrm{S}_n]$ over a polynomial f is given by S_n acting on the set of roots of f. This identifies Frobenius acting on

the set of roots with a permutation whose cycle type is given by the degrees of the irreducible factors of f. This explains why factorization statistics are connected to the cohomology of X_n. But how is X_n related to $\mathrm{PConf}_n(\mathbb{R}^3)$?

2.3.3 Our second claim (Theorem 3.3.5) is that there exists an S_n–equivariant map

$$f\colon \mathrm{PConf}_n(\mathbb{R}^3) \to X_n(\mathbb{C})$$

which is a fiber bundle with fiber $\mathbb{R}^n$ and in particular a weak homotopy equivalence. Let us define this map f. A point in $\mathbb{R}^3$ has an x–, a y–, and a z–coordinate; by forgetting the z–coordinate we obtain a point in $\mathbb{R}^2 \cong \mathbb{C}$. In this way we obtain from a configuration of n distinct ordered points in $\mathbb{R}^3$ a configuration of n points in $\mathbb{C}$, where the points may now coincide with each other. But note that if some subset of these points in $\mathbb{C}$ are equal, then their preimages in $\mathbb{R}^3$ must have distinct z–coordinates. It follows that the z–coordinate defines a total ordering on this set of coinciding points. In this way we obtain from each point of $\mathrm{PConf}_n(\mathbb{R}^3)$ a point of $X_n(\mathbb{C})$. We remark that this construction is quite similar to the stratification of $\mathrm{Conf}_n(\mathbb{C})$ considered by Fox–Neuwirth [9] and Fuks [10, Section 3].

2.3.4 The two claims combined together give a "philosophical" explanation of Hyde's theorem: the sum $\sum_{f\in \mathrm{Pol}_n(\mathbb{F}_q)} \chi(f)$ can be evaluated in terms of the S_n–equivariant étale cohomology of X_n, and over the complex numbers the space X_n has the same cohomology as $\mathrm{PConf}_n(\mathbb{R}^3)$. If we want to actually re-prove Hyde's theorem by purely geometric considerations we need two further ingredients:

- By the Artin comparison theorem, and general constructibility and base change results, the Betti cohomology of $X_n(\mathbb{C})$ is isomorphic to the étale cohomology of $X_{n,\overline{\mathbb{F}}_q}$ away from finitely many characteristics. To recover Hyde's theorem in all characteristics we need to know that the sheaves $R^q\pi_*\mathbb{Q}_\ell$ are locally constant over $\mathrm{Spec}(\mathbb{Z}[1/\ell])$, where $\pi\colon X_n \to \mathrm{Spec}(\mathbb{Z})$ is the structural morphism.
- To apply the Grothendieck–Lefschetz formula, we need to know the Galois action on $H^i(X_{n,\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell)$; specifically, what is needed for Hyde's theorem is the statement that H^i is pure Tate of weight i.

Just as in the squarefree case, we will resolve these issues by finding a direct geometric method to compute the cohomology of X_n that works independently of the characteristic and which also keeps track of the Galois action.

2.4 A cellular decomposition

2.4.1 There are many examples of algebraic varieties Z with the property that $H^i(Z)$ is pure Tate of weight i (and then nonzero only for even i), such as smooth Schubert varieties and smooth projective toric varieties. The most hands-on way in which this can happen is when Z is smooth and admits an *algebraic cell decomposition*, a certain kind of decomposition into locally closed subschemes each of which is isomorphic to an affine space $\mathbb{A}^d$ (Definition 3.5.1). If this holds then H^{2i} is pure Tate of weight $2i$ for all i, and if Z is d–dimensional then $\dim H^{2i}(Z)$ is given by the number of cells of dimension $d-i$ (Proposition 3.6.7).

2.4.2 We now claim that the schemes X_n admit such an algebraic cell decomposition. For example, X_2 is the plane with doubled diagonal, which can be decomposed as the union of a plane and a line:

$$X_2 = \mathbb{A}^2 \cup \mathbb{A}^1,$$

where $\mathbb{A}^2$ is the locus $\{x_1 \neq x_2\} \cup \{x_1 = x_2, 1 \leq 2\}$ and $\mathbb{A}^1$ is the line $\{x_1 = x_2, 2 \leq 1\}$. Similarly X_3 looks like $\mathbb{A}^3$ except the planes defined by $x = y$, $x = z$ and $y = z$ have all been doubled, and there are six copies of the line $x = y = z$; this gives rise to a decomposition

$$X_3 = \mathbb{A}^3 \cup \mathbb{A}^2 \cup \mathbb{A}^2 \cup \mathbb{A}^2 \cup \mathbb{A}^1 \cup \mathbb{A}^1.$$

What will happen in general is indeed that each X_n admits an algebraic cell decomposition (which is far from unique or canonical, as can be seen even for X_2). Although this statement might seem simple, there is in fact a nontrivial combinatorial theorem lurking in the background here, which is a theorem of Delucchi [7] and Lofano and Paolini [25] stating that the set of chambers of an arrangement of real hyperplanes always admits a "*valid order*". Each choice of valid order gives rise to an algebraic cell decomposition of X_n, so it follows that the Betti numbers of X_n in étale cohomology are independent of the characteristic of the base field, and the i^{th} étale cohomology group of X_n is pure Tate of weight i in any characteristic. Moreover, *even though the cell decomposition is not* S_n*–equivariant*, it follows that the decomposition of the étale cohomology into irreducible S_n–representations is the same in any characteristic. Indeed, the ℓ–adic sheaf on $\operatorname{Spec}(\mathbb{Z}[1/\ell])$ given by the q^{th} étale cohomology of X_n with $\mathbb{Q}_\ell$–coefficients is the direct sum of its isotypical components under the action of S_n, by Maschke's theorem. But this sheaf is locally constant, and a direct summand of a locally constant sheaf is locally constant, so each isotypical component must itself be locally constant over $\operatorname{Spec}(\mathbb{Z}[1/\ell])$. Hence the étale cohomology of $[X_n/\mathrm{S}_n]$ with coefficients in V_χ, which is identified with this isotypic component, is locally constant as well.

2.5 Conclusion

2.5.1 We are now in a position to assemble what has been said so far into a proof. More precisely, putting these three ingredients together gives a completely geometric proof of Hyde's theorem:

- The identification between the Grothendieck–Lefschetz trace formula for $[X_n/\mathrm{S}_n]$ and the left-hand side of Hyde's formula (Section 2.3.2).
- The weak equivalence between $\mathrm{PConf}_n(\mathbb{R}^3)$ and $X_n(\mathbb{C})$ (Theorem 3.3.5).
- The space X_n having an algebraic cell decomposition over the integers, so that its cohomology is pure of Tate type, with Betti numbers independent of characteristic of the base field (Corollary 3.5.8 and Proposition 3.6.7).

2.5.2 Remark Amusingly, while $|\mathrm{PConf}_n(\mathbb{A}^1)(\mathbb{F}_q)| = q\cdot(q-1)\cdots(q-(n-1))$ has its point counts given by a falling Pochhammer symbol $(q)_{\underline{n}}$, the point counts of $X_n(\mathbb{F}_q)$ are given by a *rising* Pochhammer symbol $(q)^{\underline{n}} = q(q+1)\cdots(q+n-1)$. We can see this directly by considering the map $X_n(\mathbb{F}_q)\to X_{n-1}(\mathbb{F}_q)$ which forgets the last point, and has fiber of size $q+(n-1)$. This may be compared with the map of configuration spaces $\mathrm{PConf}_n(\mathbb{A}^1)\to\mathrm{PConf}_{n-1}(\mathbb{A}^1)$, which can be used to derive in a similar fashion the number of $\mathbb{F}_q$–points of $\mathrm{PConf}_n(\mathbb{A}^1)$. Furthermore, just as we may consider the configuration space $\mathrm{PConf}_n(Y)$ for any space Y, we may construct a variant of X_n where $\mathbb{A}^1$ is replaced by an arbitrary Y. If we take Y to be the finite set $[k]=\{1,\dots,k\}$, then we obtain a set with $(k)^{\underline{n}}$ elements, on which S_n acts freely, with quotient exactly the set of multisubsets of $[k]$. In some sense, this whole paper is an elaboration on this bijective proof that $[k]$ has $(k)^{\underline{n}}/n!$ multisubsets.

2.5.3 Remark As a final remark on the proof let us point out that in our argument we applied Artin's comparison theorem between étale and singular cohomology to X_n, and not to the quotient $[X_n/\mathrm{S}_n]$. One reason is psychological — the scheme X_n is easier to visualize than the bug-eyed algebraic space $[X_n/\mathrm{S}_n]$ — but another reason is that it makes the argument significantly technically simpler. The reason is that a finite type algebraic space over $\mathbb{C}$ admits an analytification if and only if it is locally separated [6, Theorem 2.2.5], so the space $[X_n/\mathrm{S}_n]$ doesn't even *have* an analytification to which Artin's theorem could have been applied. The fact that $[X_n/\mathrm{S}_n]$ does not admit an analytification is perhaps not surprising, if one accepts that no sensible map of analytic spaces could possibly correspond to the bug-eyed cover $[X_n/\mathrm{S}_n]\to\mathbb{A}^n$. Thus, if we

wanted to apply Artin's theorem to $[X_n/\mathrm{S}_n]$ we would have to form the analytification in a larger category of groupoids in complex-analytic spaces, to which a version of Artin's theorem could have been applied only after a cohomological descent argument (which to our knowledge is not written out in the literature).

2.6 General Weyl groups

2.6.1 An advantage of our geometric approach to Hyde's theorem is that it generalizes to a general Weyl group, as we now discuss. Suppose that we are given an n–dimensional root system, and that W is the associated reflection group acting on $\mathbb{A}^n$. We will associate to this data a nonseparated scheme X_W obtained by gluing together copies of $\mathbb{A}^n$ along complements of reflecting hyperplanes, in such a way that when W is the symmetric group S_n acting on $\mathbb{A}^n$ by permuting coordinates, then we recover the scheme X_n. The correct definition of X_W is clear once one realizes that a total order on the set $\{1,\dots,n\}$ is precisely the same thing as a Weyl chamber for the symmetric group S_n. More generally, a partition of the set $\{1,\dots,n\}$ corresponds to a parabolic subgroup of S_n, and a total order on each block of the partition is a Weyl chamber for that parabolic subgroup. (We recall that a *parabolic subgroup* of a reflection group W is a subgroup which is the stabilizer of some point of $\mathbb{A}^n$.) Thus we define in general X_W to be the nonseparated scheme parametrizing a point $y \in \mathbb{A}^n$ together with a Weyl chamber for the parabolic subgroup $\mathrm{Stab}(y)$. The stabilizer $\mathrm{Stab}(y)$ is itself a reflection group; it is the subgroup generated by reflections in hyperplanes containing y, so we can think of the data of a Weyl chamber for $\mathrm{Stab}(y)$ as a chamber of the subarrangement consisting of the reflecting hyperplanes through x. Just as X_n is glued from $n!$ copies of $\mathbb{A}^n$, we may glue together X_W from $|W|$ copies of $\mathbb{A}^n$, one for each Weyl chamber of W.

2.6.2 Let us now discuss how to generalize the statement of Hyde's theorem. Indeed, in formulating Hyde's theorem we used that to any element of $\mathrm{Pol}_n(\mathbb{F}_q) = (\mathbb{A}^n/\mathrm{S}_n)(\mathbb{F}_q)$ we can associate a well-defined conjugacy class in S_n. In the squarefree case we could take this conjugacy class to be the action of Frobenius on the set of roots of the polynomial, ie on the fiber of $\mathbb{A}^n \to \mathbb{A}^n/\mathrm{S}_n$, but for a polynomial with repeated roots this does not prescribe a unique conjugacy class and our recipe was instead that we took the conjugacy class specified by the partition given by the degrees of the irreducible factors. How should this procedure be generalized?

2.6.3 First, we take a step back and consider a finite group G acting *freely* on a scheme X over $\mathbb{F}_q$. Then any $\mathbb{F}_q$–point x of X/G determines a conjugacy class of elements of G. Indeed, we lift $x \in (X/G)(\mathbb{F}_q)$ to a point $y \in X(\overline{\mathbb{F}}_q)$, and there is

a unique element $g \in G$ such that $g \cdot y = \mathrm{Frob}(y)$. If we had chosen another lift y then the element g would have been changed by conjugation, so the conjugacy class of g is independent of choices made. Alternatively, the conjugacy class is the image of Frobenius under the composite

$$\widehat{\mathbb{Z}} \cong \pi_1^{\text{ét}}(\mathrm{Spec}(\mathbb{F}_q)) \xrightarrow{x} \pi_1^{\text{ét}}(X/G) \to G,$$

where we omit basepoints and the second homomorphism classifies the G–torsor $X \to X/G$. In any case, if G does not act freely on X then we can still lift $x \in (X/G)(\mathbb{F}_q)$ to $y \in X(\overline{\mathbb{F}}_q)$, but now an element $g \in G$ carrying y to $\mathrm{Frob}(y)$ is only well defined as a left coset modulo the subgroup $\mathrm{Stab}(y)$; we do not get a well-defined conjugacy class in G but rather a conjugacy class of cosets, and in general this is the best we can say.

2.6.4 Now suppose however that we are given a Weyl group W acting on $\mathbb{A}^n$ as above. As already mentioned, a parabolic subgroup $\mathrm{Stab}(y)$ is itself a reflection group; it is the subgroup generated by reflections through hyperplanes containing y. In particular, $\mathrm{Stab}(y)$ acts freely and transitively on the set of chambers of the arrangement of reflecting hyperplanes through y. An immediate consequence is that if we fix the choice of a Weyl chamber C_0 for W, then any coset of a parabolic subgroup has a distinguished representative: every left coset of $\mathrm{Stab}(y)$ contains a unique element w with the property that $w \cdot C_0$ and C_0 are on the same side of every reflecting hyperplane passing through y. This element w is called the *minimal representative* of the coset. It can also be characterized as the unique element of the coset of smallest Coxeter length with respect to C_0. See eg [17, Sections I.5-1 and I.5-2]. It follows that in the Weyl group case we *can*, in a canonical manner, associate to an $\mathbb{F}_q$–point x of $\mathbb{A}^n/W$ a conjugacy class of elements of W. Indeed, we may lift x to a point $y \in \mathbb{A}^n(\overline{\mathbb{F}}_q)$, there is a unique left coset modulo $\mathrm{Stab}(y)$ carrying y to $\mathrm{Frob}(y)$, and this coset has a unique minimal representative w once we choose a base chamber C_0. Choosing a different base chamber or a different lift would have the effect of changing w by conjugation, so we get a well-defined conjugacy class in W associated to x. If $\chi\colon W \to \mathbb{C}$ is a class function, then we denote by $\chi(x)$ the value of χ on the conjugacy class associated to $x \in (\mathbb{A}^n/W)(\mathbb{F}_q)$.

2.6.5 We may now state the general form of Hyde's result. We state the result for a Weyl group, for the convenience of having everything defined over $\mathbb{Z}$; for a general finite Coxeter group, the scheme X_W would be defined over the ring of integers of a number field, and the statement of the following theorem would become more complicated.

2.6.6 Theorem *Let $W \subset \mathrm{GL}(n,\mathbb{Z})$ be a Weyl group acting by reflections on affine n–space. Let $\chi\colon W \to \mathbb{C}$ be a class function. Let M_W be the complement of the associated codimension-3 subspace arrangement in $(\mathbb{R}^3)^n$. For any finite field $\mathbb{F}_q$, we have*

$$\sum_{x \in (\mathbb{A}^n/W)(\mathbb{F}_q)} \chi(x) = q^n \sum_{i \geq 0} q^{-i} \langle \chi, \mathrm{ch}\, H^{2i}(M_W, \mathbb{Q}) \rangle_W .$$

2.6.7 The proof of the theorem follows, just as the proof of Hyde's theorem, by applying the Grothendieck–Lefschetz trace formula to $[X_W/W]$ and combining the following ingredients:

- The W–equivariant weak homotopy equivalence $M_W \simeq X_W(\mathbb{C})$ (Theorem 3.3.5).
- The space X_W admits an algebraic cell decomposition over the integers, so that its cohomology is pure of Tate type, with Betti numbers independent of characteristic of the base field (Corollary 3.5.8 and Proposition 3.6.7).
- W acts freely on X_W, and the algebraic space $[X_W/W]$ is a bug-eyed cover of $\mathbb{A}^n/W \cong \mathbb{A}^n$, so that there is a bijection of $\mathbb{F}_q$–points $[X_W/W](\mathbb{F}_q) \xrightarrow{\sim} (\mathbb{A}^n/W)(\mathbb{F}_q)$.
- Under this bijection, the W–conjugacy class associated to an element of the space $[X_W/W](\mathbb{F}_q)$ (as in Section 2.6.3, using that W acts freely on X_W) coincides with the conjugacy class associated to the corresponding element in $(\mathbb{A}^n/W)(\mathbb{F}_q)$ (as in Section 2.6.4).

The first three ingredients work out in exactly the same way as for Hyde's theorem, so let us only comment on the last one. Pick a point $x \in (\mathbb{A}^n/W)(\mathbb{F}_q) = [X_W/W](\mathbb{F}_q)$. An $\overline{\mathbb{F}}_q$–point of X_W over x is given by a pair (y, C) where $y \in \mathbb{A}^n(\overline{\mathbb{F}}_q)$ is in the fiber of $\mathbb{A}^n \to \mathbb{A}^n/W$ over x, and C is a Weyl chamber of $\mathrm{Stab}(y)$ acting on $\mathbb{A}^n$. The Galois group acts as

$$\mathrm{Frob}((y, C)) = (\mathrm{Frob}(y), C),$$

and since W acts freely on X_W, there is a unique $w \in W$ such that $w \cdot (y, C) = (\mathrm{Frob}(y), C)$. The claim is that if we choose a base chamber C_0 for the Coxeter group W contained inside C, then w is the minimal representative inside the coset consisting of elements which move y to $\mathrm{Frob}(y)$. Indeed, $w \cdot C_0$ and C_0 being on the same side of every reflecting hyperplane for the parabolic subgroup $\mathrm{Stab}(y)$ implies that w preserves the Weyl chamber of this parabolic subgroup containing C_0, ie C. This completes the proof of the generalized version of Hyde's theorem.

2.6.8 Remark A cute application of our construction for a general reflection group is as follows. Consider a finite reflection group W acting on affine space. Consider on one hand our space X_W, which parametrizes pairs (y, C) of a point in affine space and C a chamber of $\mathrm{Stab}(y)$, and on the other hand the "inertia variety" I_W parametrizing pairs (y, w) with y a point in affine space and w an element of $\mathrm{Stab}(y)$. Since $\mathrm{Stab}(y)$ acts freely and transitively on its set of chambers, we obtain a set-theoretic bijection between X_W and I_W after choosing a base chamber C_0 of W: indeed, such a choice fixes also the choice of a base chamber for each parabolic $\mathrm{Stab}(y)$, and hence a bijection between $\mathrm{Stab}(y)$ and its set of chambers for every point y. Now

$$I_W \cong \coprod_{w \in W} \mathrm{Fix}(w),$$

where $\mathrm{Fix}(w) \subseteq \mathbb{A}^n$ denotes the linear subspace fixed by w, and each subset $\mathrm{Fix}(w)$ is carried isomorphically to a locally closed subscheme of X_W under this bijection. In particular we may partition both X_W and I_W into locally closed subschemes which are pairwise isomorphic; both spaces X_W and I_W have the same class in the Grothendieck ring of varieties, and we obtain a calculation of $[X_W] \in K_0(\mathrm{Var})$. Thus for any algebraic cell decomposition of X_W, the number of d–dimensional cells in the decomposition is just the number of $w \in W$ whose set of fixed points has dimension d. Since the Betti numbers of the complement of the arrangement of reflecting hyperplanes can be read off from a cell decomposition, we obtain the formula

$$\sum_{w \in W} q^{\operatorname{codim} \mathrm{Fix}(w)} = \sum_i \dim H^{2i}(M_W, \mathbb{Q}) \cdot q^i.$$

The result is of course not new: both sides are equal to $\prod_i (1 + m_i q)$, where m_i are the exponents of the group, by classical results of Shephard–Todd [33, Theorem 5.3] and Brieskorn [4, Théorème 6(ii)], respectively. Strictly speaking, Brieskorn's results concern the cohomology of the complement of the *complexified* arrangement, so we also need to know that the complements of the arrangements in $\mathbb{C}^n$ and $(\mathbb{R}^3)^n$ have the same Betti numbers up to a degree shift, which is, for example, a consequence of the Goresky–MacPherson formula [13].

2.6.9 Remark The construction of the previous paragraph produces a decomposition of X_W into pairwise disjoint locally closed subschemes isomorphic to affine spaces, indexed by elements $w \in W$, after the choice of a base chamber. At an earlier stage of this project we expected this decomposition to be an algebraic cell decomposition in the sense of Definition 3.5.1. This is not true, however: for the scheme X_4, the "cells"

corresponding to the permutations (14) and (13)(24) both have the property that they intersect each other's closure nontrivially. According to Corollary 3.5.8 it is still true that X_W admits an algebraic cell decomposition, but the construction is more nontrivial.

2.6.10 Remark As explained by Hyde [14, Section 2.5], one can evaluate the $q = 1$ specialization of Hyde's theorem to produce an expression for the direct sum of all cohomology groups of $\mathrm{PConf}_n(\mathbb{R}^3)$ as a representation of S_n, and the result is that

$$\bigoplus_k H^k(\mathrm{PConf}_n(\mathbb{R}^3), \mathbb{Q}) \cong \mathbb{Q}[\mathrm{S}_n],$$

ie one obtains the regular representation. (As Hyde points out, the result is not new.) It is natural to ask whether our geometric approach can shed light on this observation, in particular given that setting $q = 1$ amounts to counting the number of cells in an algebraic cell decomposition. Unfortunately we do not see an easy way to see this result (in particular since the cell decomposition is not equivariant), but we remark that the analogous statement remains true for any finite reflection group; see for example [27, Theorem 1.4(b)]. Let us give a short proof. Suppose W acts by reflections on $\mathbb{R}^n$, and for any $d \geq 1$ let $M^{(d)}$ denote the associated codimension-d subspace arrangement in $(\mathbb{R}^d)^n$. We claim that the W–equivariant Euler characteristic of $M^{(d)}$ depends only on the parity of d. Indeed, writing $\mathbb{R}^d = \mathbb{C} \times \mathbb{R}^{d-2}$ we get an induced $\mathbb{C}^*$–action on $M^{(d)}$ with fixed-point set $M^{(d-2)}$, and by localization the two Euler characteristics coincide. More explicitly, one may apply Mayer–Vietoris to the open cover of $M^{(d)}$ given by the complement of $M^{(d-2)}$ (whose Euler characteristic vanishes because $\mathbb{C}^*$ acts freely) and the $W \times \mathbb{C}^*$–equivariant tubular neighborhood of $M^{(d-2)}$ given by $\mathbb{C}^n \times M^{(d-2)}$. When $d = 1$ we have that $M^{(1)}$ is a disjoint union of contractible components which are freely and transitively permuted by W, so in this case the equivariant Euler characteristic is the regular representation of W. Then the same is true for $M^{(3)} = M_W$, and since M_W has no odd cohomology, the claim follows.

2.7 Type B and even polynomials

2.7.1 Let us examine our formula more closely in the type B case. The type B Weyl group is the hyperoctahedral group $B_n := \mathrm{S}_n \ltimes \mathbb{Z}/2^n$, which we can think of as the automorphism group of the $(\mathbb{Z}/2)$–set $[n]^{\pm} = \{1, -1, \ldots, n, -n\}$. The group B_n acts on $\mathbb{A}^n = \mathrm{Spec}\,\mathbb{Z}[z_1, \ldots, z_n]$, and its ring of invariants is well known to be generated by

$$w_d(z) := \sum_{1 \leq i_1 < \cdots < i_d \leq n} z_{i_1}^2 \cdots z_{i_d}^2$$

for $d = 1, \dots, n$. These invariants give coordinates on the quotient space $\mathbb{A}^n/B_n = \operatorname{Spec}(\mathbb{Z}[w_1, \dots, w_n])$. We can interpret the coordinate w_d as the coefficient of t^{2d} in the product

$$\prod_{i=1}^{n}(t^2 - z_i^2) = \prod_{i=1}^{n}(t - z_i)(t + z_i).$$

Over an algebraically closed field, a polynomial $f(t)$ admits a factorization of this form if and only if it is *even*. Thus we see that $\mathbb{A}^n/B_n$ is a parameter space for monic degree-$2n$ even polynomials, where w_d records the $2d^{\text{th}}$ coefficient, and $\mathbb{A}^n \to \mathbb{A}^n/B_n$ is the finite map that takes $(z_1, \dots, z_n)$ to the even monic degree-$2n$ polynomial $\prod_i(t^2 - z_i^2)$. In these terms, our space X_{B_n} parametrizes even monic degree-$2n$ polynomials, together with a choice of labeling of the set of roots by the set $[n]^{\pm}$, subject to the condition that $\alpha_{-j} = -\alpha_j$ for all $j \in [n]^{\pm}$.

2.7.2 Let $f \in \mathrm{Pol}_{2n}^{\mathrm{even}}(\mathbb{F}_q)$ be an even monic degree-$2n$ polynomial. Its factorization into irreducibles takes the form

$$f(t) = \prod_{g \text{ even}} g(t) \prod_{h \text{ not even}} h(t)h(-t),$$

where we have used that the factorization into irreducibles must be invariant under the transformation $t \mapsto -t$. We associate to f a double integer partition of n, which we call its *factorization type*:

$$(\lambda, \mu) = (1^{u_1}2^{u_2} \cdots, \overline{1}^{v_1}\overline{2}^{v_2} \cdots), \quad d = \sum_i u_i i + \sum_j v_j j.$$

Here u_i is the number of degree-$2i$ irreducible even factors of f, and v_j is the number of pairs $\{h(t), h(-t)\}$ of degree-j irreducible noneven factors of f. Similarly, a signed permutation $\sigma \in B_n$ gives rise to a $(\mathbb{Z}/2)$–equivariant partition of the set $[n]^{\pm}$ into cycles, and the cycle type of σ may be encoded by a double partition,

$$(1^{a_1}2^{a_2} \cdots, \overline{1}^{b_1}\overline{2}^{b_1} \cdots),$$

where a_i denotes the number of $(\mathbb{Z}/2)$–invariant blocks of size $2i$ in the partition, and b_j denotes the number of pairs of blocks of size j that are interchanged by the $(\mathbb{Z}/2)$–action. This construction determines a bijection between conjugacy classes of B_n and double partitions. A polynomial f has factorization type (λ, μ) if and only if Frobenius permutes a signed labeling of its roots in $X_{B_n}(\overline{\mathbb{F}}_q)$ by an element $\sigma \in B_n$ in the conjugacy class (λ, μ). Thus we have proved the following.

2.7.3 Theorem *Let χ be a character of B_n, considered as a function on the set of double integer partitions of n. We write $\chi(f)$ for the value of χ on the factorization type of f. Then*

$$\sum_{f \in \mathrm{Pol}_{2d}^{\mathrm{even}}(\mathbb{F}_q)} \chi(f) = q^n \sum_{i \geq 0} q^{-i} \langle \chi, \mathrm{ch}\, H^{2i}(\mathrm{PConf}_n^{\mathbb{Z}/2}(\mathbb{R}^3), \mathbb{Q}) \rangle_{B_n}.$$

Here $\mathrm{PConf}_n^{\mathbb{Z}/2}(\mathbb{R}^3)$ denotes the orbit configuration space, which parametrizes $(\mathbb{Z}/2)$–equivariant injections of $[n]^{\pm}$ into $\mathbb{R}^3$, considered with its $(\mathbb{Z}/2)$–action $x \mapsto -x$.

2.7.4 Thus we see that in the type B case, our generalized version of Hyde's theorem gives a formula for the factorization statistics of even polynomials. These statistics record the number of even and noneven factors in the factorization type of p. As mentioned previously, Jiménez Rolland and Wilson [16] studied a type B analogue of squarefree factorization statistics. Their formulation of factorization statistics is different, but equivalent to the one used here. Jiménez Rolland and Wilson consider monic squarefree polynomials $q(t)$ of degree d with nonzero constant term, and their statistics concern the factorization of $q(t)$ into irreducibles and whether the roots of q are quadratic residues. The statistics we consider are related to theirs by the transformation $f(t) = q(t^2)$. Indeed, let $i(t)$ be a degree-d irreducible polynomial. Then $i(t^2)$ either remains irreducible or factors as $h(t)h(-t)$. Further, $i(t^2)$ factors if and only if all of the roots of $i(t)$ in $\mathbb{F}_{q^{\deg i}}$ are quadratic residues, and $i(t^2)$ remains irreducible if and only if all of the roots of $i(t)$ are nonresidues.

3 Non-Hausdorff models for complements of arrangements

3.1 Conventions and terminology on hyperplane arrangements

3.1.1 Let A be a finite set of affine hyperplanes in $\mathbb{R}^n$. We let $\mathcal{L}(A)$ denote the partially ordered set consisting of all nonempty intersections of hyperplanes in A, including $\mathbb{R}^n$ itself as the empty intersection, ordered by reverse inclusion. Elements of $\mathcal{L}(A)$ are called *flats*. If $K \in \mathcal{L}(A)$ is a flat, then $K \setminus \bigcup_{K<L} L$ is a finite union of convex open regions inside K called *faces*. We denote by $\mathcal{F}(A)$ the set of all faces of the arrangement A. We partially order the set of faces according to

$$F \leq G \iff \overline{F} \supseteq G.$$

If E is a subset of affine n–space (typically a flat), then we denote by $\mathrm{Supp}(E) \subseteq A$ the set of hyperplanes containing E.

3.1.2 The minimal elements of $\mathcal{F}(A)$ are called *chambers*, and we denote the set of chambers by $\mathrm{Ch}(A)$. For $C, D \in \mathrm{Ch}(A)$ we denote by $\mathrm{Sep}(C, D) \subseteq A$ the subset of hyperplanes H which separate C from D, ie such that C and D are contained in different connected components of $V \setminus H$.

3.1.3 We define $\mathrm{Strat}(A)$ to be the partially ordered set consisting of pairs (K, C) where $K \in \mathcal{L}(A)$ and $C \in \mathrm{Ch}(\mathrm{Supp}(K))$. The partial order relation is given by $(K, C) \preceq (L, D)$ if $K \leq L$ and $C \supseteq D$. This poset will parametrize the strata for a natural stratification of the scheme X_A, which we will define shortly.

3.1.4 The poset $\mathrm{Strat}(A)$ is an analogue of the *Salvetti poset* $\mathrm{Sal}(A)$, which is the partially ordered set consisting of pairs (F, C) where $F \in \mathcal{F}(A)$ and $C \in \mathrm{Ch}(\mathrm{Supp}(F))$. The partial order relation is given by $(F, C) \preceq (G, D)$ if $F \leq G$ and $C \supseteq D$. Usually in the literature elements of $\mathrm{Sal}(A)$ are described as pairs (F, C) where $F \in \mathcal{F}(A)$ and $C \in \mathrm{Ch}(A)$ is a chamber adjacent to F (ie whose closure contains F); this is equivalent to the definition given here, since every chamber of $\mathrm{Supp}(F)$ contains a unique chamber of A that is adjacent to F. This poset was introduced by Salvetti [31], who proved that the complement of the complexification of the arrangement A deformation retracts onto a regular CW complex whose poset of cells is precisely the Salvetti poset. The similarity between the Salvetti poset and our stratification of X_A is no coincidence, as we explain in Remark 3.4.3.

3.2 Construction of the scheme X_A

3.2.1 Let A be as above. Following Proudfoot [28], we are going to associate to A a nonseparated scheme X_A over $\mathrm{Spec}(\mathbb{R})$, which will be obtained by gluing together a copy of the affine space $\mathbb{A}^n$ for each chamber of the arrangement. If all hyperplanes in A are defined over a subring $R \subset \mathbb{R}$, then X_A is defined over R, too.

3.2.2 Let us first recall Zariski gluing of schemes in general. Let $X = \bigcup_{i \in I} X_i$ be an open cover of a scheme X. Let $X_{ij} = X_i \cap X_j$. Then the identity map on X induces a collection of isomorphisms $X_{ij} \xrightarrow{\sim} X_{ji}$ for all i and j, and X can be uniquely reconstructed from the collection $\{X_i\}$ and the isomorphisms $X_{ij} \cong X_{ji}$. In the other direction, let $\{X_i\}_{i \in I}$ be an arbitrary family of schemes. Suppose we are given a Zariski open subset $X_{ij} \subset X_i$ for every $i, j \in I$, and an isomorphism $\phi_{ij} : X_{ij} \to X_{ji}$ for every $i, j \in I$. This family of schemes $\{X_i\}$ may be glued together along the isomorphisms ϕ_{ij} to a scheme X, which is then unique up to canonical isomorphism, if and only if the gluing data satisfy the following two axioms:

- $X_{ii} = X_i$, and $\phi_{ii} = \mathrm{id}$ for all $i \in I$.
- ϕ_{ij} restricts to an isomorphism from $X_{ij} \cap X_{ik}$ to $X_{ji} \cap X_{jk}$, and $\phi_{jk} \circ \phi_{ij} = \phi_{ik}$ on $X_{ij} \cap X_{ik}$, for every $i, j, k \in I$ (cocycle condition).

3.2.3 We can now write down gluing data for the scheme X_A. For every $C \in \mathrm{Ch}(A)$ we let $X_C = \mathbb{A}^n$. For any pair $D, C \in \mathrm{Ch}(A)$ we define

$$X_{CD} = \mathbb{A}^n \setminus \left(\bigcup_{H \in \mathrm{Sep}(C,D)} H \right) \subseteq X_C.$$

The identity map on $\mathbb{A}^n$ gives an isomorphism $\phi_{CD} \colon X_{CD} \cong X_{DC}$, and the resulting gluing data is easily checked to satisfy the cocycle condition, since

$$\mathrm{Sep}(C, E) \subset \mathrm{Sep}(C, D) \cup \mathrm{Sep}(D, E).$$

We denote the resulting scheme by X_A.

3.2.4 Remark A reader may wonder why we took care to say that we glued together copies of the scheme $\mathbb{A}^n$ to obtain a scheme over $\mathrm{Spec}(\mathbb{R})$, instead of just saying that we are gluing together copies of $\mathbb{R}^n$ or $\mathbb{C}^n$. The reason is that once we have a *scheme* X_A over a base ring R then we can talk freely about its set of points over an arbitrary R–algebra. In particular, the scheme X_A will be defined over $\mathbb{Z}$ whenever all the hyperplanes have integer coefficients, and in this case we can make sense of $X_A(\mathbb{F}_q)$ for any prime power q, which is at the core of our intended application to factorization statistics. If K is a field and an R–algebra, then a K–valued point of X_A is an element $(b_1, \dots, b_n)$ of $K^n = \mathbb{A}^n(K)$, together with the choice of a chamber of $\mathrm{Supp}(\{(b_1, \dots, b_n)\})$, ie the subarrangement of A consisting of all hyperplanes containing $(b_1, \dots, b_n)$.

3.3 Construction of the fiber bundle

3.3.1 Let A be a finite arrangement of hyperplanes in $V \cong \mathbb{R}^n$. We write $V_{\mathbb{C}}$ for the complexification of V, and $A_{\mathbb{C}} = \{H_{\mathbb{C}} : H \in A\}$ for the induced arrangement of complex hyperplanes in $V_{\mathbb{C}}$. The complement of $A_{\mathbb{C}}$ in $V_{\mathbb{C}}$ will be denoted by $M(A_{\mathbb{C}})$. Similarly we write $V_{(3)} = V_{\mathbb{C}} \times V$, we set $A_{(3)}$ for the codimension-3 subspace arrangement consisting of the subspaces $\{H_{\mathbb{C}} \times H : H \in A\}$, and we let $M(A_{(3)})$ denote the complement of this codimension-3 arrangement.

3.3.2 Example If A consists of the diagonal hyperplanes $x_i = x_j$ in $\mathbb{R}^n$, then $M(A_{\mathbb{C}})$ is the configuration space of n distinct ordered points in $\mathbb{C}$ and $M(A_{(3)})$ is the configuration space of n distinct ordered points in $\mathbb{R}^3$.

3.3.3 If $C \in \mathrm{Ch}(A)$ is a chamber, then define

$$U_C = \big\{(x,z) \in V_{\mathbb{C}} \times V : \text{ if } x \in H_{\mathbb{C}}, \text{ then } z \text{ is in the component of } V \setminus H \text{ containing } C\big\}.$$

Observe that the collection $\{U_C\}_{C \in \mathrm{Ch}(A)}$ is an open cover of $M(A_{(3)})$. Write f for the projection from $V_{\mathbb{C}} \times V$ onto $V_{\mathbb{C}}$. Note that

$$f(U_C \cap U_D) = V_{\mathbb{C}} \setminus \bigcup_{H \in \mathrm{Sep}(C,D)} H_{\mathbb{C}}. \tag{6}$$

3.3.4 Recall from Section 3.2.3 the open cover $X_A = \bigcup_{C \in \mathrm{Ch}(A)} X_C$ of the scheme X_A, where each X_C is an affine space $\mathbb{A}^n$. We will now construct a continuous map $M(A_{(3)}) \to X_A(\mathbb{C})$:

- Cover $M(A_{(3)})$ by the open sets U_C.
- Map each open set U_C onto $V_{\mathbb{C}} \cong X_C(\mathbb{C})$ via f.
- By comparing equation (6) and the description of the intersections $X_C \cap X_D = X_{CD}$ from Section 3.2.3, observe that these glue to a map $M(A_{(3)}) \to X_A(\mathbb{C})$.

3.3.5 Theorem *The map $M(A_{(3)}) \to X_A(\mathbb{C})$ is a fiber bundle with fiber $\mathbb{R}^n$.*

Proof The question is local on the base, so it suffices to show that we get a fiber bundle over each set of the open cover $X_A(\mathbb{C}) = \bigcup_{C \in \mathrm{Ch}(A)} X_C(\mathbb{C})$, ie that $U_C \to V_C \cong \mathbb{C}^n$ is a fiber bundle for every $C \in \mathrm{Ch}(A)$. Now U_C is an open subset of the trivial vector bundle $\mathbb{R}^{3n} \cong V_{\mathbb{C}} \times V \to V_{\mathbb{C}}$. Moreover, the fiber of $U_C \to V_{\mathbb{C}}$ over an element $x \in V_{\mathbb{C}}$ is nothing but the chamber of the arrangement $\mathrm{Supp}(x)$ containing C, considered as an open subspace of V. In particular, the fibers are nonempty and convex, which by the following Lemma 3.3.6 implies that the map is a fiber bundle. □

3.3.6 Lemma *Let $p\colon E \to X$ be a rank-n vector bundle over a metrizable topological space, and $U \subset E$ an open subspace such that $U \cap p^{-1}(x)$ is convex and nonempty for all $x \in X$. Then $U \to X$ is an $\mathbb{R}^n$–fiber bundle, and U is homeomorphic to E as a bundle over X.*

Proof By a partition of unity there is a section contained in U, which we may assume to be the zero section. The strategy will be to construct an exhaustion of U by fiberwise compact subsets

$$\{0\} \Subset B_1 \Subset B_2 \Subset B_3 \Subset \cdots \quad \text{with} \quad \bigcup_{k \geq 1} B_k = U,$$

together with a compatible system of homeomorphisms between B_k and the disk bundle of radius k inside E, for some bundle metric which we fix from now on. Let $S \to X$ be the unit-sphere bundle inside E. Define a map $\phi\colon S \to \mathbb{R}_{>0} \cup \{\infty\}$ by

$$\phi(z) = \sup\{t \in \mathbb{R} : t \cdot z \in U\},$$

so that if we identify E^* (the complement of the zero section in E) with $S \times \mathbb{R}_{>0}$ then

$$U^* = \{(z,t) \in S \times \mathbb{R}_{>0} : t < \phi(z)\}.$$

Since U is open, ϕ is lower semicontinuous, and then there exists a strictly increasing sequence of continuous functions $\psi_1, \psi_2, \ldots : S \to \mathbb{R}_{>0}$ converging pointwise to ϕ; see eg [30, Theorem 2.1.3]. Define

$$B_k^* = \{(z,t) \in S \times \mathbb{R}_{>0} : t \leq \psi_k(z)\}.$$

If we then let B_k be the union of B_k^* with the zero section, then each B_k is homeomorphic to the radius-k disk bundle by rescaling, these homeomorphisms may be chosen compatibly with each other, and by construction the bundles B_k exhaust U. □

3.3.7 Remark It is not hard to see that the construction of Lemma 3.3.6 can, with only a little extra care, be done smoothly, so that if the base X were a smooth manifold then E and U would be diffeomorphic.

3.3.8 Remark The arguments given here generalize to show that similarly $M(A_{\mathbb{C}})$ is an $\mathbb{R}^n$–fiber bundle over the set of *real* points $X_A(\mathbb{R})$ of X_A, which recovers (by a somewhat different argument) a theorem of Proudfoot [28]. Even more generally, the induced c–arrangement in $(\mathbb{R}^c)^n$ is an $\mathbb{R}^n$–bundle over a non-Hausdorff topological space obtained by gluing together copies of the euclidean space $(\mathbb{R}^{c-1})^n$ along complements of subspaces in exactly the same way that we glued together X_A from copies of affine space $\mathbb{A}^n$. The only difference when $c > 3$ is that the base of the fiber bundle can not be interpreted as the real or complex points of an algebraic scheme.

3.4 Stratification of X_A

3.4.1 Let X be a topological space (or a scheme). A *stratification* of X is a decomposition of X into finitely many nonempty pairwise disjoint locally closed subspaces (subschemes) X_α, which are called *strata*, such that the closure of a stratum is a union of strata. We partially order the set of strata by the rule

$$\alpha \preceq \beta \iff \overline{X}_\alpha \supseteq X_\beta.$$

Observe that a stratification is uniquely determined by the set of *closed strata* $\overline{X}_\alpha$, and we will sometimes find it convenient to describe a stratification of a space by specifying the closed strata.

3.4.2 In particular, if A is an arrangement of hyperplanes in $\mathbb{R}^n$, then the collection of flats $\mathcal{L}(A)$ constitute the set of closed strata for a stratification of $\mathbb{A}^n$ into subschemes. When we form the scheme X_A then we glue together copies of $\mathbb{A}^n$ along unions of strata, which means in particular that the stratification of $\mathbb{A}^n$ induces a stratification of the scheme X_A. Explicitly, we can understand X_A and its stratification by starting from the scheme

$$\coprod_{C \in \mathrm{Ch}(A)} \mathbb{A}^n,$$

together with its natural stratification with strata indexed by $\mathcal{L}(A) \times \mathrm{Ch}(A)$. To obtain the scheme X_A we need to impose an appropriate gluing relation on strata: the stratum (K, C) is identified with (L, D) if $K = L$, and the chambers C and D are not separated by any hyperplane containing K. Thus the poset of strata of X_A is the poset $\mathrm{Strat}(A)$ introduced in Section 3.1.3.

3.4.3 Remark Each stratum (K, C) of X_A is itself the complement of an arrangement of hyperplanes inside a vector space of dimension $\dim(K)$. This implies in particular that the set of *real* points of each stratum in X_A is disconnected, with each connected component contractible. More precisely, a connected component of the real points of the stratum (K, C) is just a face in $\mathcal{F}(A)$ which is open in K. Thus $X_A(\mathbb{R})$ has a more refined stratification indexed by pairs (F, C) with $F \in \mathcal{F}(A)$ and $C \in \mathrm{Ch}(\mathrm{Supp}(F))$, so we have recovered the Salvetti poset of Section 3.1.4. Now it is not hard to verify that the closures of strata are contractible and that their intersections are empty or contractible, so that by the nerve lemma it follows that $X_A(\mathbb{R})$ is weakly homotopy equivalent to the order complex of $\mathrm{Sal}(A)$. On the other hand we have the $\mathbb{R}^n$–fiber bundle $M(A_{\mathbb{C}}) \to X_A(\mathbb{R})$ of Remark 3.3.8, so $M(A_{\mathbb{C}})$ is homotopic to the order complex of $\mathrm{Sal}(A)$, as originally proven by Salvetti. A more direct way of understanding this connection is that the stratification of $X_A(\mathbb{R})$ pulls back to a stratification of $M(A_{\mathbb{C}})$, and the stratum inside $M(A_{\mathbb{C}})$ corresponding to a pair $(F, C) \in \mathrm{Sal}(A)$ is precisely the cartesian product

$$F \times C.$$

This coincides with the stratification of $M(A_{\mathbb{C}})$ described by Björner and Ziegler in terms of "sign vectors" [3].

3.5 Algebraic cell decomposition

3.5.1 Definition An *algebraic cell decomposition* of a scheme X is a filtration of X by open subschemes

$$\varnothing = U^{(0)} \subset U^{(1)} \subset \cdots \subset U^{(t)} = X,$$

where for each j there is an isomorphism $U^{(i)} \setminus U^{(i-1)} \cong \mathbb{A}^{d_i}$. We call $U^{(i)} \setminus U^{(i-1)}$ a *cell* of dimension d_i.

3.5.2 We will now show that X_A admits an algebraic cell decomposition. Our strategy will be to make use of the open cover $X_A = \bigcup_{C \in \mathrm{Ch}(A)} X_C$ considered in Section 3.2.3, where $X_C \cong \mathbb{A}^d$ for all $C \in \mathrm{Ch}(A)$. To be more precise, suppose that $C_1, C_2, \ldots, C_t$ is an enumeration of the elements of $\mathrm{Ch}(A)$ in some order. Define for each $0 \leq i \leq t$ the open subset $U^{(i)} \subseteq X_A$ by the formula

$$U^{(i)} := X_{C_1} \cup X_{C_2} \cup \cdots \cup X_{C_i}.$$

Clearly we have a chain of inclusions

$$\varnothing = U^{(0)} \subset U^{(1)} \subset \cdots \subset U^{(t)} = X_A,$$

so all that is needed for this to be an algebraic cell decomposition is that $U^{(i)} \setminus U^{(i-1)}$ is isomorphic to an affine space for every index i. Now we have

$$U^{(i)} \setminus U^{(i-1)} = X_{C_i} \setminus \bigcup_{j<i} (X_{C_i} \cap X_{C_j}),$$

and if we also note that

$$X_{C_i} \cap X_{C_j} = \mathbb{A}^d \setminus \bigcup_{H \in \mathrm{Sep}(C_i, C_j)} H, \tag{7}$$

then we see that an enumeration $C_1, C_2, \ldots, C_t$ of the chambers of the arrangement A will give us an algebraic cell decomposition if and only if the following condition is satisfied:

($\star$) For any $i = 1, \ldots, t$, $\displaystyle\bigcap_{j<i} \bigcup_{H \in \mathrm{Sep}(C_i, C_j)} H$ is an affine subspace of $\mathbb{A}^d$.

3.5.3 Definition An enumeration $C_1, C_2, \ldots, C_t$ of $\mathrm{Ch}(A)$ is called a *valid order* if the condition ($\star$) is satisfied.

3.5.4 Perhaps surprisingly, the notion of a valid order is not new; when expressed combinatorially, it coincides with a notion introduced by Delucchi. One has the following theorem:

3.5.5 Theorem (Delucchi, Lofano–Paolini) *Every real hyperplane arrangement admits a valid order.*

3.5.6 Theorem 3.5.5 was proven by Delucchi in the case of a central hyperplane arrangement [7, Theorem 4.15]. The general case of Theorem 3.5.5 is due to Lofano and Paolini [25, Theorem 5.6], who also introduced the terminology "valid order". Lofano–Paolini's construction of valid orders is particularly simple: let x_0 be a generic point of our ambient real vector space, and suppose that the chambers are enumerated in any weakly increasing order by their euclidean distance from the point x_0, ie

$$d(C_i, x_0) < d(C_j, x_0) \implies i < j.$$

Then this is always a valid order. We remark that there is a small gap in Delucchi's argument for Theorem 3.5.5: he uses an induction on the number of hyperplanes in the arrangement, and that a total order on $\mathrm{Ch}(A)$ induces a total order on $\mathrm{Ch}(A \setminus \{H\})$ for any $H \in A$. Unfortunately the family of total orders considered by Delucchi is not closed under the operation of removing a hyperplane, which breaks the inductive strategy. Delucchi has communicated to us that if one runs the same argument using, for instance, the euclidean orderings of Lofano–Paolini, which are closed under removing a hyperplane, then the argument works.

3.5.7 Combining Theorem 3.5.5 with the observations of Section 3.5.2, we immediately deduce the following:

3.5.8 Corollary *For any real hyperplane arrangement A, the associated nonseparated scheme X_A admits an algebraic cell decomposition.*

3.5.9 Remark The cell decomposition of X_A obtained from Corollary 3.5.8 is a *refinement* of the stratification of X_A, in the sense that each affine cell is a union of strata. In this way we can think of the algebraic cell decomposition as an increasing filtration of the poset $\mathrm{Strat}(A)$, such that the union of strata that are in the k^{th} step of the filtration but not the $(k-1)^{\text{st}}$ form a cell.

3.5.10 Remark At first sight it is perhaps surprising that the question whether there always exists a total order on $\mathrm{Ch}(A)$ satisfying the strange condition $(\star)$ has already been considered (and answered) in the literature. But in fact Delucchi's motivation for introducing the condition $(\star)$ was not unrelated to ours. Recall that a CW complex is said to be *minimal* if its k^{th} Betti number coincides with the number of k–cells.

By a theorem due independently to Randell [29] and Dimca and Papadima [8], the complement of any complex hyperplane arrangement has the homotopy type of a minimal CW complex; see also [1]. On the other hand, as we have already mentioned, the complement of a complexified *real* arrangement of hyperplanes deformation retracts onto the *Salvetti complex*, and it is natural to try to make minimality explicit in this case by writing down an explicit perfect discrete Morse function on the Salvetti complex. Delucchi [7] showed that a valid order on $\mathrm{Ch}(A)$ induces a perfect discrete Morse function on $\mathrm{Sal}(A)$, and this was his motivation for introducing the notion of a valid order. What we have explained above is precisely that a valid order on $\mathrm{Ch}(A)$ induces a filtration of the poset $\mathrm{Strat}(A)$ — and hence also of $\mathrm{Sal}(A)$ — such that the union of strata within each successive difference is homeomorphic to an open cell. Delucchi showed instead that the poset of strata within each successive difference admits an acyclic matching with a single critical cell, so that these can be "glued together" to a discrete Morse function on the whole poset $\mathrm{Sal}(A)$. For another approach to minimality via discrete Morse theory on the Salvetti complex, see [32].

A related fact is that the filtration that we obtain is "minimal" in the sense that the cells form a basis for the Borel–Moore homology of the complement of the arrangement. See [15] for another quite similar construction of such stratifications of complements of complexified real hyperplane arrangements. In fact it is not hard to show that more is true: the cells make up a CW decomposition of the one-point compactification of the complement, in which the attaching maps are nullhomotopic by a straight-line homotopy, so that the one-point compactification is a wedge of spheres in a very explicit way. This works for both $M(A_{(3)})$ and $M(A_{\mathbb{C}})$.

3.5.11 Remark One can think of the notion of a valid ordering as an analogue of the notion of a shelling of a simplicial complex. Let us explain this. We consider all posets to have the Alexandroff topology, in which the open subsets are the downwards closed subsets. Then the poset $\mathrm{Strat}(A)$ has an evident open cover indexed by $\mathrm{Ch}(A)$, with each open subset isomorphic to $\mathcal{L}(A)$: explicitly, to $C \in \mathrm{Ch}(A)$ we associate the subposet

$$S_C = \{(K, D) \in \mathrm{Strat}(A) : C \subseteq D\} \cong \mathcal{L}(A).$$

The open cover $\mathrm{Strat}(A) = \bigcup_{C \in \mathrm{Ch}(A)} S_C$ corresponds geometrically to the open cover of the scheme X_A by the subschemes X_C of Section 3.2.3. In these terms we can define a valid order as an enumeration of these open sets, say $S_1, \dots, S_t$, such that the poset

$$S_i \setminus \bigcup_{j<i} S_j$$

has a unique minimum element for all i. If instead F is the face poset of a finite purely d–dimensional simplicial complex, then F has an open cover by subposets each isomorphic to the boolean poset of rank d (corresponding to the maximal-dimensional faces), and a shelling is an enumeration of the sets in this open cover satisfying precisely the same condition. However, we will not pursue this analogy further here.

3.6 Cohomology of nonseparated schemes via algebraic cell decompositions

3.6.1 In this section we explain why the cohomology of a smooth scheme with an algebraic cell decomposition (Definition 3.5.1) is torsion free, and why one can read off its Betti numbers from the number of cells in each dimension. This result is certainly well known, but we do not know a reference stating it in the generality we need, ie for the étale cohomology of potentially nonseparated schemes. But let us first recall the (easy) argument in the case of a complex algebraic variety.

3.6.2 The arguments are naturally phrased in terms of the *Borel–Moore homology* groups $\bar{H}_i(X,\mathbb{Z})$, which can be defined as the reduced homology groups of the one-point compactification of X. If X is smooth then Poincaré–Lefschetz duality gives an isomorphism $\bar{H}_i(X,\mathbb{Z}) \cong H^{2d-i}(X,\mathbb{Z})$, where $d = \dim_{\mathbb{C}}(X)$.

3.6.3 Proposition *Let X be a complex algebraic variety with an algebraic cell decomposition. Then the odd-degree Borel–Moore homology groups of X vanish, and for all i there is an isomorphism $\bar{H}_{2i}(X,\mathbb{Z}) \cong \mathbb{Z}^{\oplus n_i}$, where n_i denotes the number of cells of dimension i. In particular, if X is, in addition, smooth and of pure dimension d, then we also have that the odd cohomology groups of X vanish, and $H^{2i}(X,\mathbb{Z}) \cong \mathbb{Z}^{\oplus n_{d-i}}$, by Poincaré duality.*

Proof This is proven by induction on the number of cells, the base case being $X = \varnothing$. For the induction step we consider the filtration

$$\varnothing = U^{(0)} \subset U^{(1)} \subset \cdots \subset U^{(t)} = X$$

of Definition 3.5.1, with $U^{(t)} \setminus U^{(t-1)} \cong \mathbb{A}^d$. Then we know the result for $U^{(t-1)}$ by induction. To conclude the result for $X = U^{(t)}$ we use the long exact sequence in Borel–Moore homology

$$\cdots \to \bar{H}_i(U^{(t-1)},\mathbb{Z}) \to \bar{H}_i(X,\mathbb{Z}_\ell) \to \bar{H}_i(\mathbb{C}^d,\mathbb{Z}) \to \bar{H}_{i-1}(U^{(t-1)},\mathbb{Z}) \to \cdots$$

and the fact that $\bar{H}_i(\mathbb{C}^d,\mathbb{Z}) = 0$ for $i \neq 2d$, and $\bar{H}_{2d}(\mathbb{C}^d,\mathbb{Z}) \cong \mathbb{Z}$. □

3.6.4 Remark A common variant of Proposition 3.6.3 in the literature is to assume X proper but not necessarily smooth. For proper varieties Borel–Moore homology coincides with ordinary homology, and under the hypotheses of Proposition 3.6.3 one has $H^{2i}(X,\mathbb{Z}) \cong \mathbb{Z}^{\oplus n_i}$. If X is neither proper nor smooth, there is no reason that one should be able to determine its cohomology groups from an algebraic cell decomposition.

3.6.5 In generalizing Proposition 3.6.3 to the non-Hausdorff setting, one runs into the problem that Borel–Moore homology (or its dual, cohomology with compact support), and the Verdier duality theory, are typically only defined for Hausdorff spaces. (But see [26] for a basic theory of compactly supported sheaf cohomology for non-Hausdorff manifolds.) In this respect the situation in étale cohomology is in fact significantly nicer. Although textbook treatments of étale cohomology only define the functors $Rf_!$, $f^!$ for separated morphisms, this can now be completely sidestepped using the work of Laszlo and Olsson [22; 23]. Their main focus is to construct an ℓ–adic formalism of "six functors" for stacks, but even for schemes their lack of separation hypotheses is new. Their approach to the functors $Rf_!$ and $f^!$ is to first construct the Verdier duality functor $\mathbb{D}$, and then define

$$Rf_! = \mathbb{D} \circ Rf_* \circ \mathbb{D} \quad \text{and} \quad f^! = \mathbb{D} \circ f^* \circ \mathbb{D},$$

and this, crucially, does not require f to be separated.

3.6.6 Let us recall how to define Borel–Moore homology in terms of Grothendieck's six functors. For $f : X \to \operatorname{Spec}(k)$ a finite type scheme, say over a separably closed field, its étale Borel–Moore homology is defined as $\bar{H}_i^{\text{ét}}(X,\mathbb{Z}_\ell) = \mathcal{H}^{-i}(Rf_* f^! \mathbb{Z}_\ell)$; compare this with the usual étale cohomology $H^i_{\text{ét}}(X,\mathbb{Z}_\ell) = \mathcal{H}^i(Rf_* f^* \mathbb{Z}_\ell)$. This theory has the following properties:

(1) If X is smooth and of pure dimension d, then there is a Poincaré duality isomorphism

$$H^i_{\text{ét}}(X,\mathbb{Z}_\ell) \otimes \mathbb{Z}_\ell(d) \cong \bar{H}^{\text{ét}}_{2d-i}(X,\mathbb{Z}_\ell).$$

(2) If $U \subset X$ is open with closed complement Z, then there is a long exact sequence

$$\cdots \to \bar{H}_i^{\text{ét}}(Z,\mathbb{Z}_\ell) \to \bar{H}_i^{\text{ét}}(X,\mathbb{Z}_\ell) \to \bar{H}_i^{\text{ét}}(U,\mathbb{Z}_\ell) \to \bar{H}_{i-1}^{\text{ét}}(Z,\mathbb{Z}_\ell) \to \cdots.$$

The first property is immediate from our definition of Borel–Moore homology and the fact that $f^!(-) \cong f^*(-) \otimes \mathbb{Z}_\ell(d)[2d]$ if f is smooth of relative dimension d;

see [23, Lemma 9.1.2]. The second property can be proven by considering the distinguished triangle

$$i_* i^! \mathcal{F} \to \mathcal{F} \to R j_* j^! \mathcal{F} \xrightarrow{+1},$$

where $j : U \to X$ and $i : Z \to X$ are the inclusions, setting $\mathcal{F} = f^! \mathbb{Z}_\ell$ where $f : X \to \mathrm{Spec}(k)$ is the structural morphism, and applying Rf_*. The distinguished triangle is constructed in case of finite coefficients in [22] as equation (4.10.ii), and the same reasoning works in the ℓ–adic case.

3.6.7 Proposition *Let X be a finite type scheme over a base field k equipped with an algebraic cell decomposition. Then the odd-degree Borel–Moore homology groups of X vanish, and for all i there is an isomorphism $\bar{H}^{\text{ét}}_{2i}(X, \mathbb{Z}_\ell) \cong \mathbb{Z}_\ell(i)^{\oplus n_i}$, where n_i denotes the number of cells of dimension i. In particular, if X is in addition smooth and of pure dimension d, then we also have that the odd cohomology groups of X vanish, and $H^{2i}_{\text{ét}}(X, Z_\ell) \cong \mathbb{Z}_\ell(-i)^{\oplus n_{d-i}}$, by Poincaré duality.*

Proof This can now be proven in an identical manner to Proposition 3.6.3. □

3.6.8 Remark A fortiori we also see that the analogue of Proposition 3.6.3 remains valid for the singular cohomology of nonseparated complex schemes, by applying the ℓ–adic Proposition 3.6.7 and Artin's comparison isomorphism.

References

[1] **K A Adiprasito**, *Combinatorial stratifications and minimality of 2–arrangements*, J. Topol. 7 (2014) 1200–1220 MR Zbl

[2] **K A Behrend**, *Derived l–adic categories for algebraic stacks*, Mem. Amer. Math. Soc. 774, Amer. Math. Soc., Providence, RI (2003) MR Zbl

[3] **A Björner**, **G M Ziegler**, *Combinatorial stratification of complex arrangements*, J. Amer. Math. Soc. 5 (1992) 105–149 MR Zbl

[4] **E Brieskorn**, *Sur les groupes de tresses (d'après V I Arnold)*, from "Séminaire Bourbaki, 1971/1972", Lecture Notes in Math. 317, Springer (1973) exposé 401, 21–44 MR Zbl

[5] **T Church**, **J S Ellenberg**, **B Farb**, *Representation stability in cohomology and asymptotics for families of varieties over finite fields*, from "Algebraic topology: applications and new directions" (U Tillmann, S Galatius, D Sinha, editors), Contemp. Math. 620, Amer. Math. Soc., Providence, RI (2014) 1–54 MR Zbl

[6] **B Conrad**, **M Temkin**, *Non-Archimedean analytification of algebraic spaces*, J. Algebraic Geom. 18 (2009) 731–788 MR Zbl

[7] **E Delucchi**, *Shelling-type orderings of regular CW-complexes and acyclic matchings of the Salvetti complex*, Int. Math. Res. Not. 2008 (2008) art. id. rnm167 MR Zbl

[8] **A Dimca**, **S Papadima**, *Hypersurface complements, Milnor fibers and higher homotopy groups of arrangements*, Ann. of Math. 158 (2003) 473–507 MR Zbl

[9] **R Fox**, **L Neuwirth**, *The braid groups*, Math. Scand. 10 (1962) 119–126 MR Zbl

[10] **D B Fuchs**, *Cohomology of the braid group* mod 2, Funkcional. Anal. i Priložen. 4 (1970) 62–73 MR Zbl In Russian; translated in Funct. Anal. Appl. 4 (1970) 143–151

[11] **W Fulton**, **R MacPherson**, *A compactification of configuration spaces*, Ann. of Math. 139 (1994) 183–225 MR Zbl

[12] **N Gadish**, *A trace formula for the distribution of rational G–orbits in ramified covers, adapted to representation stability*, New York J. Math. 23 (2017) 987–1011 MR Zbl

[13] **M Goresky**, **R MacPherson**, *Stratified Morse theory*, Ergeb. Math. Grenzgeb. 14, Springer (1988) MR Zbl

[14] **T Hyde**, *Polynomial factorization statistics and point configurations in $\mathbb{R}^3$*, Int. Math. Res. Not. 2020 (2020) 10154–10179 MR Zbl

[15] **K-K Ito**, **M Yoshinaga**, *Semi-algebraic partition and basis of Borel–Moore homology of hyperplane arrangements*, Proc. Amer. Math. Soc. 140 (2012) 2065–2074 MR Zbl

[16] **R Jiménez Rolland**, **J C H Wilson**, *Stability for hyperplane complements of type B/C and statistics on squarefree polynomials over finite fields*, Q. J. Math. 70 (2019) 565–602 MR Zbl

[17] **R Kane**, *Reflection groups and invariant theory*, CMS Books Math. 5, Springer (2001) MR Zbl

[18] **M Kim**, *Weights in cohomology groups arising from hyperplane arrangements*, Proc. Amer. Math. Soc. 120 (1994) 697–703 MR Zbl

[19] **M Kisin**, **G I Lehrer**, *Equivariant Poincaré polynomials and counting points over finite fields*, J. Algebra 247 (2002) 435–451 MR Zbl

[20] **D Knutson**, *Algebraic spaces*, Lecture Notes in Math. 203, Springer (1971) MR Zbl

[21] **J Kollár**, *Cone theorems and bug-eyed covers*, J. Algebraic Geom. 1 (1992) 293–323 MR Zbl

[22] **Y Laszlo**, **M Olsson**, *The six operations for sheaves on Artin stacks, I: Finite coefficients*, Publ. Math. Inst. Hautes Études Sci. 107 (2008) 109–168 MR Zbl

[23] **Y Laszlo**, **M Olsson**, *The six operations for sheaves on Artin stacks, II: Adic coefficients*, Publ. Math. Inst. Hautes Études Sci. 107 (2008) 169–210 MR Zbl

[24] **G I Lehrer**, *The l–adic cohomology of hyperplane complements*, Bull. Lond. Math. Soc. 24 (1992) 76–82 MR Zbl

[25] **D Lofano**, **G Paolini**, *Euclidean matchings and minimality of hyperplane arrangements*, Discrete Math. 344 (2021) art. id. 112232 MR Zbl

[26] **I Moerdijk**, **J Mrčun**, *Lie groupoids, sheaves and cohomology*, from "Poisson geometry, deformation quantisation and group representations" (S Gutt, J Rawnsley, D Sternheimer, editors), Lond. Math. Soc. Lect. Note Ser. 323, Cambridge Univ. Press (2005) 145–272 MR Zbl

[27] **D Moseley**, *Equivariant cohomology and the Varchenko–Gelfand filtration*, J. Algebra 472 (2017) 95–114 MR Zbl

[28] **N Proudfoot**, *A non-Hausdorff model for the complement of a complexified hyperplane arrangement*, Proc. Amer. Math. Soc. 135 (2007) 3989–3994 MR Zbl

[29] **R Randell**, *Morse theory, Milnor fibers and minimality of hyperplane arrangements*, Proc. Amer. Math. Soc. 130 (2002) 2737–2743 MR Zbl

[30] **T Ransford**, *Potential theory in the complex plane*, Lond. Math. Soc. Stud. Texts 28, Cambridge Univ. Press (1995) MR Zbl

[31] **M Salvetti**, *Topology of the complement of real hyperplanes in $\mathbb{C}^N$*, Invent. Math. 88 (1987) 603–618 MR Zbl

[32] **M Salvetti**, **S Settepanella**, *Combinatorial Morse theory and minimality of hyperplane arrangements*, Geom. Topol. 11 (2007) 1733–1766 MR Zbl

[33] **G C Shephard**, **J A Todd**, *Finite unitary reflection groups*, Canad. J. Math. 6 (1954) 274–304 MR Zbl

Department of Mathematics, Stockholm University
Stockholm, Sweden

Department of Mathematics, University of Chicago
Chicago, IL, United States

dan.petersen@math.su.se, ptoste@math.uchicago.edu

Proposed: Mladen Bestvina Received: 17 June 2020
Seconded: David M Fisher, Dan Abramovich Revised: 13 October 2020

GEOMETRY & TOPOLOGY

msp.org/gt

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

The subscription price for 2021 is US $635/year for the electronic version, and $910/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Geometry & Topology is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

Geometry & Topology (ISSN 1465-3060 printed, 1364-0380 electronic) is published 7 times per year and continuously online, by Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840.

GT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

GEOMETRY & TOPOLOGY

Volume 25 Issue 7 (pages 3257–3753) 2021