Factorization Statistics and the
Twisted Grothendieck-Lefschetz formula

Trevor Hyde*

Dept. of Mathematics
University of Michigan
Ann Arbor, MI 48109-1043

Abstract. Factorization statistics are functions defined on the set $\mathrm{Poly}_d(\mathbb{F}_q)$ of all monic degree d polynomials with coefficients in $\mathbb{F}_q$ which only depend on the degrees of the irreducible factors of a polynomial. We show that the expected values of factorization statistics are determined by the representation theoretic structure of the cohomology of point configurations in $\mathbb{R}^3$. This *twisted Grothendieck-Lefschetz formula for* Poly_d is analogous to a result of Church, Ellenberg, and Farb for *squarefree* polynomials. Our proof uses formal power series methods which also lead to a new proof of the Church, Ellenberg, and Farb result circumventing algebraic geometry.

Keywords: arithmetic statistics, symmetric group representations, configuration space

1 Introduction

What is the probability that a random integer m in the interval $[1, n]$ is prime? The Prime Number Theorem tells us that

$$\mathrm{Prob}(m \in [1, n] \text{ is prime}) \approx \frac{1}{\log(n)}$$

for sufficiently large n. Following a classic analogy between $\mathbb{Z}$ and $\mathbb{F}_q[x]$, we ask: what is the probability that a random monic degree d polynomial $f(x) \in \mathbb{F}_q[x]$ is irreducible? One can show that

$$\mathrm{Prob}(f(x) \text{ monic degree } d \text{ is irreducible}) \approx \frac{1}{d} \tag{1.1}$$

for large values of q. Note that the number of monic degree d polynomials in $\mathbb{F}_q[x]$ is q^d, hence $\frac{1}{d} = \frac{1}{\log_q(q^d)}$ parallels the result for $\mathbb{Z}$. This is the beginning of a motivating theme: analogous arithmetic statistical questions for $\mathbb{Z}$ and $\mathbb{F}_q[x]$ have analogous answers.

Often we may determine the exact values of statistics on the $\mathbb{F}_q[x]$ side of the analogy which seem out of reach for $\mathbb{Z}$. For example, the number of irreducible monic degree d polynomials in

---

*tghyde@umich.edu

$\mathbb{F}_q[x]$ is given by dth *necklace polynomial*,

$$\#\{f(x) \text{ irreducible monic degree } d\} = \frac{1}{d} \sum_{e|d} \mu(d/e)q^e,$$

where μ is the Möbius function. Therefore,

$$\text{Prob}(f(x) \text{ monic degree } d \text{ is irreducible}) = \frac{1}{d} \sum_{e|d} \frac{\mu(d/e)}{q^{d-e}}. \tag{1.2}$$

From an analytic point of view there is an impulse to focus on the leading terms for a probability like (1.1). However, on closer inspection of the precise formula (1.2), we see that each term has a structural interpretation. Consider the case of (1.2) when $d = 6$,

$$\text{Prob}(f(x) \text{ monic degree } 6 \text{ is irreducible}) = \frac{1}{6}\left(1 - \frac{1}{q^3} - \frac{1}{q^4} + \frac{1}{q^5}\right).$$

The four terms in this expression correspond to the intermediate fields of the degree 6 extension $\mathbb{F}_{q^6}/\mathbb{F}_q$ and the coefficients encode how these fields fit together.

This brings us to our main thesis: the exact expressions for arithmetic statistical questions in $\mathbb{F}_q[x]$ reflect hidden structure which is not apparent from approximations. In other words, *there are no error terms*, each term has an interpretation and together they tell a complete story. Our main result (Theorem 1.1) supports this claim for the expected values of functions on $\mathbb{F}_q[x]$.

A *factorization statistic* P is a function defined on the set $\text{Poly}_d(\mathbb{F}_q)$ of monic degree d polynomials in $\mathbb{F}_q[x]$ such that $P(f)$ depends only on the partition of d given by the degrees of the irreducible factors of f. P may also be viewed as a function defined on partitions of d, or as a class function of the symmetric group S_d. Let ψ_d^k be the character of the S_d-representation $H^{2k}(\text{PConf}_d(\mathbb{R}^3),\mathbb{Q})$ where $\text{PConf}_d(\mathbb{R}^3)$ is the ordered configuration space of d distinct points in $\mathbb{R}^3$ (see Section 3.)

Theorem 1.1 (Twisted Grothendieck-Lefschetz for Poly_d)**.** *If P is a factorization statistic, then the expected value $E_d(P)$ of P on $\text{Poly}_d(\mathbb{F}_q)$ is given by*

$$E_d(P) := \frac{1}{q^d} \sum_{f \in \text{Poly}_d(\mathbb{F}_q)} P(f) = \sum_{k=0}^{d-1} \frac{\langle P, \psi_d^k \rangle}{q^k},$$

where $\langle P, \psi_d^k \rangle = \frac{1}{d!} \sum_{\sigma \in S_d} P(\sigma)\psi_d^k(\sigma)$ is the standard inner product of class functions of the symmetric group S_d.

The twisted Grothendieck-Lefschetz formula provides a bridge between representation theory and topology on the one hand and the combinatorics of finite fields on the other. We explore this interplay through examples in Section 4.

2   Factorization statistics

The *factorization type* of a polynomial $f(x) \in \mathbb{F}_q[x]$ is the partition of $\deg(f)$ given by the degrees of the irreducible factors of $f(x)$. Let $\mathrm{Poly}_d(\mathbb{F}_q)$ denote the set of monic degree d polynomials in $\mathbb{F}_q[x]$. A *factorization statistic P* is a function defined on polynomials $f(x) \in \mathrm{Poly}_d(\mathbb{F}_q)$ which only depends on the factorization type of $f(x)$.

Example 2.1.

1. Consider the polynomials $g(x), h(x) \in \mathrm{Poly}_5(\mathbb{F}_3)$ with irreducible factorizations

$$g(x) = x^2(x+1)(x^2+1) \qquad h(x) = (x+1)(x-1)(x^3-x+1).$$

   The factorization type of $g(x)$ is the partition $[2,1,1,1]$ and the factorization type of $h(x)$ is $[3,1,1]$. Note that the factorization type does not record the multiplicity of a specific factor so that x^2 and $x(x+1)$ both have the same factorization type $[1,1]$.

2. Let $R(f)$ be the number of $\mathbb{F}_q$-roots of $f(x) \in \mathrm{Poly}_d(\mathbb{F}_q)$. Then $R(f)$ depends only on the number of linear factors of $f(x)$, hence is a factorization statistic. Referring to the two polynomials above we have $R(g) = 3$ and $R(h) = 2$.

3. Say a polynomial $f(x)$ has *even type* if the factorization type of $f(x)$ is an even partition. More specifically, say $\lambda = (1^{m_1} 2^{m_2} 3^{m_3} \cdots)$ is the factorization type of $f(x)$ and define $\mathrm{sgn}(\lambda)$ by

$$\mathrm{sgn}(\lambda) = \prod_{j \geq 1} (-1)^{m_j(j-1)},$$

   then $f(x)$ has even type if $\mathrm{sgn}(\lambda) = 1$. The function ET defined by

$$ET(f) = \begin{cases} 1 & f(x) \text{ has even type} \\ 0 & \text{otherwise,} \end{cases}$$

   is a factorization statistic. Continuing our examples, $ET(g) = 0$ and $ET(h) = 1$.

4. Define the *quadratic excess $Q(f)$* of a polynomial f to be

$$Q(f) = \#\{\text{reducible quadratic factors of } f(x)\} - \#\{\text{irreducible quadratic factors of } f(x)\}.$$

   Then $Q(f)$ depends only on the number of linear and irreducible quadratic factors of $f(x)$, hence is a factorization statistic. Since $g(x)$ has 3 linear factors and 1 irreducible quadratic factor, we have

$$Q(g) = \binom{3}{2} - 1 = 2.$$

   The polynomial $h(x)$ has 2 linear factors and 0 irreducible quadratic factors, hence

$$Q(h) = \binom{2}{2} - 0 = 1.$$

Let $E_d(P)$ denote the expected value of a factorization statistic P on the set $\text{Poly}_d(\mathbb{F}_q)$ of all monic degree d polynomials. More precisely,

$$E_d(P) := \frac{1}{q^d} \sum_{f \in \text{Poly}_d(\mathbb{F}_q)} P(f).$$

By counting the number of polynomials with a given factorization type (e.g. using unique factorization and necklace polynomials) we can explicitly compute $E_d(P)$ for any particular P and d as a function of q. For example, here are some computations of $E_d(Q)$ where Q is the quadratic excess statistic defined in Example 2.1 (4).

d	$E_d(Q)$
3	$\frac{2}{q} + \frac{1}{q^2}$
4	$\frac{2}{q} + \frac{2}{q^2} + \frac{2}{q^3}$
5	$\frac{2}{q} + \frac{2}{q^2} + \frac{4}{q^3} + \frac{2}{q^4}$
6	$\frac{2}{q} + \frac{2}{q^2} + \frac{4}{q^3} + \frac{4}{q^4} + \frac{3}{q^5}$
10	$\frac{2}{q} + \frac{2}{q^2} + \frac{4}{q^3} + \frac{4}{q^4} + \frac{6}{q^5} + \frac{6}{q^6} + \frac{8}{q^7} + \frac{8}{q^8} + \frac{5}{q^9}$

There are some remarkable features of these expected values: $E_d(Q)$ is a polynomial in $\frac{1}{q}$ of degree $d - 1$ with *positive integer coefficients*—one should expect the coefficients to be rational numbers, but both the positivity and integrality are not a priori evident. Evaluating the polynomial $E_d(Q)$ at $q = 1$ gives the binomial coefficient $\binom{d}{2}$. The coefficients of $E_d(Q)$ appear to stabilize as d increases with a clear pattern emerging already for $d = 10$, suggesting that the expected values $E_d(Q)$ converge coefficientwise as $d \to \infty$. We return to this example in Section 4.1 to explain these observations in the light of our results.

3   Twisted Grothendieck-Lefschetz formulas

We briefly detour from our discussion of factorization statistics and finite fields to review some topology. If X is a topological space, then $\text{PConf}_d(X)$ is

$$\text{PConf}_d(X) := \{(x_1, x_2, \ldots, x_d) \in X^d : x_i \neq x_j\},$$

the *ordered configuration space of d points on X*. The symmetric group S_d acts on $\text{PConf}_d(X)$ by permuting coordinates; this action is free given that all coordinates are distinct. Let $\text{Conf}_d(X)$ be the quotient of $\text{PConf}_d(X)$ by this action. $\text{Conf}_d(X)$ is the space of *unordered configurations of d points on X*. Note that for each $k \geq 0$ the singular cohomology $H^k(\text{PConf}_d(X), \mathbb{Q})$ is, by functoriality, a finite dimensional S_d-representation.

Our main result establishes a surprising connection between the expected values of factorization statistics on $\mathrm{Poly}_d(\mathbb{F}_q)$ and the sequence of S_d-representations $H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$. The cohomology $H^*(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$ is supported in even degrees, vanishing beyond degree $2(d-1)$.

Theorem 3.1 (Twisted Grothendieck-Lefschetz formula for Poly_d). *Let P be a factorization statistic and let ψ_d^k be the character of the S_d-representation $H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$. Then the expected value $E_d(P)$ of P on the set $\mathrm{Poly}_d(\mathbb{F}_q)$ of polynomials is given by*

$$E_d(P) := \frac{1}{q^d} \sum_{f \in \mathrm{Poly}_d(\mathbb{F}_q)} P(f) = \sum_{k=0}^{d-1} \frac{\langle P, \psi_d^k \rangle}{q^k},$$

where $\langle P, \psi_d^k \rangle = \frac{1}{d!} \sum_{\sigma \in S_d} P(\sigma) \psi_d^k(\sigma)$ is the standard inner product of S_d-class functions.

Theorem 3.1 shows that the coefficients of the expected value $E_d(P)$ are determined by the representation theoretic structure of $H^*(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$ for any factorization statistic P. Note that the factorization statistic P plays different roles on each side of this equation: on the left it acts as a function on the set $\mathrm{Poly}_d(\mathbb{F}_q)$ of polynomials; on the right it acts as a class function of the symmetric group S_d.

The sequence of representations $H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$ has another interpretation as the *higher Lie representations* Lie_k [5, Sec. 2.6]. We express Theorem 3.1 in terms of the cohomology of point configurations in $\mathbb{R}^3$ to parallel the following result of Church, Ellenberg, and Farb.

Theorem 3.2 ([3, Prop. 4.1]). *Let P be a factorization statistic, and let ϕ_d^k be the character of the S_d-representation $H^k(\mathrm{PConf}_d(\mathbb{R}^2), \mathbb{Q})$. If $\mathrm{Poly}_d^{\mathrm{sf}}(\mathbb{F}_q)$ is the set of squarefree polynomials of degree d in $\mathbb{F}_q[x]$, then*

$$\frac{1}{q^d} \sum_{f \in \mathrm{Poly}_d^{\mathrm{sf}}(\mathbb{F}_q)} P(f) = \sum_{k=0}^{d-1} \frac{(-1)^k \langle P, \phi_d^k \rangle}{q^k}.$$

The proof of Theorem 3.2 in [3] uses algebraic geometry: viewing PConf_d as a scheme defined over $\mathbb{Z}$, the Grothendieck-Lefschetz trace formula for étale cohomology with "twisted coefficients" expresses the weighted point counts on $\mathrm{Conf}_d(\mathbb{F}_q)$ in terms of the trace of Frobenius. This combined with a purity result and a comparison theorem between étale and singular cohomology yields Theorem 3.2.

To see the connection between squarefree polynomials and point configurations in $\mathbb{R}^2$, we view the plane as $\mathbb{C}$ and note there is a natural correspondence between squarefree polynomials over $\mathbb{C}$ of degree d and unordered configurations of d distinct points in $\mathbb{C}$:

$$(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_d) \quad \longleftrightarrow \quad \{\alpha_1, \alpha_2, \ldots, \alpha_d\}$$

This correspondence extends to an isomorphism of schemes $\mathrm{Poly}_d^{\mathrm{sf}} \cong \mathrm{Conf}_d$.

The geometric perspective behind their proof appears to break down in the case of Theorem 3.1, as there is no clear connection between general polynomials of degree d and point configurations in $\mathbb{R}^3$. Instead we prove Theorem 3.1 using a generating function argument. Our approach also leads to a new proof of Theorem 3.2, circumventing the methods of algebraic geometry entirely.

We outline the strategy for Theorem 3.1. A complete proof appears in [7]. Consider the *splitting measure* ν defined on partitions $\lambda \vdash d$ by

$$\nu(\lambda) = \mathrm{Prob}(f \in \mathrm{Poly}_d(\mathbb{F}_q) \text{ has factorization type } \lambda).$$

One may show that $\nu(\lambda)$ is a polynomial in $1/q$ with rational coefficients for any partition λ. The connection between $\mathrm{Poly}_d(\mathbb{F}_q)$ and $H^*(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$ is made through the following result.

Theorem 3.3 ([7, Thm. 1.4]). *Let ψ_d^k be the character of $H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$. If $\lambda \vdash d$ is a partition, let $z_\lambda = \prod_{j \geq 1} j^{m_j} m_j!$ when $\lambda = (1^{m_1} 2^{m_2} \cdots)$. Then*

$$\nu(\lambda) = \frac{1}{z_\lambda} \sum_{k=0}^{d-1} \frac{\psi_d^k(\lambda)}{q^k}.$$

Theorem 3.3 is deduced with the help of a beautiful product formula for the cycle index series of the family $H^*(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$ of representations which may be found in Hersh and Reiner [5, Thm. 2.7]. Once we have this result, Theorem 3.1 follows by a change in the order of summation.

Proof of Theorem 3.1. Since factorization statistics depend only on the factorization type of a polynomial, we may rewrite the expected value in terms of the splitting measure,

$$E_d(P) = \frac{1}{q^d} \sum_{f \in \mathrm{Poly}_d(\mathbb{F}_q)} P(f) = \sum_{\lambda \vdash d} P(\lambda)\nu(\lambda).$$

Then Theorem 3.3 implies,

$$\sum_{\lambda \vdash d} P(\lambda)\nu(\lambda) = \sum_{\lambda \vdash d} \frac{1}{z_\lambda} \sum_{k=0}^{d-1} \frac{P(\lambda)\psi_d^k(\lambda)}{q^k} = \sum_{k=0}^{d-1} \frac{1}{q^k} \left(\sum_{\lambda \vdash d} \frac{P(\lambda)\psi_d^k(\lambda)}{z_\lambda} \right) = \sum_{k=0}^{d-1} \frac{\langle P, \psi_d^k \rangle}{q^k}.$$

$\square$

Church, Ellenberg, and Farb combine Theorem 3.2 with the *representation stability* of $H^k(\mathrm{PConf}_d(\mathbb{R}^2), \mathbb{Q})$ to deduce the asymptotic stability of squarefree factorization statistics. A sequence V_d of S_d-representations is called representation stable when the decomposition of V_d into irreducibles stabilizes as $d \to \infty$. We refer the reader to [4] for a precise description of representation stability. For us the important fact is the following: Let x_j be the function defined on partitions where $x_j(\lambda)$ is the number of parts of λ os size j. Suppose P is an element in $\mathbb{Q}[x_1, x_2, \ldots]$, then P defines a function on partitions called a *character polynomial*. If V_d is a representation stable sequence with character χ_d and P is a character polynomial, then the following limit exists

$$\langle P, \chi \rangle = \lim_{d \to \infty} \langle P, \chi_d \rangle.$$

Furthermore, the sequence $\langle P, \chi_d \rangle$ is eventually constant (see [3, Sec. 3.4].) It follows from a general result of Church [2, Thm. 1] that for each k the sequence $H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$ is representation stable. Thus we have the following corollary of Theorem 3.1.

Corollary 3.4 (Asymptotic stability of expected values)**.** *If P is a factorization statistic given by a character polynomial, then the following limit converges coefficientwise in the ring of formal power series in $1/q$:*

$$\lim_{d \to \infty} E_d(P) = \sum_{k=0}^{\infty} \frac{\langle P, \psi^k \rangle}{q^k},$$

where $\langle P, \psi^k \rangle := \lim_{d \to \infty} \langle P, \psi_d^k \rangle$.

4   Examples

In this section we explore the interplay provided by Theorem 3.1 between finite field combinatorics, representation theory, and topology through examples.

4.1   Quadratic excess

Recall the quadratic excess factorization statistic Q from Section 2: $Q(f)$ is defined as the difference between the number of reducible versus irreducible quadratic factors of f. Rephrasing this in terms of partitions we see that Q is given by the character polynomial

$$Q(\lambda) = \binom{x_1(\lambda)}{2} - \binom{x_2(\lambda)}{1}.$$

Let $\mathbb{Q}[d]$ be the permutation representation of the symmetric group with basis $\{e_1, e_2, \ldots, e_d\}$ and consider the linear representation given by the second exterior power $\bigwedge^2 \mathbb{Q}[d]$. This representation has dimension $\binom{d}{2}$ with basis given by $\{e_i \wedge e_j : i < j\}$. If $\sigma \in S_d$ is a permutation, then the trace of σ on $\bigwedge^2 \mathbb{Q}[d]$ is

$$\mathrm{Trace}(\sigma) = \#\{\{i,j\} : \sigma \text{ fixes } i \text{ and } j\} - \#\{\{i,j\} : \sigma \text{ transposes } i \text{ and } j\}$$
$$= \binom{x_1(\sigma)}{2} - \binom{x_2(\sigma)}{1}$$
$$= Q(\sigma).$$

Thus Q, viewed as a class function of S_d, is the character of $\bigwedge^2 \mathbb{Q}[d]$. It follows that $\langle Q, \psi_d^k \rangle$ is a non-negative integer for all $d, k \geq 0$. This together with Theorem 3.1 explains the non-negative integral coefficients of $E_d(Q)$. That the degree of $E_d(Q)$ is $d - 1$ reflects that $2(d - 1)$ is the largest non-vanishing degree of cohomology for $\mathrm{PConf}_d(\mathbb{R}^3)$. The coefficientwise convergence of $E_d(Q)$ follows from Corollary 3.4.

The coefficientwise convergence of $E_d(P)$ holds in much greater generality for functions P defined on $\mathrm{Conf}_d(V)$ for V an affine or projective variety defined over $\mathbb{F}_q$ which only depend on the cycle structure of Frobenius, even when there is no apparent representation stability present; we do not pursue this further here but refer the reader to [1, Cor. 10]. One benefit of the combinatorial approach is that we can explicitly compute the limits of $E_d(P)$ as a rational function of q. For example,

$$
\begin{aligned}
\lim_{d\to\infty} E_d(Q) &= \frac{1}{2}\left(1+\frac{1}{q}\right)\left(\frac{1}{1-\frac{1}{q}}\right)^2 - \frac{1}{2}\left(1-\frac{1}{q}\right)\left(\frac{1}{1-\frac{1}{q^2}}\right) \\
&= \frac{2}{q} + \frac{2}{q^2} + \frac{4}{q^3} + \frac{4}{q^4} + \frac{6}{q^5} + \frac{6}{q^6} + \frac{8}{q^7} + \frac{8}{q^8} + \frac{10}{q^9} + \cdots
\end{aligned}
$$

4.2   Constraint on total cohomology

The next result gives a constraint on the total cohomology of $\mathrm{PConf}_d(\mathbb{R}^3)$.

Theorem 4.1. *For each $d \geq 0$ there is an isomorphism of S_d-representations*

$$
\bigoplus_{k=0}^{d-1} H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q}) \cong \mathbb{Q}[S_d], \tag{4.1}
$$

where $\mathbb{Q}[S_d]$ is the regular representation of S_d.

Proof. Let ρ be the character of $\bigoplus_{k=0}^{d-1} H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$. Then

$$
\rho = \sum_{k=0}^{d-1} \psi_d^k,
$$

where ψ_d^k is the character of $H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$. By Theorem 3.3 we have

$$
\nu(\lambda) = \frac{1}{z_\lambda} \sum_{k=0}^{d-1} \frac{\psi_d^k(\lambda)}{q^k},
$$

where ν is the splitting measure. Let ν_1 denote the splitting measure evaluated at $q = 1$. Then $\nu_1(\lambda) = \frac{\rho(\lambda)}{z_\lambda}$. On the other hand we can compute $\nu_1(\lambda)$ directly. The number of irreducible polynomials in $\mathrm{Poly}_d(\mathbb{F}_q)$ is given by

$$
M_d(q) = \frac{1}{d} \sum_{e|d} \mu(e) q^{d/e}.
$$

Hence by unique factorization in $\mathbb{F}_q[x]$,

$$
\nu(\lambda) = \frac{1}{q^d} \prod_{j \geq 1} \binom{M_j(q) + m_j - 1}{m_j}.
$$

Since $M_j(1) = 0$ for $j > 1$ and $M_1(1) = 1$ it follows that

$$\nu_1(\lambda) = \prod_{j \geq 1} \binom{M_j(1) + m_j - 1}{m_j} = \begin{cases} 1 & \lambda = [1^d] \\ 0 & \text{otherwise.} \end{cases}$$

Since $z_{[1^d]} = d!$, $\nu_1(\lambda) = \frac{\rho(\lambda)}{z_\lambda}$ implies

$$\rho(\lambda) = \begin{cases} d! & \lambda = [1^d] \\ 0 & \text{otherwise,} \end{cases}$$

which is the character of the regular representation. □

The right hand side of (4.1) is well-understood: the irreducible representations of S_d are indexed by partitions $\lambda \vdash d$, each irreducible $\mathcal{S}_\lambda$ is a direct summand of $\mathbb{Q}[S_d]$ with multiplicity $f_\lambda := \dim \mathcal{S}_\lambda$. Thus Theorem 4.1 tells us that all the irreducible components of $\mathbb{Q}[S_d]$ are distributed among the various degrees of cohomology on the left hand side of (4.1). Theorem 3.1 implies that this filtration of the regular representation completely determines and is determined by the expected values of factorization statistics on $\mathrm{Poly}_d(\mathbb{F}_q)$. We use this information to locate some of the irreducible S_d-representations in the cohomology of $\mathrm{PConf}_d(\mathbb{R}^3)$.

4.2.1 Trivial representation

Let $\mathbf{1} = \mathcal{S}_{[d]}$ be the trivial representation of S_d. Recall that the trivial representation $\mathbf{1}$ is one dimensional with constant character equal to 1. By Theorem 4.1 there is precisely one k such that $\mathbf{1}$ is a summand of $H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$. Interpreting the character of $\mathbf{1}$ as a factorization statistic we have $E_d(\mathbf{1}) = 1$ and Theorem 3.1 implies

$$1 = E_d(\mathbf{1}) = \sum_{k=0}^{d-1} \frac{\langle \mathbf{1}, \psi_d^k \rangle}{q^k}.$$

Comparing coefficients of $1/q^k$ we conclude that $\mathbf{1}$ is a summand of $H^0(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$. On the other hand, $\mathrm{PConf}_d(\mathbb{R}^3)$ is path connected so the degree 0 cohomology is one dimensional. Thus

$$H^0(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q}) \cong \mathbf{1}.$$

Note that any factorization statistic P is a class function of S_d and the irreducible characters of S_d form a $\mathbb{Q}$-basis for the vector space of all class functions. Thus there are $a_\lambda \in \mathbb{Q}$ such that

$$P = \sum_{\lambda \vdash d} a_\lambda \chi_\lambda,$$

where χ_λ is the character of the irreducible representation $\mathcal{S}_\lambda$. In particular if $a_1 := a_{[d]}$ is the coefficient of the trivial character in this decomposition, then we have the following corollary.

Corollary 4.2. *If P is any factorization statistic and a_1 is the coefficient of the trivial character in the canonical expression for P as a linear combination of irreducible characters, then*

$$a_1 = \lim_{q \to \infty} E_d(P).$$

Hence $a_1 = 0$ if and only if the expected value of P approaches 0 for large q.

Our table of values for $E_d(Q)$ with Q the quadratic excess show that $\lim_{q \to \infty} E_d(Q) = 0$ for each d, hence the representation $\bigwedge^2 Q[d]$ has no trivial component.

4.2.2 Sign representation

The only other one dimensional irreducible representation of S_d is the sign representation **Sgn** $:=$ $\mathcal{S}_{[1^d]}$ whose character we write as sgn. Viewing sgn as a factorization statistic Theorem 3.1 implies

$$E_d(\text{sgn}) = \frac{1}{q^k}$$

for some $k > 0$, but which value of k is it?

Theorem 4.3. *For each $d \geq 0$,*

$$E_d(\text{sgn}) = \frac{1}{q^{\lfloor d/2 \rfloor}}.$$

*Hence $H^{2\lfloor d/2 \rfloor}(\text{PConf}_d(\mathbb{R}^3), Q)$ is the unique cohomological degree with a **Sgn** summand.*

We prove Theorem 4.3 using *liminal reciprocity* which relates factorization statistics in $\text{Poly}_d(\mathbb{F}_q)$ with the limiting values of *squarefree* factorization statistics for $\mathbb{F}_q[x_1, x_2, \ldots, x_n]$ as the number of variables n tends to infinity. See [6] for details.

Theorem 4.3 has a surprising consequence. Recall that ET is the *even type* factorization statistic defined as $ET(f) = 1$ when the factorization type of f is an even partition and 0 otherwise. Thus the expected value $E_d(ET)$ is the probability of a random polynomial in $\text{Poly}_d(\mathbb{F}_q)$ having even factorization type. One might guess that a polynomial should be just as likely to have an even versus odd factorization type. However, notice that as class functions of S_d we have

$$ET = \tfrac{1}{2}(1 + \text{sgn}).$$

It follows by the linearity of expectation that

$$E_d(ET) = \tfrac{1}{2}(E_d(1) + E_d(\text{sgn})) = \tfrac{1}{2}\left(1 + \tfrac{1}{q^{\lfloor d/2 \rfloor}}\right).$$

The leading term of this probability is $1/2$ as we expected, but there is a bias toward a polynomial having even factorization type coming from the sign representation and the degree of cohomology in which it appears. For comparison we remark that in the squarefree case the probability of a random polynomial in $\text{Poly}_d^{\text{sf}}(\mathbb{F}_q)$ having even factorization type is exactly

$$E_d^{\text{sf}}(ET) = \tfrac{1}{2},$$

matching our original guess.

4.2.3 Standard representation

Recall the factorization statistic R from Example 2.1 where $R(f)$ is the number of $\mathbb{F}_q$-roots of $f(x)$. In [7] we use generating functions to compute the expected number of roots $E_d(R)$ of a degree d polynomial to be

$$E_d(R) = \frac{1 - \frac{1}{q^d}}{1 - \frac{1}{q}} = 1 + \frac{1}{q} + \frac{1}{q^2} + \frac{1}{q^3} + \ldots + \frac{1}{q^{d-1}}. \tag{4.2}$$

Viewed as a class function of S_d, $R(\sigma)$ is the number of fixed points of σ. Hence R is the character of the permutation representation $\mathbb{Q}[d]$. It is well known that the irreducible decomposition of $\mathbb{Q}[d]$ is

$$\mathbb{Q}[d] \cong \mathbf{1} \oplus \mathbf{Std},$$

where $\mathbf{Std} := \mathcal{S}_{[d-1,1]}$ is the *standard representation* of S_d of dimension $d - 1$. We already determined that $H^0(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q}) \cong \mathbf{1}$, explaining the constant term in (4.2). Thus Theorem 3.1 implies that each $H^{2k}(\mathrm{PConf}_d(\mathbb{R}^3), \mathbb{Q})$ has a single $\mathbf{Std}$ component for $1 \leq k \leq d - 1$, accounting for all copies of $\mathbf{Std}$. For comparison we note that

$$E_d^{\mathrm{sf}}(R) = \frac{1 - \frac{(-1)^{d-1}}{q^{d-1}}}{1 + \frac{1}{q}} = 1 - \frac{1}{q} + \frac{1}{q^2} - \frac{1}{q^3} + \ldots + (-1)^{d-1}\frac{1}{q^{d-2}}.$$

4.2.4 Evaluating at $q = 1$

Recall that the inner product $\langle \chi, \psi \rangle$ of symmetric group class functions is bilinear. If P is any factorization statistic, then by Theorem 3.1 we have the following evaluation of $E_d(P)$ at $q = 1$,

$$E_d(P)_{q=1} = \sum_{k=0}^{d-1} \langle P, \psi_d^k \rangle = \langle P, \sum_{k=0}^{d-1} \psi_d^k \rangle.$$

Passing to characters in Theorem 4.1 gives

$$\sum_{k=0}^{d-1} \psi_d^k = \chi_{\mathrm{reg}},$$

where χ_{reg} is the character of the regular representation $\mathbb{Q}[S_d]$. If P is a character of an S_d-representation V, then it follows from the general representation theory of finite groups that $\langle P, \chi_{\mathrm{reg}} \rangle = \dim V$. Therefore,

$$E_d(P)_{q=1} = \dim V.$$

If Q is the quadratic excess factorization statistic, then earlier we showed that Q is the character of the $\binom{d}{2}$-dimensional representation $\wedge^2 \mathbb{Q}[d]$. Hence

$$E_d(Q)_{q=1} = \binom{d}{2},$$

which was observed in the table of values for $E_d(Q)$. We also showed that the root statistic R was the character of the permutation representation $\mathbb{Q}[d]$, hence

$$E_d(R)_{q=1} = d.$$

Acknowledgements

References

[1]   Weiyan Chen. "Analytic number theory for 0-cycles". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Cambridge University Press. 2017, pp. 1–24.

[2]   Thomas Church. "Homological stability for configuration spaces of manifolds". In: *Inventiones mathematicae* 188.2 (2012), pp. 465–504.

[3]   Thomas Church, Jordan Ellenberg, and Benson Farb. "Representation stability in cohomology and asymptotics for families of varieties over finite fields". In: *Contemporary Mathematics* 620 (2014), pp. 1–54.

[4]   Thomas Church and Benson Farb. "Representation theory and homological stability". In: *Advances in Mathematics* 245 (2013), pp. 250–314.

[5]   Patricia Hersh and Victor Reiner. "Representation Stability for Cohomology of Configuration Spaces in". In: *International Mathematics Research Notices* 2017.5 (2016), pp. 1433–1486.

[6]   Trevor Hyde. "Liminal reciprocity and factorization statistics". In: *ArXiv e-prints* (Mar. 2018). arXiv: 1803.08438 [math.NT].

[7]   Trevor Hyde. "Polynomial factorization statistics and point configurations in $\mathbb{R}^3$". In: *ArXiv e-prints* (Jan. 2018). arXiv: 1802.00305 [math.RT].