THE UNIVERSITY OF CHICAGO

A PARAMETER SPACE FOR TSCHIRNHAUS TRANSFORMATIONS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

BY

NATHANIEL MAYER

CHICAGO, ILLINOIS

AUGUST 2022

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am very grateful to my advisor, Benson Farb, for his patience and support. He impressed upon me that there is beauty in the concrete as well as the abstract, and showed me the power of a fully worked-out example to reveal deeper patterns. He also reminded me, when I needed to hear it, that there is more to life than math.

Thank you also to the Farb research group, members present and past, for the many fascinating conversations. A particularly big thank you to Ronno Das; any time I had a question he was there with sharp insight and genuine interest, ready to drop whatever he was doing to work through it with me.

I could not have made it to this point without many excellent friends. Thank you Jon DeWitt, Liam Mazurowski, and Michael Neaton, for getting me through a stressful first year of classes and a more stressful second year of beerskits. To Kailee Zingler, for being the most generous person I know. To the Crossword Crew, Tim Black, Alan Chang, Ronno Das again, Dylan Quintana, Mariya Sardarli, Karl Schaeffer, Isabella Scott, Eric Stubley, and Noah Taylor, for sharing the joy of solving so many good puzzles. To Margaret Nichols, Nick Salter, and Raluca Havarneanu, for good advice over good beer. To Ben O'Connor and the Eckhart Sabres, for sloppy softball fun.

Thank you to my family, my parents Jim Mayer and Faith Katkin, my brother Burt and sister Elena, for your endless love and support. Most of all, thank you to my wife, Chloe Li, for being by my side through every step, through all the highs and lows, even from a thousand miles away.

ABSTRACT

In this thesis we explore the topology of a space $\mathcal{T}_n$ parameterizing Tschirnhaus transformations. We begin with an exposition of Tschirnhaus transformations in modern language, and a demonstration of their use in solving the depressed cubic equation. We then construct $\mathcal{T}_n$ and show that it has the structure of a $\mathrm{PConf}_n(\mathbb{C})$ bundle over $\mathrm{UConf}_n(\mathbb{C})$. We see that rational maps from a variety Y into $\mathcal{T}_n$ correspond to primitive elements in degree n field extensions of the function field $K(Y)$. We give a formula for the rational cohomology of $\mathcal{T}_n$, see that it is homologically stable as $n \to \infty$, and explicitly compute the dimensions of the stable cohomology groups in low degree, along the way computing character polynomials for $H^i(P_n; \mathbb{Q})$ as an S_n representation for $i \leq 5$. Finally, we give a brief historical overview of progress in solving low degree polynomials, leading into the active research fields of essential dimension, resolvent degree and the algebraic form of Hilbert's 13th problem, topics to which we hope this parameter space and our results on it might apply and prove useful.

1 INTRODUCTION

Introduced by Ehrenfried Walter von Tschirnhaus in 1683 [23], the Tschirnhaus transformation is a powerful tool for reducing the number of parameters in polynomial expressions. We begin by illustrating the technique, as Tschirnhaus himself did, to solve the depressed cubic

$$p(x) = x^3 + a_2 x + a_3 = 0.$$

The idea is to perform a sort of indirect substitution, replacing x with a new variable

$$y = s(x) = x^2 + b_2 x + b_3.$$

(In general $s(x)$ may also have a leading coefficient b_1, but it is not necessary for this computation.) Then y also satisfies a monic cubic polynomial,

$$q(y) = y^3 + c_1 y^2 + c_2 y + c_3 = 0.$$

Any of the resulting polynomial $q(y)$, the substitution polynomial $s(x)$, or the entire process may be described as a *Tschirnhaus transformation* of $p(x)$. The goal is to choose b_2 and b_3 cleverly so that $q(y) = 0$ may be solved for y by simpler means; here we will try to eliminate the two parameters $c_1 = c_2 = 0$.

The c_i can be computed in terms of the a_i and b_i in either of two ways:

1. Using Vieta's formulas: Let x_1, x_2, x_3 be the roots of $p(x)$, and $y_i = s(x_i)$ be the individually transformed roots. Then

   $$a_i = (-1)^i \sigma_i(x_1, x_2, x_3) \quad \text{and} \quad c_i = (-1)^i \sigma_i(y_1, y_2, y_3) = (-1)^i \sigma_i(s(x_1), s(x_2), s(x_3))$$

   where σ_i are the elementary symmetric polynomials. We can then expand the c_i as polynomials in the b_i and x_i, which are symmetric in the x_i, so by the fundamental theorem of symmetric polynomials we can write them as polynomials the a_i and b_i (readily done by hand or with a computer algebra system).

2. Using linear algebra: Making liberal use of the relation $p(x) = 0$, rewrite $y^2 = s(x)^2$ and $y^3 = s(x)^3$ as quadratic polynomials in x, with coefficients in terms of the a_i and b_i. Combining those with the definitional $y = s(x)$ and trivial $1 = 1$, the four equations constitute a linear map from the space of cubic polynomials in y to the space of quadratics in x. Computing the kernel gives a relation between 1, y, y^2, and y^3, precisely the desired $q(y) = 0$.

By either method, we arrive at

$$\begin{aligned} q(y) = y^3 &+ (2a_2 - 3b_3)y^2 + (a_2^2 + 3a_3b_2 - 4a_2b_3 + a_2b_2^2 + 3b_3^2)y \\ &+ (-a_3^2 + a_2a_3b_2 - a_2^2b_3 - 3a_3b_2b_3 + 2a_2b_3^2 + a_3b_2^3 - a_2b_2^2b_3 - b_3^3) = 0. \end{aligned}$$

Next we set the first two coefficients equal to 0 and solve for b_2 and b_3 in terms of the original coefficients a_2, a_3:

$$b_2 = \frac{-3a_3 - 6\sqrt{\frac{a_3^2}{4} + \frac{a_2^3}{27}}}{2a_2}, \qquad b_3 = \frac{2}{3}a_2$$

Then we plug those expressions back into $q(y)$ and solve for y solely in terms of a_2 and a_3:

$$y = \sqrt[3]{-c_3} = \frac{6}{a_2}\sqrt{\frac{a_3^2}{4} + \frac{a_2^3}{27}}\sqrt[3]{-\frac{a_3}{2} + \sqrt{\frac{a_3^2}{4} + \frac{a_2^3}{27}}}$$

Finally, we reverse the substitution, solving for x in terms of y. In this low degree case, the quadratic $y = s(x)$ can easily be solved for x directly (though one of the two solutions is parasitic), but in general we can use linear algebra to avoid the issue: Following method (2) above (but leaving out y^3), we have a linear isomorphism from the space of quadratic polynomials in y to the quadratics in x, which we can invert to get x in terms of y. After simplification, we arrive at

$$\begin{aligned} x &= -\frac{\frac{a_3}{2} + \sqrt{\frac{a_3^2}{4} + \frac{a_2^3}{27}}}{4\left(\frac{a_3^2}{4} + \frac{a_2^3}{27}\right)} y^2 + \frac{a_2}{6\sqrt{\frac{a_3^2}{4} + \frac{a_2^3}{27}}} y \\ &= \sqrt[3]{-\frac{a_3}{2} + \sqrt{\frac{a_3^2}{4} + \frac{a_2^3}{27}}} - \frac{a_2}{3\sqrt[3]{-\frac{a_3}{2} + \sqrt{\frac{a_3^2}{4} + \frac{a_2^3}{27}}}} \end{aligned}$$

(with the two cube roots equal to each other), which one may recognize as Cardano's formula.

In the centuries following his original paper, Tschirnhaus transformations have been used extensively as both a computational and theoretical tool for solving polynomials. In particular, they are central to Hilbert's 13th problem, which asks if a general degree 7 polynomial can be solved by means of algebraic functions of only two variables. We give a restatement of the problem in modern language and a brief summary of the relevant history in section 4.

A variant of Hilbert's 13th problem using continuous functions, rather than algebraic ones, was solved in 1957 by Arnold [3] [4] based on work of Kolmogorov [19]. More than a decade later, Arnold continued working on the algebraic version. In a pair of papers ([6], [7]), he explored the topology of algebraic functions, meaning the roots of polynomials over function fields, thought of as a function of the coefficients. Using the equivalence between the category of finitely generated field extensions of a base field k and the category of varieties over k and dominant rational maps, every algebraic function corresponds to a branched cover of varieties, up to rational equivalence. The *universal degree n algebraic function*

$$p_n(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n$$

corresponds to the cover:

$$\begin{tikzcd} \widetilde{\mathbb{A}^n}_{p_n} = \{(p(x), r) \in \mathbb{A}^n_{poly} \times \mathbb{A}^1 \mid p(r) = 0\} \arrow[d, "\pi_{p_n}"] \\ \mathbb{A}^n \arrow[r, equal] & \{p(x) \in k[x] \text{ monic, degree } n\} \end{tikzcd}$$

It is universal in the sense that every degree n cover is pulled back from this one via a canonical classifying map; given a variety Y with function field $K = K(Y)$ and a degree n separable algebraic function $p(x) \in K[x]$, evaluating the coefficient functions at a point $y \in Y$ (after removing their indeterminacy loci) gives a pullback square:

$$\begin{tikzcd} \widetilde{Y}_p \arrow[r] \arrow[d, "\pi_p"'] \arrow[dr, phantom, "\lrcorner", very near start] & \widetilde{\mathbb{A}^n}_{p_n} \arrow[d, "\pi_{p_n}"] \\ Y \arrow[r, "f_p"'] & \mathbb{A}^n \\ \Cup & \Cup \\ y \arrow[r, mapsto] & p(y, x) \end{tikzcd}$$

Removing the branch locus from $\mathbb{A}^n$ leaves $\mathrm{Poly}_n(k)$, the space of square-free monic degree n polynomials over k. When $k = \mathbb{C}$, identifying a polynomial with its set of roots gives an isomorphism of $\mathrm{Poly}_n(\mathbb{C})$ to $\mathrm{UConf}_n(\mathbb{C})$, the space of unordered configurations of n distinct points in the plane, which is a $K(\pi, 1)$ space for the braid group B_n. The space of ordered configurations of n points in the plane, $\mathrm{PConf}_n(\mathbb{C})$, is a $K(\pi, 1)$ space for the pure braid group P_n. The natural map $\mathrm{PConf}_n(\mathbb{C}) \to \mathrm{UConf}_n(\mathbb{C})$ forgetting the order on the configuration can be identified with the Galois closure of the cover π_{p_n}.

By universality, the cohomology of the braid group defines characteristic classes for algebraic functions, pulled back from $\mathrm{Poly}_n(\mathbb{C})$ along the classifying maps. Arnold used these characteristic classes to obstruct the reduction of parameters in certain polynomials, though under conditions that are somewhat artificial to Hilbert's problem. In [7] he uses regular functions rather than rational functions throughout, and he finds an obstruction to solving $p_n(x) = 0$ only if you disallow parasitic solutions. His theorem in that paper would suggest that the quartic $p_4(x)$ is unsolvable, since the solution in radicals includes parasitic solutions. Still, the idea of searching for obstructions in the cohomology of relevant universal spaces is a promising one.

Our inspiration for this paper comes from another paper of Arnold, [5], in which he defines the Tschirnhaus transformation geometrically as a map

$$T : \mathbb{A}^n_a \times \mathbb{A}^n_b \to \mathbb{A}^n_c,$$

with the three affine spaces parameterizing the polynomials $p(x)$, $s(x)$, and $q(y)$ described in the example above. Just as the topologically interesting universal space $\mathrm{Poly}_n(k)$ is obtained by cutting out the discriminant locus Σ from $\mathbb{A}^n$, we define

$$\mathcal{T}_n = \mathbb{A}^n_a \times \mathbb{A}^n_b \setminus T^{-1}(\Sigma),$$

a universal parameter space for Tschirnhaus transformations. Our results describe $\mathcal{T}_n$ topologically, determine a universal property, and compute its cohomology.

Theorem 2.1. *For $k = \mathbb{C}$, $\mathcal{T}_n$ has the structure of a fiber bundle:*

$$\begin{array}{ccc} \mathrm{PConf}_n(\mathbb{C}) & \longrightarrow & \mathcal{T}_n \\ & & \downarrow \\ & & \mathrm{UConf}_n(\mathbb{C}) \end{array}$$

The monodromy action of $\pi_1(\mathrm{UConf}_n(\mathbb{C})) = B_n$ factors through S_n, acting on the fiber $\mathrm{PConf}_n(\mathbb{C})$ by permuting the points.

Theorem 2.2. *Let Y be a variety, $K = K(Y)$ its function field, $p(x) \in K[x]$ a degree n separable polynomial with classifying map $f_p : Y \dashrightarrow \mathrm{Poly}_n$, defining the field extension*

$L = K[x]/p(x)$. *Then rational maps* $\sigma : Y \dashrightarrow \mathcal{T}_n$ *such that the triangle*

$$\begin{tikzcd} & \mathcal{T}_n \arrow[d, "\pi"] \\ Y \arrow[ur, dashed, "\sigma"] \arrow[r, dashed, "f_p"'] & \mathrm{Poly}_n \end{tikzcd}$$

commutes are in bijection with primitive elements of L/K.

Theorem 3.3. *Over* $k = \mathbb{C}$,

$$H^\ell(\mathcal{T}_n; \mathbb{Q}) \cong \bigoplus_{i=0}^{\ell} H^i(P_n; \mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^{\ell-i}(P_n; \mathbb{Q}).$$

The summands $H^i(P_n; \mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n; \mathbb{Q})$ *stabilize for sufficiently large* n. *Computed values for* $i, j \leq 5$ *are shown in the following table:*

j							
5	0	$\mathbb{Q}^{11}$	$\mathbb{Q}^{156}$	$\mathbb{Q}^{1345}$	$\mathbb{Q}^{8119}$	$\mathbb{Q}^{37324}$	
4	0	$\mathbb{Q}^{8}$	$\mathbb{Q}^{87}$	$\mathbb{Q}^{541}$	$\mathbb{Q}^{2363}$	$\mathbb{Q}^{8119}$	
3	0	$\mathbb{Q}^{5}$	$\mathbb{Q}^{44}$	$\mathbb{Q}^{184}$	$\mathbb{Q}^{541}$	$\mathbb{Q}^{1345}$	
2	0	$\mathbb{Q}^{4}$	$\mathbb{Q}^{18}$	$\mathbb{Q}^{44}$	$\mathbb{Q}^{87}$	$\mathbb{Q}^{156}$	
1	$\mathbb{Q}$	$\mathbb{Q}^{3}$	$\mathbb{Q}^{4}$	$\mathbb{Q}^{5}$	$\mathbb{Q}^{8}$	$\mathbb{Q}^{11}$	
0	$\mathbb{Q}$	$\mathbb{Q}$	0	0	0	0	
	0	1	2	3	4	5	i

2 CONSTRUCTION

From here on, fix the base field $k = \mathbb{C}$. As in the introduction, consider three general polynomials,

$$
\begin{aligned}
p(x) &= x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n, \\
s(x) &= b_1 x^{n-1} + \cdots + b_{n-1}x + b_n, \\
q(y) &= y^n + c_1 y^{n-1} + \cdots + c_{n-1}y + c_n,
\end{aligned}
$$

with corresponding affine spaces $\mathbb{A}^n_a$, $\mathbb{A}^n_b$, and $\mathbb{A}^n_c$ parameterizing values for the coefficients in $\mathbb{C}$. Let $x_1, ..., x_n$ and $y_1, ..., y_n$ denote the roots of $p(x)$ and $q(y)$ respectively, parameterized by yet more affine spaces $\mathbb{A}^n_x$ and $\mathbb{A}^n_y$, with canonical maps down to $\mathbb{A}^n_a$ and $\mathbb{A}^n_c$ respectively sending a set of roots to the monic polynomial with those roots.

The *Tschirnhaus transformation* is a map $T : \mathbb{A}^n_a \times \mathbb{A}^n_b \to \mathbb{A}^n_c$ fitting into the following commutative square:

$$
\begin{tikzcd}
((x_1, ..., x_n), s(x)) \arrow[rrr, mapsto] \arrow[ddd, mapsto] \arrow[dr, phantom, "\in"] & & & (s(x_1), ..., s(x_n)) \arrow[ddd, mapsto] \\
& \mathbb{A}^n_x \times \mathbb{A}^n_b \arrow[r, "\widetilde{T}"] \arrow[d] & \mathbb{A}^n_y \arrow[d] \arrow[ur, phantom, "\ni"] & \\
& \mathbb{A}^n_a \times \mathbb{A}^n_b \arrow[r, "T"'] & \mathbb{A}^n_c, \arrow[dr, phantom, "\ni"] & \\
(p(x), s(x)) \arrow[rrr, mapsto] \arrow[ur, phantom, "\in"] & & & q(y) = \prod_{i=1}^n (y - s(x_i))
\end{tikzcd}
$$

The coefficients of $q(y)$ are polynomials in the b_i and x_i, symmetric in the x_i, so can be rewritten as polynomials in the b_i and the elementary symmetric functions

$$\sigma_i(x_1, ..., x_n) = (-1)^i a_i,$$

which are the coefficients of $p(x)$. Thus T is a well-defined regular map.

Our parameter space consists of those pairs $(p(x), s(x))$ such that the resulting $q(y)$ has n distinct roots. That is,

$$\mathcal{T}_n \coloneqq T^{-1}(\mathrm{Poly}_n(\mathbb{C})).$$

Theorem 2.1. *$\mathcal{T}_n$ has the structure of a fiber bundle:*

$$\begin{tikzcd} \mathrm{PConf}_n(\mathbb{C}) \arrow[r] & \mathcal{T}_n \arrow[d, "\pi"] \\ & \mathrm{UConf}_n(\mathbb{C}) \end{tikzcd}$$

The monodromy action of $\pi_1(\mathrm{UConf}_n(\mathbb{C})) = B_n$ factors through S_n, acting on the fiber $\mathrm{PConf}_n(\mathbb{C})$ by permuting the points.

Proof. Using the notation from above, the roots $s(x_i)$ of $q(y) = T(p(x), s(x))$ can be distinct only if the original roots x_i of $p(x)$ were distinct. Thus the projection $\pi : \mathcal{T}_n \to \mathbb{A}^n_a$ onto the first factor has image contained in $\mathrm{Poly}_n(\mathbb{C}) \cong \mathrm{UConf}_n(\mathbb{C})$. Given any $p(x) \in \mathrm{Poly}_n(\mathbb{C})$, the pair $(p(x), x)$ is in $\mathcal{T}_n$, since $s(x) = x$ is the identity Tschirnhaus transformation,

$$T(p(x), x) = p(x).$$

Thus π in fact surjects onto $\mathrm{Poly}_n(\mathbb{C})$.

To determine the fiber, fix a polynomial $p(x) \in \mathrm{Poly}_n(\mathbb{C}) \subset \mathbb{A}^n_a$. Choose an arbitrary ordering of its roots $\vec{x}_p = (x_1, ..., x_n)$, and consider the restriction $\widetilde{T}_{\vec{x}_p}$ of $\widetilde{T}$ to $\{\vec{x}_p\} \times \mathbb{A}^n_b$:

$$\begin{aligned} &\widetilde{T}_{\vec{x}_p} : \mathbb{A}^n_b \to \mathbb{A}^n_y \\ &\widetilde{T}(s(x)) = (s(x_1), ..., s(x_n)). \end{aligned}$$

The transformation $T(p(x), s(x))$ lies in $\mathrm{Poly}_n(\mathbb{C})$ precisely when the transformed roots $s(x_i)$ are distinct, that is when $s(x) \in \widetilde{T}^{-1}_{\vec{x}_p}(\mathrm{PConf}_n(\mathbb{C}))$.

$\widetilde{T}_{\vec{x}_p}$ is a linear map in the coefficients b_i of $s(x)$. Its matrix is the Vandermonde matrix

$$\begin{pmatrix} x_1^{n-1} & \cdots & 1 \\ \vdots & \ddots & \vdots \\ x_n^{n-1} & \cdots & 1 \end{pmatrix},$$

whose determinant is well known to be a square root of the discriminant of $p(x)$, which is nonzero since $p(x) \in \mathrm{Poly}_n(\mathbb{C})$. Therefore $\widetilde{T}_{\vec{x}_p}$ is an isomorphism, and we conclude that

$$\pi^{-1}(p(x)) = (\{p(x)\} \times \mathbb{A}_b^n) \cap T^{-1}(\mathrm{Poly}_n(\mathbb{C})) \cong \widetilde{T}_{\vec{x}_p}^{-1}(\mathrm{PConf}_n(\mathbb{C})) \cong \mathrm{PConf}_n(\mathbb{C}).$$

That $\mathcal{T}_n$ is in fact a fiber bundle follows immediately from the fact that

$$\mathrm{PConf}_n(\mathbb{C}) \to \mathrm{UConf}_n(\mathbb{C}) \cong \mathrm{Poly}_n(\mathbb{C})$$

is a covering map. On an evenly covered neighborhood $U \subset \mathrm{UConf}_n(\mathbb{C})$ we can choose a locally consistent ordering of the roots x_i, as well defined continuous functions on U. This allows a continuous local assignment $p \mapsto \widetilde{T}_{\vec{x}_p} \in GL_n(\mathbb{C})$, and a local trivialization of the bundle

$$\pi^{-1}(U) \to U \times \mathrm{PConf}_n(\mathbb{C})$$
$$(p(x), s(x)) \mapsto (p(x), \widetilde{T}_{\vec{x}_p}(s(x))).$$

The monodromy action of B_n in the usual cover $\mathrm{PConf}_n(\mathbb{C}) \to \mathrm{Poly}_n(\mathbb{C})$ factors through the natural map $B_n \to S_n$, permuting the roots x_i of $p(x)$. This subsequently permutes the rows of the matrix for $\widetilde{T}_{\vec{x}_p}$, and thus the coordinates of the fiber $\pi^{-1}(p(x))$ which $\widetilde{T}_{\vec{x}_p}$ identifies with $\mathrm{PConf}_n(\mathbb{C})$. $\square$

Theorem 2.2. *Let Y be a variety, $K = K(Y)$ its function field, $p(x) \in K[x]$ a degree n separable polynomial with classifying map $f_p : Y \dashrightarrow \mathbb{A}^n_a$, defining the field extension $L = K[x]/p(x)$. Then rational maps $\sigma : Y \dashrightarrow \mathcal{T}_n$ such that the triangle*

$$\begin{tikzcd} & \mathcal{T}_n \arrow[d, "\pi"] \\ Y \arrow[ur, dashed, "\sigma"] \arrow[r, dashed, "f_p"'] & \mathrm{Poly}_n(\mathbb{C}) \end{tikzcd}$$

commutes are in bijection with primitive elements of L/K.

Proof. Recall from the construction that, as a bundle, $\mathcal{T}_n$ sits inside of the trivial bundle $\mathrm{Poly}_n(\mathbb{C}) \times \mathbb{A}^n_b$, where the second factor parameterizes polynomials $s(x)$ of degree at most $n-1$. Abusing notation slightly, the data of a rational map $s : Y \dashrightarrow \mathbb{A}^n_b$ is precisely a polynomial $s(x)$ of degree at most $n-1$ with coefficients in K. Let α be the residue of x in $L = K[x]/p(x)$. We then use $s(x)$ as a Tschirnhaus transformation, producing $s(\alpha) \in L$. This gives a bijection between such maps s and elements of L, since the powers of the primitive element α form a basis for L as a vector space over K. The content of the theorem is then that $s(\alpha)$ is a primitive element for L/K if and only if the image of the section $\sigma = (f_p, s)$ (restricted to an appropriate Zariski open set in Y) lands in $\mathcal{T}_n$.

Let

$$\widetilde{Y} = \{(y, r) \in Y \times \mathbb{C} \mid p(y, r) = 0\}$$

be the degree n branched cover of Y with $K(\widetilde{Y}) = L$. Up to Galois conjugation, α is the projection onto the second factor

$$\widetilde{Y} \hookrightarrow Y \times \mathbb{C} \xrightarrow{\pi_2} \mathbb{C}.$$

The fact that α is primitive for L/K is equivalent to $\widetilde{Y}$ being an n-sheeted, connected branched cover of Y.

The polynomial $s(x)$ defines a different map on $\widetilde{Y}$,

$$\begin{aligned} \widetilde{Y} &\xrightarrow{\hat{s}} Y \times \mathbb{C} \\ (y, r) &\longmapsto (y, s(y, r)) \end{aligned}$$

so that

$$s(\alpha) = \pi_2 \circ \hat{s}.$$

The transformed $s(\alpha)$ is primitive for L/K if and only if the branched cover $\hat{s}(\widetilde{Y}) \to Y$ is still degree n. Over a suitable Zariski open $U \subset Y$, both $\widetilde{Y}$ and $\hat{s}(\widetilde{Y})$ are true covering spaces, and $\hat{s}$ is a map of covers. $\hat{s}(\widetilde{Y})$ is degree n if and only if the fiberwise functions $s(y, x)$ send the n roots of $p(y, x)$ to distinct values, i.e. if $\sigma(y) = (p(y, x), s(y, x)) \in \mathcal{T}_n$. $\square$

3 COHOMOLOGY COMPUTATION

To compute the cohomology of $\mathcal{T}_n$, we must first briefly review the cohomology of $\mathrm{PConf}_n(\mathbb{C})$ (equivalently the cohomology of the pure braid group P_n, since $\mathrm{PConf}_n(\mathbb{C})$ is a $K(P_n, 1)$ space). The basic case is $n = 2$, where we have an algebraic deformation retract

$$\begin{aligned} h_t : \mathrm{PConf}_2(\mathbb{C}) &\to \mathrm{PConf}_2(\mathbb{C}) \\ (z_1, z_2) &\mapsto ((1-t)z_1, z_2 - tz_1) \end{aligned}$$

onto the subspace $\{0\} \times \mathbb{C}^*$ by simply translating the first point to the origin. Thus $H^*(P_2) \cong H^*(\mathbb{C}^*)$, with one generator $\omega \in H^1(P_2; \mathbb{Z})$ which counts the winding number of the second strand in a braid around the first.

Lemma 3.1. *The mixed Hodge structure (see [12]) on $H^1(P_2; \mathbb{Z}) \cong H^1(\mathbb{C}^*; \mathbb{Z})$ is pure of weight 2.*

Proof. Punctured curves are well-known, standard examples of mixed Hodge structures (see e.g. [13]), which can be computed using the Gysin sequence. Explicitly, the long exact sequence of the pair $(\mathbb{P}^1, \mathbb{C}^*)$ contains the segment

$$0 \longrightarrow H^1(\mathbb{C}^*; \mathbb{Q}) \longrightarrow H^2(\mathbb{P}^1, \mathbb{C}^*; \mathbb{Q}).$$

Let $\nu \cong \mathbb{C} \times \{0, \infty\}$ denote a tubular neighborhood of $\{0, \infty\}$ in $\mathbb{P}^1$. By excision and the Thom isomorphism,

$$H^2(\mathbb{P}^1, \mathbb{C}^*; \mathbb{Q}) \cong H^2(\nu, \nu \setminus \{0, \infty\}; \mathbb{Q}) \cong H^0(\{0, \infty\}; \mathbb{Q}) \cong \mathbb{Q}^2.$$

The Thom isomorphism is given by the cup product with the Thom class
$u \in H^2(\nu, \nu \setminus \{0, \infty\}; \mathbb{Q})$. By definition, the restriction of u to either fiber gives a generator of $H^2(\mathbb{C}, \mathbb{C}^*; \mathbb{Q}) \cong H^2(\mathbb{P}^1; \mathbb{Q})$ which has a pure Hodge structure of weight 2, thus u has weight 2. In the category of mixed Hodge structures, then, we have

$$H^1(\mathbb{C}^*; \mathbb{Q}) \hookrightarrow H^2(\mathbb{P}^1, \mathbb{C}^*; \mathbb{Q}) \quad \cong \quad \mathbb{Q}^2(-1),$$

where the (-1) denotes a Tate twist raising the weight by 2. We conclude that $H^1(\mathbb{C}^*; \mathbb{Q})$ is pure of weight 2. □

Lemma 3.2. *The canonical mixed Hodge structure on $H^k(P_n; \mathbb{Z})$ is pure of weight $2k$.*

Proof. There are forgetful maps $u_{ij} : \mathrm{PConf}_n(\mathbb{C}) \to \mathrm{PConf}_2(\mathbb{C})$ for each pair of indices $1 \leq i, j \leq n$, picking out only the i-th and j-th points of the configuration. Let

$$\omega_{ij} = u_{ij}^*(\omega) \in H^1(P_n; \mathbb{Z}),$$

representing the winding number of the j-th strand around the i-th strand. By Lemma 3.1, each ω_{ij} has weight 2.

Arnold showed in [2] that the classes ω_{ij} in fact generate $H^*(P_n;\mathbb{Z})$ as a ring, subject only to the relations

$$\omega_{ji} = \omega_{ij}$$

$$\omega_{ij}\omega_{\ell m} = -\omega_{\ell m}\omega_{ij}$$

$$\omega_{ij}\omega_{j\ell} + \omega_{j\ell}\omega_{\ell i} + \omega_{\ell i}\omega_{ij} = 0$$

for all choices of indices i, j, ℓ, and m. All classes in $H^k(P_n;\mathbb{Z})$ are then degree k polynomials in the ω_{ij}, therefore have weight $2k$. $\square$

Now we can proceed with the main computation.

Theorem 3.3. *Over $k = \mathbb{C}$,*

$$H^\ell(\mathcal{T}_n;\mathbb{Q}) \cong \bigoplus_{i=0}^{\ell} H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^{\ell-i}(P_n;\mathbb{Q}).$$

The summands $H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q})$ stabilize for sufficiently large n. Computed values for $i, j \leq 5$ are shown in the following table:

j							
5	0	$\mathbb{Q}^{11}$	$\mathbb{Q}^{156}$	$\mathbb{Q}^{1345}$	$\mathbb{Q}^{8119}$	$\mathbb{Q}^{37324}$	
4	0	$\mathbb{Q}^{8}$	$\mathbb{Q}^{87}$	$\mathbb{Q}^{541}$	$\mathbb{Q}^{2363}$	$\mathbb{Q}^{8119}$	
3	0	$\mathbb{Q}^{5}$	$\mathbb{Q}^{44}$	$\mathbb{Q}^{184}$	$\mathbb{Q}^{541}$	$\mathbb{Q}^{1345}$	
2	0	$\mathbb{Q}^{4}$	$\mathbb{Q}^{18}$	$\mathbb{Q}^{44}$	$\mathbb{Q}^{87}$	$\mathbb{Q}^{156}$	
1	$\mathbb{Q}$	$\mathbb{Q}^{3}$	$\mathbb{Q}^{4}$	$\mathbb{Q}^{5}$	$\mathbb{Q}^{8}$	$\mathbb{Q}^{11}$	
0	$\mathbb{Q}$	$\mathbb{Q}$	0	0	0	0	
	0	1	2	3	4	5	i

Proof. We apply the Serre spectral sequence to the fiber bundle

$$\mathrm{PConf}_n(\mathbb{C}) \to \mathcal{T}_n \to \mathrm{UConf}_n(\mathbb{C}),$$

giving E^2 page

$$\begin{aligned} E^2_{i,j} &= H^i(\mathrm{UConf}_n(\mathbb{C}); \mathcal{H}^j(\mathrm{PConf}_n(\mathbb{C}); \mathbb{Q})) \\ &\cong H^i(B_n; \mathcal{H}^j(P_n; \mathbb{Q})). \end{aligned}$$

The coefficients above are a local system, with B_n acting on $H^j(P_n;\mathbb{Q})$ by permuting the indices of the generating classes ω_{ij}, as by Theorem 2.1 the monodromy of the bundle acts on $\mathrm{PConf}_n(\mathbb{C})$ by permuting the points of the configuration. The pullback of this local system along the covering map $\mathrm{PConf}_n(\mathbb{C}) \to \mathrm{UConf}_n(\mathbb{C})$ is the constant sheaf on $\mathrm{PConf}_n(\mathbb{C})$ with stalk $H^j(\mathrm{PConf}_n(\mathbb{C});\mathbb{Q})$. By transfer,

$$\begin{aligned} E^2_{i,j} = H^i(B_n; \mathcal{H}^j(P_n; \mathbb{Q})) &\cong H^i(P_n; H^j(P_n; \mathbb{Q}))^{S_n} \\ &= (H^i(P_n; \mathbb{Q}) \otimes H^j(P_n; \mathbb{Q}))^{S_n} \\ &\cong H^i(P_n; \mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n; \mathbb{Q}) \end{aligned}$$

That last isomorphism, between invariants and coinvariants, holds for all semisimple representations.

By lemma 3.2, the Hodge structure on $E^2_{i,j}$ is then pure of weight $2(i+j)$. The differentials in the spectral sequence all go up to a higher diagonal, but they must preserve weights [1], so it follows that all the differentials vanish and $E^2_{i,j} = E^\infty_{i,j}$. The ℓ-th diagonal on the E^∞ page gives the associated graded group of $H^\ell(\mathcal{T}_n;\mathbb{Q})$, but since every term is a $\mathbb{Q}$-vector space there are no extension problems. $H^\ell(\mathcal{T}_n;\mathbb{Q})$ is simply the direct sum of all the groups on the diagonal, giving the desired result.

S_n representations are all self dual, so we may alternatively write

$$H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q}) = \mathrm{Hom}_{S_n}(H^i(P_n;\mathbb{Q}), H^j(P_n;\mathbb{Q})).$$

For each partition $\lambda \vdash n$, let $m_{i,\lambda}$ be the multiplicity of the irreducible representation V_λ in $H^i(P_n;\mathbb{Q})$. Then by Schur's lemma, the dimension of $H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q})$ is

$$\sum_{\lambda \vdash n} m_{i,\lambda} m_{j,\lambda}.$$

The cohomology groups $H^i(P_n;\mathbb{Q})$ are representation stable [11]; that is, by padding the largest segment of the partition we can treat λ as a partition of any $N \geq n$, and the multiplicities $m_{i,\lambda}$ become constant for sufficiently large N. Thus the dimensions of $H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q})$ also stabilize.

To compute those stable dimensions, we use two different descriptions of $H^*(P_n;\mathbb{Q})$ as an S_n-representation. First, Chen [10] gives a generating function for the twisted cohomology of the braid group B_n with coefficients in an arbitrary virtual S_n-representation, but a bit of background is required to state it. For $j = 1, ..., n$, define class functions $X_j : S_n \to \mathbb{Z}$, where $X_j(g)$ is the number of j-cycles in the cycle decomposition of g. All class functions of S_n can be written as polynomials in $\mathbb{Q}[X_1, ..., X_n]$, known as character polynomials. Given a tuple of nonnegative integers $\lambda = (\lambda_1, ..., \lambda_\ell)$, define the monomial

$$\binom{X}{\lambda} = \binom{X_1}{\lambda_1}\binom{X_2}{\lambda_2}\cdots\binom{X_\ell}{\lambda_\ell}.$$

Such monomials form a basis for character polynomials over $\mathbb{Q}$, and thus for all class functions. Also, let μ be the Mobius function, and let

$$M_k(z^{-1}) = \frac{1}{k}\sum_{j|k} \mu\left(\frac{k}{j}\right) z^{-j}.$$

Then from Chen [10], the twisted cohomology of P_n with coefficients in the virtual representation with character $\binom{X}{\lambda}$ is given by the generating function

$$\sum_{n=0}^{\infty}\sum_{i=0}^{\infty} \dim H^i(P_n; \binom{X}{\lambda})(-z)^i t^n = \frac{1-zt^2}{1-t}\prod_{k=1}^{\ell}\binom{M_k(z^{-1})}{\lambda_k}\left(\frac{(tz)^k}{1+(tz)^k}\right)^{\lambda_k}.$$

This generating function is perfect for computing our terms $H^i(B_n; H^j(P_n;\mathbb{Q}))$, except that we first need to know the character polynomial for $H^j(P_n;\mathbb{Q})$, which is not so easy to read off from the generating function. Instead, we use a second description of $H^*(P_n;\mathbb{Q})$, due to Lehrer and Solomon [21]:

$$H^i(P_n;\mathbb{Q}) \cong \bigoplus_{\lambda \vdash n} \mathrm{Ind}_{Z_\lambda}^{S_n}(\xi_\lambda), \tag{1}$$

where λ now ranges over partitions of n into exactly $n-i$ segments, Z_λ is the centralizer of an element of S_n with cycle type given by the partition λ, and ξ_λ is a particular character of Z_λ which we now describe.

Suppose the partition λ consists of segments

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0, \qquad \sum_{i=1}^{k}\lambda_i = n.$$

For each $1 \leq i \leq k$, let g_i be the λ_i-cycle

$$g_i = ((\lambda_1 + \cdots + \lambda_{i-1} + 1)\cdots(\lambda_1 + \cdots + \lambda_i)),$$

let

$$c_\lambda = \prod_{i=1}^{k} g_i$$

be an element with cycle type λ, and let n_j be the number of cycles in c_λ of length j. One may verify that the centralizer of c_λ is the product of wreath products

$$
\begin{aligned}
Z_\lambda &= \prod_{j=1}^{n} (\mathbb{Z}/j) \wr S_{n_j} \\
&= \prod_{j=1}^{n} (\mathbb{Z}/j)^{n_j} \rtimes S_{n_j}
\end{aligned}
$$

where the $\mathbb{Z}/j$ factors are the cyclic subgroups generated by the cycles g_i, and the S_{n_j} act by permuting cycles of equal length. The character ξ_λ is defined on the cyclic factors by

$$\xi_\lambda(g_i) = (-1)^{\lambda_i+1} e^{2\pi i/\lambda_i}$$

and on the S_{n_j} factors by $\xi_\lambda = \varepsilon^{j+1}$ where ε is the sign character. Note in particular that ξ_λ is trivial on any $\mathbb{Z}/2$ factors, but nontrivial on all longer cycles.

The induced characters are given by

$$\mathrm{Ind}(\xi_\lambda)(g) = \frac{1}{|Z_\lambda|} \sum_{h \in S_n,\, h^{-1}gh \in Z_\lambda} \xi_\lambda(h^{-1}gh).$$

With care, we can translate that into a character polynomial. For example, let $c_\lambda = (12) \in S_n$ be a single transposition, the unique conjugacy class with exactly $n-1$ cycles. Its centralizer is $Z_\lambda = \mathbb{Z}/2 \times S_{n-2}$, which contains two types of elements, $(12)(\cdots)$ and $(1)(2)(\cdots)$. A permutation g can be conjugated onto something of the form $(12)(\cdots)$ by sending a 2-cycle (ab) in g onto (12), which can be done in either order. The remaining $n-2$ elements can be permuted arbitrarily, so there are $2X_2(g)(n-2)!$ ways to do this. Similarly g can be conjugated onto $(1)(2)(\cdots)$ by choosing two 1-cycles in g and sending them to 1 and 2 in either order, with $2\binom{X_1(g)}{2}(n-2)!$ ways of doing that. Conveniently in this simple case,

$\xi_\lambda = 1$, so the character for $H^1(P_n;\mathbb{Q})$ is:

$$\begin{aligned}\chi_{H^1(P_n;\mathbb{Q})} &= \mathrm{Ind}(\xi_{(2,1,\ldots,1)}) \\ &= \frac{1}{2\cdot(n-2)!}\left(2X_2(n-2)! + 2\binom{X_1}{2}(n-2)!\right) \\ &= X_2 + \binom{X_1}{2}\end{aligned}$$

We derive further character polynomials by the same process, with some effort:

$$\begin{aligned}\chi_{H^2(P_n;\mathbb{Q})} &= \mathrm{Ind}(\xi_{(3,1,\ldots,1)}) + \mathrm{Ind}(\xi_{(2,2,1,\ldots,1)}) \\ &= 2\binom{X_1}{3} - X_3 + 3\binom{X_1}{4} + \binom{X_1}{2}X_2 - \binom{X_2}{2} - X_4\end{aligned}$$

$$\begin{aligned}\chi_{H^3(P_n;\mathbb{Q})} &= \mathrm{Ind}(\xi_{(4,1,\ldots,1)}) + \mathrm{Ind}(\xi_{(3,2,1,\ldots,1)}) + \mathrm{Ind}(\xi_{(2,2,2,1,\ldots,1)}) \\ &= 6\binom{X_1}{4} - 2\binom{X_2}{2} + 20\binom{X_1}{5} + 2\binom{X_1}{3}X_2 - \binom{X_1}{2}X_3 - X_2X_3\end{aligned}$$

$$\begin{aligned}\chi_{H^4(P_n;\mathbb{Q})} &= \mathrm{Ind}(\xi_{(5,1,\ldots,1)}) + \mathrm{Ind}(\xi_{(4,2,1,\ldots,1)}) + \mathrm{Ind}(\xi_{(3,3,1,\ldots,1)}) \\ &\quad + \mathrm{Ind}(\xi_{(3,2,2,1,\ldots,1)}) + \mathrm{Ind}(\xi_{(2,2,2,2,1,\ldots,1)}) \\ &= 24\binom{X_1}{5} - X_5 + 130\binom{X_1}{6} + 6\binom{X_1}{4}X_2 - 2\binom{X_1}{2}\binom{X_2}{2} + 2\binom{X_2}{3} - 2\binom{X_1}{3}X_3 \\ &\quad + \binom{X_3}{2} - X_6 + 210\binom{X_1}{7} + 20\binom{X_1}{5}X_2 - 2\binom{X_1}{3}\binom{X_2}{2} - 3\binom{X_1}{4}X_3 \\ &\quad - \binom{X_1}{2}X_2X_3 + \binom{X_2}{2}X_3 - 2\binom{X_1}{3}X_4 + X_3X_4 + 105\binom{X_1}{8} + 15\binom{X_1}{6}X_2 \\ &\quad - 3\binom{X_1}{4}\binom{X_2}{2} - 5\binom{X_1}{2}\binom{X_2}{3} + \binom{X_2}{4} + 3\binom{X_1}{2}\binom{X_3}{2} + 3X_2\binom{X_3}{2} \\ &\quad - 3\binom{X_1}{4}X_4 - \binom{X_1}{2}X_2X_4 + \binom{X_2}{2}X_4 - 3\binom{X_4}{2} + \binom{X_1}{2}X_6 + X_2X_6 - X_8\end{aligned}$$

$$
\begin{aligned}
\chi_{H^5(P_n;\mathbb{Q})} &= \mathrm{Ind}(\xi_{(6,1,\dots,1)}) + \mathrm{Ind}(\xi_{(5,2,1,\dots,1)}) + \mathrm{Ind}(\xi_{(4,3,1,\dots,1)}) + \mathrm{Ind}(\xi_{(4,2,2,1,\dots,1)}) \\
&\quad + \mathrm{Ind}(\xi_{(3,3,2,1,\dots,1)}) + \mathrm{Ind}(\xi_{(3,2,2,2,1,\dots,1)}) + \mathrm{Ind}(\xi_{(2,2,2,2,2,1,\dots,1)}) \\
&= 130\binom{X_1}{6}X_2 - 15\binom{X_1}{6}X_3 + 120\binom{X_1}{6} + 210\binom{X_1}{7}X_2 + 924\binom{X_1}{7} \\
&\quad + 2380\binom{X_1}{8} + 2520\binom{X_1}{9} + 24\binom{X_1}{5}X_2 + -20\binom{X_1}{5}\binom{X_2}{2} - 20\binom{X_1}{5}X_3 \\
&\quad - 20\binom{X_1}{5}X_4 - 2\binom{X_1}{3}X_2X_3 - 2\binom{X_1}{3}X_2X_4 - 3\binom{X_1}{4}X_2X_3 + X_2X_3X_4 \\
&\quad + X_2\binom{X_3}{2} - X_2X_5 - X_2X_6 - 4\binom{X_1}{3}\binom{X_2}{2} - 10\binom{X_1}{3}\binom{X_2}{3} + 6\binom{X_1}{3}\binom{X_3}{2} \\
&\quad + 2\binom{X_1}{3}X_6 - 12\binom{X_1}{4}\binom{X_2}{2} + \binom{X_1}{2}\binom{X_2}{2}X_3 + 2\binom{X_2}{2}X_3 + 2\binom{X_2}{2}X_4 \\
&\quad - 6\binom{X_1}{4}X_3 - 6\binom{X_1}{4}X_4 + 2\binom{X_1}{2}\binom{X_2}{3} + 5\binom{X_2}{3}X_3 + 8\binom{X_2}{3} + \binom{X_1}{2}X_3X_4 \\
&\quad + \binom{X_1}{2}\binom{X_3}{2} - \binom{X_1}{2}X_5 - \binom{X_1}{2}X_6 + 44\binom{X_2}{4} - X_3X_6 - 3\binom{X_3}{2} - 9\binom{X_3}{3} \\
&\quad - X_6 + 945\binom{X_1}{10} + 105\binom{X_1}{8}X_2 - 15\binom{X_1}{6}\binom{X_2}{2} - 15\binom{X_1}{4}\binom{X_2}{3} \\
&\quad + \binom{X_1}{2}\binom{X_2}{4} + 41\binom{X_2}{5} + 9\binom{X_1}{4}\binom{X_3}{2} + 3\binom{X_1}{2}X_2\binom{X_3}{2} - 3\binom{X_2}{2}\binom{X_3}{2} \\
&\quad - 15\binom{X_1}{6}X_4 - 3\binom{X_1}{4}X_2X_4 + \binom{X_1}{2}\binom{X_2}{2}X_4 + 5\binom{X_2}{3}X_4 - 3\binom{X_3}{2}X_4 \\
&\quad - 3\binom{X_1}{2}\binom{X_4}{2} - 3X_2\binom{X_4}{2} + 5\binom{X_5}{2} + 3\binom{X_1}{4}X_6 + \binom{X_1}{2}X_2X_6 - \binom{X_2}{2}X_6 \\
&\quad - X_4X_6 - \binom{X_1}{2}X_8 - X_2X_8 + X_{10}
\end{aligned}
$$

With those character polynomials in hand, Chen gives us the generating functions for $H^i(P_n; H^j(P_n;\mathbb{Q}))$. A computer algebra system can easily produce the series expansions, and we can simply read off the desired coefficients to produce the table in the theorem statement. As one may see in the alternate computational method below, these values are stable for $n > 2(i+j)$ (in fact well before that, but this is an easy upper bound). $\square$

We include here another method of computing the stable groups

$H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q})$ which is less computationally efficient, but recasts the question as an interesting combinatorial problem. From Lehrer-Solomon we have

$$\begin{aligned} H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q}) &= \mathrm{Hom}_{S_n}\left(\bigoplus_{\lambda \vdash n} \mathrm{Ind}_{Z_\lambda}^{S_n}(\xi_\lambda), \bigoplus_{\mu \vdash n} \mathrm{Ind}_{Z_\mu}^{S_n}(\xi_\mu)\right) \\ &= \bigoplus_{\lambda,\mu \vdash n} \mathrm{Hom}_{S_n}\left(\mathrm{Ind}_{Z_\lambda}^{S_n}(\xi_\lambda),\, \mathrm{Ind}_{Z_\mu}^{S_n}(\xi_\mu)\right), \end{aligned}$$

where λ and μ range over partitions of n with $n-i$ and $n-j$ segments, respectively. By Frobenius reciprocity and Theorem XVIII.7.6 in [20],

$$\begin{aligned} H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q}) &= \bigoplus_{\lambda,\mu \vdash n} \mathrm{Hom}_{S_n}\left(\xi_\lambda,\, \mathrm{Res}_{Z_\lambda}^{S_n}\mathrm{Ind}_{Z_\mu}^{S_n}(\xi_\mu)\right) \\ &= \bigoplus_{\substack{\lambda,\mu \vdash n \\ Z_\lambda\gamma Z_\mu}} \mathrm{Hom}_{S_n}\left(\xi_\lambda,\, \mathrm{Ind}_{Z_\lambda\cap\gamma Z_\mu\gamma^{-1}}^{Z_\lambda}\mathrm{Res}_{Z_\lambda\cap\gamma Z_\mu\gamma^{-1}}^{\gamma Z_\mu\gamma^{-1}}(\xi_\mu)\right) \\ &= \bigoplus_{\substack{\lambda,\mu \vdash n \\ Z_\lambda\gamma Z_\mu}} \mathrm{Hom}_{S_n}\left(\mathrm{Ind}_{Z_\lambda\cap\gamma Z_\mu\gamma^{-1}}^{Z_\lambda}\mathrm{Res}_{Z_\lambda\cap\gamma Z_\mu\gamma^{-1}}^{\gamma Z_\mu\gamma^{-1}}(\xi_\mu),\, \xi_\lambda\right) \\ &= \bigoplus_{\substack{\lambda,\mu \vdash n \\ Z_\lambda\gamma Z_\mu}} \mathrm{Hom}_{S_n}\left(\mathrm{Res}_{Z_\lambda\cap\gamma Z_\mu\gamma^{-1}}^{\gamma Z_\mu\gamma^{-1}}(\xi_\mu),\, \mathrm{Res}_{Z_\lambda\cap\gamma Z_\mu\gamma^{-1}}^{Z_\lambda}(\xi_\lambda)\right) \end{aligned}$$

where γ ranges over representatives for the double cosets $Z_\lambda\gamma Z_\mu$. Each term in the sum is the Hom space between two characters, so is either 0 or $\mathbb{Q}$. We need to count how many summands give a nonzero contribution.

Lemma 3.4. *Let $A_i \subset S_n$ be the set of elements with exactly $n-i$ cycles. The set T of ordered triples $(\lambda, \mu, Z_\lambda\gamma Z_\mu)$ as above is in bijection with the set O of orbits of the diagonal action of S_n on $A_i \times A_j$ by conjugation.*

Proof. As above, let c_λ and c_μ be chosen representatives of the conjugacy classes in S_n with cycle types λ and μ, respectively. Define a function $f : T \to O$ by

$$f(\lambda, \mu, Z_\lambda \gamma Z_\mu) = S_n \cdot (c_\lambda,\, \gamma c_\mu \gamma^{-1}).$$

To see this is well defined, let $\delta = z_\lambda^{-1} \gamma z_\mu$ be another representative of the same double coset, with $z_\lambda \in Z_\lambda$ and $z_\mu \in Z_\mu$.

$$\begin{aligned} z_\lambda \cdot (c_\lambda,\, \delta c_\mu \delta^{-1}) &= (z_\lambda c_\lambda z_\lambda^{-1},\, \gamma z_\mu c_\mu z_\mu^{-1} \gamma^{-1}) \\ &= (c_\lambda,\, \gamma c_\mu \gamma^{-1}) \end{aligned}$$

The second equality above used the fact that $z_\lambda \in Z_\lambda$ and $z_\mu \in Z_\mu$ by definition commute with c_λ and c_μ respectively. Thus both representatives γ and δ produce the same orbit in O.

In the other direction, let $(a_i, a_j) \in A_i \times A_j$ be an arbitrary pair, with cycle types λ and μ. It's orbit contains at least one element of the form (c_λ, b) where c_λ is the chosen representative of its conjugacy class, and $b = \gamma c_\mu \gamma^{-1}$ for some γ. Define a function $g : O \to T$ by $g(a_i, a_j) = (\lambda, \mu, Z_\lambda \gamma Z_\mu)$. Clearly f and g are inverses. To see that g is well defined, suppose another element in the same orbit has the desired form; for some h and $\delta \in S_n$,

$$h \cdot (c_\lambda,\, \gamma c_\mu \gamma^{-1}) = (c_\lambda,\, \delta c_\mu \delta^{-1}).$$

In the first coordinate, this tells us

$$h c_\lambda h^{-1} = c_\lambda,$$

i.e. $h \in Z_\lambda$. In the second coordinate we have

$$\begin{aligned} h \gamma c_\mu \gamma^{-1} h^{-1} &= \delta c_\mu \delta^{-1} \\ (\delta^{-1} h \gamma) c_\mu (\gamma^{-1} h^{-1} \delta) &= c_\mu, \end{aligned}$$

so $\delta^{-1}h\gamma \in Z_\mu$. Then

$$\gamma = h^{-1}\delta(\delta^{-1}h\gamma) \in Z_\lambda \delta Z_\mu,$$

so γ and δ represent the same double coset. $\square$

The lemma gives us an effective if tedious way to list out the terms in the sum; they correspond to all of the combinatorially distinct ways that two permutations c_λ and b_μ with $n-i$ and $n-j$ cycles respectively can sit relative to each other in S_n, up to conjugation. A summand is nonzero (equals $\mathbb{Q}$) if and only if the characters ξ_λ and ξ_μ agree on the intersection of centralizers $Z(c_\lambda) \cap Z(b_\mu)$. The total count will be stable for n sufficiently large that all contributing pairs c_λ, b_μ can be conjugated such that their nontrivial cycles only involve numbers up to n.

Now we turn to computations for specific i and/or j.

- $i = 0$.

  c_λ must have n cycles, so is the identity. The centralizer is $Z(c_\lambda) = S_n$, and $\xi_\lambda = 1$. Each term in the sum is nonzero if and only if ξ_μ is trivial.

  - For $j = 0$, the only term is $b_\mu = 1$, for which $\xi_\mu = 1$, so $H^0(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^0(P_n;\mathbb{Q}) = \mathbb{Q}$, for all n.
  - For $j = 1$, there is only $b_\mu = (12)$, for which $\xi_\mu = 1$, so $H^0(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^1(P_n;\mathbb{Q}) = \mathbb{Q}$, for all $n \geq 2$.
  - For $j \geq 2$, every permutation b_μ with $n-j$ cycles must either have a cycle of length at least 3 or two 2-cycles. In either case ξ_μ is nontrivial, so every summand is 0. Thus $H^0(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q}) = 0$ for $j \geq 2$.

- $i = 1$.

  λ has $n-1$ segments, which is to say c_λ is a transposition. Then $\xi_\lambda = 1$, so for a summand to contribute positively ξ_μ must be trivial on $Z(c_\lambda) \cap Z(b_\mu)$. In particular, the 2-cycle in c_λ must overlap with any cycle in b_μ of length 3 or longer, else that cycle would provide an element commuting with both c_λ and b_μ on which $\xi_\mu \neq 1$. For the same reason the 2-cycle in c_λ cannot be disjoint from two 2-cycles in b_μ. That leaves the following options:

- $b_\mu = (1 \cdots (j+1))$.

  For $j \geq 2$ we may assume without loss of generality that $c_\lambda = (1t)$. All choices of $t \geq j+2$ are equivalent up to conjugation, so we need only consider $t = j+2$. The corresponding summand is $\mathbb{Q}$, as $Z(c_\lambda) \cap Z(b_\mu) = \{1\}$.

  If $t \leq j+1$, then up to conjugation the pair (c_λ, b_μ) is determined by how far apart 1 and t are in the $(j+1)$-cycle. The order is irrelevant, as 1 and t are indistinguishable up to symmetry. The options are then

$$2 \leq t \leq 1 + \left\lfloor \frac{j+1}{2} \right\rfloor.$$

  The centralizers have trivial intersection in almost all cases, with one exception: If j is odd and $t = \frac{j+3}{2}$, then 1 and t will be precisely opposite each other in the $(j+1)$-cycle, and $(1t)$ will commute with $b_\mu^{(j+1)/2}$. Since $\xi_\lambda = 1$, this term will contribute if and only if

$$\xi_\mu(b_\mu^{(j+1)/2}) = (-e^{2\pi i/(j+1)})^{(j+1)/2} = 1,$$

  that is if $j \equiv 1 \mod 4$.

All together, b_μ of this form contribute

$$1 + \begin{cases} j/2 & \text{if } j \equiv 0 \mod 4, \\ (j+1)/2 & \text{if } j \equiv 1 \mod 4, \\ j/2 & \text{if } j \equiv 2 \mod 4, \\ (j-1)/2 & \text{if } j \equiv 3 \mod 4. \end{cases}$$

If $j = 1$, there is one additional contribution from $c_\lambda = (34)$.

– $b_\mu = (1 \cdots j)((j+1)(j+2))$ (requires $j \geq 2$).

First we treat the case $j \geq 3$. Again c_λ must overlap the j-cycle because that is where ξ_μ is nontrivial, so without loss of generality $c_\lambda = (1t)$. We get contributing terms for $t = j+1$ and $t = j+3$. The $t \leq j$ case mirrors the previous bullet point. All told this contributes

$$2 + \begin{cases} j/2 - 1 & \text{if } j \equiv 0 \mod 4, \\ (j-1)/2 & \text{if } j \equiv 1 \mod 4, \\ j/2 & \text{if } j \equiv 2 \mod 4, \\ (j-1)/2 & \text{if } j \equiv 3 \mod 4. \end{cases}$$

If $j = 2$, the above holds except that $c_\lambda = (13)$ does not contribute because the characters do not agree on $(13)(24)$.

– b_μ has only two nontrivial cycles, both of length at least 3 (requires $j \geq 4$).

Here c_λ must overlap both cycles, so without loss of generality $c_\lambda = (1j)$ The different combinations here come from the different lengths that the cycles can have, totalling $j+2$. The length of the shorter cycle ranges from 3 to $1 + \left\lfloor \frac{j}{2} \right\rfloor$.

Every such term contributes except if the two cycles have equal, even length, occurring only when $j \equiv 2 \mod 4$. Contribution:

$$\begin{cases} j/2-1 & \text{if } j \equiv 0 \mod 4, \\ (j-1)/2-1 & \text{if } j \equiv 1 \mod 4, \\ j/2-2 & \text{if } j \equiv 2 \mod 4, \\ (j-1)/2-1 & \text{if } j \equiv 3 \mod 4. \end{cases}$$

– Finally, b_μ may have three nontrivial cycles (requires $j \geq 3$), at least one of which must be of length 2, without loss of generality $((j+2)(j+3))$ (if all three were longer, c_λ couldn't overlap them all).

  As in the previous case we may take $c_\lambda = (1j)$, overlapping the two longest cycles, again the combinations come from the different lengths of the first two cycles, totalling $j+1$, and again every term contributes positively except one where the first two cycles have equal even length, occurring when $j \equiv 3 \mod 4$. Contribution:

$$\begin{cases} j/2-1 & \text{if } j \equiv 0 \mod 4, \\ (j+1)/2-1 & \text{if } j \equiv 1 \mod 4, \\ j/2-1 & \text{if } j \equiv 2 \mod 4, \\ (j+1)/2-2 & \text{if } j \equiv 3 \mod 4. \end{cases}$$

Totalling everything, for $j \geq 1$,

$$\dim H^1(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q}) = \begin{cases} 2j & \text{if } j \equiv 0 \mod 4, \\ 2j+1 & \text{if } j \equiv 1 \mod 4, \\ 2j & \text{if } j \equiv 2 \mod 4, \\ 2j-1 & \text{if } j \equiv 3 \mod 4. \end{cases}$$

Some parts of the derivation did not apply to $j = 1$, 2, or 3, but if you trace back it so happens that the extra terms all cancel out in those cases, so the formula still holds. These values are stable for $n \geq j+3$, as that is the largest number which appeared in a nontrivial cycle of c_λ or b_μ in any contributing term.

- Systematic analysis like that for $i = 1$ is more challenging for $i, j \geq 2$, but further groups may be computed by exhaustively listing all possible ways that c_λ and b_μ could overlap, and determining in each case whether ξ_λ and ξ_μ agreed on the intersection of the centralizers. As an example, here is the full list of the 18 contributing terms for $i = j = 2$:

c_λ	b_μ
(123)	(123)
	(124)
	(142)
	(145)
	(12)(34)
	(12)(45)
	(14)(25)
	(14)(56)

c_λ	b_μ
(12)(34)	(123)
	(125)
	(135)
	(156)
	(12)(34)
	(12)(35)
	(12)(56)
	(13)(25)
	(15)(36)
	(15)(67)

Problem 3.5. *Find a simplified expression for* $\dim H^i(P_n;\mathbb{Q}) \otimes_{\mathbb{Q}[S_n]} H^j(P_n;\mathbb{Q})$ *or* $\dim H^k(\mathcal{T}_n;\mathbb{Q})$*, in the stable range* $n \gg 0$*.*

An approach using generating functions may be more tractable here, but as shown above there is combinatorial subtlety in writing out the character polynomials for $H^i(P_n;\mathbb{Q})$, or, as in the alternate method, in determining whether ξ_λ and ξ_μ agree on the intersection of the centralizers.

4 HISTORICAL CONTEXT

In his paper introducing the method, Tschirnhaus asserted that it could be used "for determining analytically the roots of all equations of any degree" [23], which to him would have meant solving for the roots in radicals. Of course Galois, nearly 150 years later, showed that task to be impossible; indeed eliminating many terms from a high degree polynomial by Tschirnhaus transformation requires you to solve a system of equations for the b_i which can be as or more difficult than solving the original equation. Though Tschirnhaus himself was unable to make progress beyond Cardano's and Ferrari's solutions of the cubic and quartic respectively, others would push forward essentially using his method.

The next leap came from Bring [15] and Jerrard [18], who independently demonstrated that the general quintic polynomial, though not solvable in radicals, can by an appropriate Tschirnhaus transformation (with the b_i solvable in radicals) be reduced to the Bring radical, a root of

$$x^5 + x + a = 0,$$

an algebraic function in the single free parameter a. By the same method one may eliminate four parameters from any polynomial of degree $n \leq 5$.

This was the state of the art when Hilbert posed the 13th of his famous problems [16]. Though it was originally phrased as a question about nomography, a good statement in modern language requires a detour to introduce the notions of *essential dimension* (see, e.g., [9]) and *resolvent degree* [14]:

Fix a base field k and consider field extensions $E/F/k$, with E/F algebraic of degree n. We say that E/F is *defined over* a subfield $F_0 \subset F$ if there exists an extension E_0/F_0 of degree n inside of E, such that $E = E_0F$. Equivalently, in the case that E/F is separable, E/F is defined over F_0 if there exists a primitive element whose minimal polynomial is in $F_0[x]$. The *essential dimension* of E/F is then

$$\mathrm{ed}_k(E/F) = \min\{\mathrm{trdeg}_k(F_0) \mid E/F \text{ is defined over } F_0\}.$$

Classically, if $p(x) \in F[x]$ and $E = F[x]/p(x)$, we think of $p(x) = 0$ as expressing a root x as a multivariate function of $m = \mathrm{trdeg}_k(F)$ variables, a transcendence basis for F/k. Applying a rational substitution or Tschirnhaus transformation to $p(x)$ to reduce the number of free parameters is then "solving" for x, constructing the field extension E/F, as a function of fewer variables. $\mathrm{ed}(E/F)$ is the minimum number of variables one can reduce down to.

To illustrate, once again consider the universal degree n polynomial

$$p_n(x) = x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_n,$$

and the fields $K_n = k(a_1, ..., a_n)$ and $L_n = K_n[x]/p_n(x)$.

The quadratic equation is a demonstration that $p_2(x) = 0$, thought of as a 2-variable function $x(a_1, a_2)$, can be reduced to the 1-variable square root function $y(c_2) = \sqrt{c_2}$, the root of the simpler polynomial

$$y^2 - c_2 = 0.$$

This shows that $\mathrm{ed}_k(L_2/K_2) = 1$.

Cardano's formula is a bit more complicated; it shows that we can "solve" the 3-variable cubic $p_3(x) = 0$ using 1-variable radicals, but it requires both a cube root and a square root.

Indeed it is impossible to solve a general cubic using only a cube root, as the Galois group for the general cubic is S_3 while that of $y^3 - c_3$ is only $\mathbb{Z}/3$. Thus Cardano's formula is not a direct demonstration that $\mathrm{ed}_{\mathbb{C}}(L_3/K_3) = 1$ (though that is true; $p_3(x)$ can be reduced to the simpler $y^3 + cy + c = 0$.) Rather, Cardano's formula represents a two step process, a tower of field extensions, forming a diagram:

$$\begin{tikzcd}
 & F_2 \arrow[dl, dash] \arrow[d, dash, "\sqrt[3]{\ }"] \\
L_3 \arrow[d, dash] & F_1 \arrow[dl, dash, "\sqrt{\ }"] \\
K_3 \arrow[d, dash] & \\
\mathbb{C} &
\end{tikzcd}$$

Note also that the radicals overshoot the mark, generating the larger field F_2 which contains but is not generated by the desired solution to $p_3(x) = 0$. This motivates the definition of the *resolvent degree* of a field extension, $\mathrm{RD}_k(E/F)$, as the minimal d such that there exists a tower $F_m/\cdots/F_1/F_0 = F$ with $E \subset F_m$ and $\mathrm{ed}_k(F_i/F_{i-1}) \leq d$ for all i. Intuitively, if $E = F[x]/p(x)$, it is the minimal d such that $p(x) = 0$ may be "solved" in terms of other algebraic functions of d or fewer variables, along with field operations.

To abbreviate, write $\mathrm{RD}(n) := \mathrm{RD}_{\mathbb{C}}(L_n/K_n)$. From the quadratic formula and the work of Cardano, Ferrari, Bring and Jerrard, we know that

$$\mathrm{RD}(n) = 1 \quad \text{for } n \leq 5.$$

In this language, Hilbert's sextic conjecture is that

$$\mathrm{RD}(6) = 2,$$

and Hilbert's 13th problem is the conjecture

$$\mathrm{RD}(7) = 3.$$

These conjectures remain wide open. Some progress has been made in lowering the upper bounds for $\mathrm{RD}(n)$, by Hilbert himself [17], Segre [22], Brauer [8], and recently Wolfson [24]. No nontrivial lower bounds have been proven; the current state of the art leaves open the possibility that $\mathrm{RD}(n) = 1$ for all n.

Given the centrality of Tschirnhaus transformations to Hilbert's problem, we hope that the cohomology of $\mathcal{T}_n$ may be useful in meaningfully obstructing the reduction of parameters. More work is warranted to explore the potential of this space, and more broadly to build a bridge, as Arnold began to do, to bring more topological tools to bear on these hard algebraic problems.

REFERENCES

[1] Donu Arapura. The leray spectral sequence is motivic. *Inventiones mathematicae*, 160(3):567–589, Dec 2004.

[2] Vladimir I. Arnold. The cohomology ring of the colored braid group. *Mathematical notes of the Academy of Sciences of the USSR*, 5:138–140, 1969.

[3] Vladimir I. Arnold. On functions of three variables. In Alexander B. Givental et al., editor, *Collected Works: Representations of Functions, Celestial Mechanics and KAM Theory, 1957–1965*, pages 5–8. Springer Berlin, Heidelberg, 2009.

[4] Vladimir I. Arnold. Representation of continuous functions of three variables by the superposition of continuous functions of two variables. In Alexander B. Givental et al., editor, *Collected Works: Representations of Functions, Celestial Mechanics and KAM Theory, 1957–1965*, pages 47–133. Springer Berlin, Heidelberg, 2009.

[5] Vladimir I. Arnold. On cohomology classes of algebraic functions invariant under tschirnhausen transformations. In Alexander B. Givental et al., editor, *Vladimir I. Arnold - Collected Works: Hydrodynamics, Bifurcation Theory, and Algebraic Geometry 1965-1972*, pages 187–190. Springer Berlin, Heidelberg, 2014.

[6] Vladimir I. Arnold. On some topological invariants of algebraic functions. In Alexander B. Givental et al., editor, *Vladimir I. Arnold - Collected Works: Hydrodynamics, Bifurcation Theory, and Algebraic Geometry 1965-1972*, pages 199–221. Springer Berlin, Heidelberg, 2014.

[7] Vladimir I. Arnold. Topological invariants of algebraic functions ii. In Alexander B. Givental et al., editor, *Vladimir I. Arnold - Collected Works: Hydrodynamics, Bifurcation Theory, and Algebraic Geometry 1965-1972*, pages 223–230. Springer Berlin, Heidelberg, 2014.

[8] Richard Brauer. A note on systems of homogeneous algebraic equations. *Bulletin of the American Mathematical Society*, 51:749–755, 1945.

[9] J. Buhler and Z. Reichstein. On the essential dimension of a finite group. *Compositio Mathematica*, 106(2):159–179, Apr 1997.

[10] Weiyan Chen. Twisted cohomology of configuration spaces and spaces of maximal tori via point-counting, 2016. arXiv:1603.03931.

[11] Thomas Church and Benson Farb. Representation theory and homological stability. *Advances in Mathematics*, 245:250–314, Oct 2013.

[12] Pierre Deligne. Théorie de hodge ii. *Publications Mathématiques de l'IHÉS*, 40:5–57, 1971.

[13] Alan H. Durfee. A naive guide to mixed hodge theory. In Peter Orlik, editor, *Singularities, Part 1*, pages 313–320. American Mathematical Society, 1983.

[14] Benson Farb and Jesse Wolfson. Resolvent degree, Hilbert's 13th Problem and geometry. *L'Enseignement Mathématique*, 65(3):303–376, 2019.

[15] Yang-Hui He, John McKay, and Alexander Chen. Erland samuel bring's "transformation of algebraic equations". *inSTEMM Journal*, 1(S1):50–60, Jul 2022.

[16] David Hilbert. Mathematical problems. *Bulletin of the American Mathematical Society*, 8(10):437–479, Jul 1902.

[17] David Hilbert. Über die Gleichung neunten Grades. *Mathematische Annalen*, 97:243–250, 1927.

[18] G.B. Jerrard. *An essay on the resolution of equations.* London: Taylor and Francis, 1859.

[19] Andrey Kolmogorov. On the representations of functions of several variables by means of superpositions of continuous functions of a smaller number of variables. *Doklady Akademii Nauk SSSR*, 108:179–182, 1956.

[20] Serge Lang. *Algebra (3. ed.).* Addison-Wesley, 1993.

[21] G.I Lehrer and Louis Solomon. On the action of the symmetric group on the cohomology of the complement of its reflecting hyperplanes. *Journal of Algebra*, 104(2):410 – 424, 1986.

[22] Beniamino Segre. Sulle irrationalità da cui può farsi dipendere la determinazione di S_k appartenenti a varietà intersezioni complete di forme. *Rendiconti Lincei*, 4:149–154, 1948.

[23] Ehrenfried Walter von Tschirnhaus and R. F. Green. A method for removing all intermediate terms from a given equation. *SIGSAM Bulletin*, 37(1):1–3, Mar 2003.

[24] Jesse Wolfson. Tschirnhaus transformations after Hilbert. *L'Enseignement Mathématique*, 66(3):489–540, 2020.