Séminaire BOURBAKI Juin 2019
71e année, 2018–2019, no 1164, p. 469 à 497
doi:10.24033/ast.1143

HOMOLOGY OF HURWITZ SPACES AND THE COHEN-LENSTRA HEURISTIC FOR FUNCTION FIELDS [after Ellenberg, Venkatesh, and Westerland]

by **Oscar RANDAL-WILLIAMS**

INTRODUCTION

Ellenberg and Venkatesh [**7**] introduced the idea of analyzing the function field analogue of the conjecture of Malle (on the distribution of number fields with given Galois group G) by relating Malle's conjectural upper bound with the asymptotics of $\mathbb{F}_q$-point counts on Hurwitz schemes $\mathsf{H}_{G,n}$. Under the heuristic that each $\mathbb{F}_q$-rational component of $\mathsf{H}_{G,n}$ contains q^n $\mathbb{F}_q$-points they were able to precisely reproduce the upper bound in Malle's conjecture.

In a breakthrough paper, Ellenberg, Venkatesh, and Westerland [**8**] applied similar reasoning to relate the function field analogue of the Cohen-Lenstra heuristic (on the distribution of imaginary number fields with ℓ-parts of their class groups isomorphic to a fixed group A) with the asymptotics of $\mathbb{F}_q$-point counts on certain Hurwitz schemes $\mathsf{Hn}^c_{G,n}$, with $G = A \rtimes \mathbb{Z}^\times$ a generalized dihedral group and $c \subset G$ the conjugacy class of involutions. In this case they were—remarkably—able to *justify* the heuristic that each $\mathbb{F}_q$-rational component of $\mathsf{Hn}^c_{G,n}$ contains q^n $\mathbb{F}_q$-points, by using the Grothendieck-Lefschetz trace formula and a comparison between étale and singular cohomology to reduce it to a problem in algebraic topology, and then solving this problem.

The topological problem concerns the singular homology of the corresponding Hurwitz spaces $\mathsf{Hn}^c_{G,n}(\mathbb{C})^{\mathrm{an}}$. It is easy to show that the total dimension of the homology of these spaces is at most $(2|c|)^n$, but in order to show that the main term in the Grothendieck-Lefschetz trace formula is not overwhelmed as $n \to \infty$ one must show that there is not too much homology in homological degrees which are small compared with n. Ellenberg, Venkatesh, and Westerland accomplish this by proving a *homological stability theorem* for these Hurwitz spaces.

The phenomenon of homological stability was discovered by Quillen, in his analysis of the homology of general linear groups in relation to algebraic K-theory. This is the phenomenon that for many natural sequences of spaces

$$X_1 \longrightarrow X_2 \longrightarrow X_3 \longrightarrow \cdots$$

the induced maps $H_d(X_{n-1}) \to H_d(X_n)$ are isomorphisms as long as $d \leqslant f(n)$, for some divergent function f. In this case $H_d(X_n)$ agrees with the direct limit $H_d(X_\infty) = \operatorname{colim}_i H_d(X_i)$ for all but finitely many n. There is a remarkable range of families $\{X_n\}$ for which homological stability is known to hold. When the X_n are Eilenberg-MacLane spaces for groups G_n, one may take: symmetric groups, braid groups, general linear groups over rings of finite stable rank [**27**], mapping class groups of surfaces [**24**], automorphism groups of free groups [**25**], Higman-Thompson groups [**35**], Coxeter groups [**26**], and many others. When the X_n are moduli spaces, broadly interpreted, one may take: configuration spaces [**31**], classifying spaces for fiber bundles [**22**], classifying spaces for fibrations or block bundles [**2**], and many variants of these.

In a related direction, the more recent development of *representation stability* [**4**]—in which there is a sequence of groups Γ_n acting on the X_n in a compatible manner and the eventual behavior of $H_d(X_n)$ as a Γ_n-representation is studied—may be applied to study asymptotics of weighted point counts (i.e., moments) of sequences of schemes, cf. [**3**].

There is a more or less standard pattern in most proofs of homological stability,(1) in which one constructs an approximation to X_n from the spaces $\{X_i\}_{i<n}$ in a standard way (cf. [**32**]), and then is left with the problem of proving that it is a good approximation, which invariably leads one to analyze the connectivity of certain simplicial complexes associated to the situation in hand. Ellenberg, Venkatesh, and Westerland follow this general strategy, but because the Hurwitz spaces $\mathsf{Hn}^c_{G,n}(\mathbb{C})^{\mathrm{an}}$ are disconnected a new kind of difficulty arises. To surmount this difficulty, they invent a clever piece of homological algebra.

In this exposition of [**8**] I will present their argument differently to the way it appears in that paper, closer to the framework of [**20**] than to the classical approach to homological stability described above. While many of the key steps are unchanged, I find that this streamlined argument clarifies the essential points.

(1) Though Galatius, Kupers, and I have recently proposed another [**20**], which in fortunate circumstances can provide information beyond classical homological stability, for example in the case of mapping class groups of surfaces [**19**].

1. THE COHEN-LENSTRA HEURISTIC FOR FUNCTION FIELDS

Let ℓ be a prime number. The *Cohen-Lenstra distribution* is the probability measure μ on the set of isomorphism classes of finite abelian ℓ-groups given by

$$\mu(A) = \frac{\prod_{i\geqslant 1}(1-\ell^{-i})}{|\mathrm{Aut}(A)|}.$$

The numerator is simply a normalization to make μ into a probability measure: what is important is that an abelian ℓ-group is counted with weight proportional to the reciprocal of the size of its automorphism group, as one does in the cardinality of groupoids.

The original Cohen-Lenstra heuristic [**5**] suggests that when ℓ is odd the ℓ-part of the class groups of imaginary quadratic extensions of $\mathbb{Q}$ is distributed according to μ. The analogue for function fields was first considered by Friedman and Washington [**16**]. In the function field case, $K = \mathbb{F}_q(t)$, a quadratic extension $L \supset K$ is called *imaginary* if it is ramified at infinity, or equivalently if it is of the form $K(\sqrt{f(t)})$ with f a squarefree polynomial of odd degree n.

For n odd let $\mathfrak{S}_n$ denote the set of such imaginary quadratic extensions $L \supset K$ up to K-isomorphism, and for a fixed finite abelian ℓ-group A let $\iota : \mathfrak{S}_n \to \{0,1\}$ denote the indicator function for those L with ℓ-part of their class group isomorphic to A. Define the upper and lower densities

$$\delta^+(q) := \limsup_{n\to\infty} \frac{\sum_{L\in\mathfrak{S}_n} \iota(L)}{|\mathfrak{S}_n|} \quad \text{and} \quad \delta^-(q) := \liminf_{n\to\infty} \frac{\sum_{L\in\mathfrak{S}_n} \iota(L)}{|\mathfrak{S}_n|}.$$

The formulation of the Cohen-Lenstra heuristic proved by Ellenberg, Venkatesh, and Westerland is as follows, where a prime power q is called *good for* ℓ if q is odd and neither q nor $q-1$ is divisible by ℓ.

Theorem 1.1 (Ellenberg-Venkatesh-Westerland). — *Suppose ℓ is odd. As $q \to \infty$ with q good for ℓ, both $\delta^+(q)$ and $\delta^-(q)$ converge to $\mu(A)$.*

I will mainly discuss the solution of the topological problem that Theorem 1.1 reduces to, but will first briefly outline how this topological problem arises.

2. REDUCTION TO COUNTING POINTS OF HURWITZ SCHEMES

The key property of the Cohen-Lenstra distribution μ is that for any finite abelian ℓ-group A the expected number of surjections $A' \to A$ is 1 when A' is distributed according to μ, and in fact this property characterizes μ [**8**, Lemma 8.2]. For $L \in \mathfrak{S}_n$ one writes $m_A(L)$ for the number of surjections from the class group of L to A. Using the above characterisation of the measure μ, Ellenberg, Venkatesh, and Westerland show [**8**, p. 777] that to prove Theorem 1.1 it suffices to prove the following.

Theorem 2.1 (Theorem 8.8 of [**8**]). — *Suppose ℓ is odd and q is good for ℓ. There is a constant $B = B(A)$ such that*

$$\left| \frac{\sum_{L \in \mathfrak{S}_n} m_A(L)}{|\mathfrak{S}_n|} - 1 \right| \leqslant \frac{B}{\sqrt{q}}$$

for all n and q with $\sqrt{q} > B$, $n > B$, and n odd.

A pair (G, c) of a finite group G and a conjugation-invariant subset $c \subset G$ is called *admissible* if c generates G and if whenever $g \in c$ then $g^n \in c$ for all n coprime to $|G|$. If (G, c) is admissible then there are *Hurwitz schemes* $\mathsf{Hn}^c_{G,n}$ over $\mathsf{Spec}(\mathbb{Z}[1/|G|])$ which parametrise connected branched Galois G-covers of the affine line with n branch points and monodromy in the class c. These schemes are formed out of similar schemes $\mathsf{H}_{G,n}$ parametrising branched Galois G-covers of the projective line, which have been constructed by Romagny and Wewers [**33**].

The crucial relation between the Cohen-Lenstra heuristic for function fields and these Hurwitz schemes, which was discovered by Yu [**36**], is as follows. For an odd prime ℓ and an abelian ℓ-group A, form the semi-direct product $G = A \rtimes \mathbb{Z}^\times$, where $\mathbb{Z}^\times$ acts on A by inversion, and let $c \subset G$ denote the conjugacy class of involutions. The pair (G, c) is admissible in the sense defined above, and using Proposition 8.7 of [**8**], which for this choice of (G, c) relates surjections $\mathrm{Cl}(\mathcal{O}_L) \to A$ to branched Galois G-covers of the affine line $\mathsf{A}^1_{\mathbb{F}_q}$ with monodromy in the class c, one shows that

$$\sum_{L \in \mathfrak{S}_n} m_A(L) = 2 \cdot \#\mathsf{Hn}^c_{G,n}(\mathbb{F}_q).$$

On the other hand, the number of squarefree polynomials of degree n is $(q-1)(q^n - q^{n-1})$, but the sets of $\frac{q-1}{2}$ polynomials which differ only by a square in $\mathbb{F}_q^\times$ define isomorphic quadratic extensions, so

$$|\mathfrak{S}_n| = 2 \cdot (q^n - q^{n-1}).$$

To prove Theorem 1.1 it therefore suffices to prove the following.

Statement 2.2. — *Suppose ℓ is odd and q is good for ℓ. There is a constant $B = B(A)$ such that*

$$\left| \frac{\#\mathsf{Hn}^c_{G,n}(\mathbb{F}_q)}{q^n} - 1 \right| \leqslant \frac{B}{\sqrt{q}}$$

for all n and q with $\sqrt{q} > B$, $n > B$, and n odd.

3. POINT COUNTING AND HOMOLOGICAL STABILITY

3.1. Example of the method: squarefree polynomials

I will first illustrate how algebraic topology may be used to prove results such as Statement 2.2 with a much simpler example. The squarefree, monic, degree n polynomials over $\mathbb{F}_q$ are the $\mathbb{F}_q$-points $\mathsf{C}_n(\mathbb{F}_q)$ of a scheme C_n over $\mathsf{Spec}(\mathbb{Z})$. Parametrising monic degree n polynomials by their n coefficients, C_n may be described as the complement in A^n of the zero-locus of the discriminant morphism $\Delta : \mathsf{A}^n \to \mathsf{A}^1$.

As a squarefree, monic, degree n polynomial over $\mathbb{C}$ is determined by its unordered set of n distinct roots, the set of complex points in the analytic topology $\mathsf{C}_n(\mathbb{C})^{\mathrm{an}}$ is precisely the space of configurations of n distinct unordered points in $\mathbb{C}$. This space is well-studied [2] and its $\mathbb{Q}$-homology can be computed by many methods (originally by Arnol′d [**1**]): for $n \geqslant 2$ it is

$$H_i(\mathsf{C}_n(\mathbb{C})^{\mathrm{an}};\mathbb{Q}) \cong \begin{cases} \mathbb{Q} & \text{if } i \text{ is } 0 \text{ or } 1, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Furthermore, the discriminant restricts to a morphism $\Delta : \mathsf{C}_n \to \mathsf{A}^1 \setminus \{0\}$ which on complex points gives a continuous map $\mathsf{C}_n(\mathbb{C})^{\mathrm{an}} \to \mathbb{C}^\times$, and this map induces the above isomorphism on $\mathbb{Q}$-homology.

To evaluate the number $\#\mathsf{C}_n(\mathbb{F}_q)$ of squarefree, monic, degree n polynomials over $\mathbb{F}_q$ one may try to apply the Grothendieck-Lefschetz trace formula to the smooth n-dimensional scheme C_n, in the form

$$\#\mathsf{C}_n(\mathbb{F}_q) = q^n \sum_{i=0}^{n} (-1)^i \mathrm{Tr}(\mathrm{Frob}_q : H^i_{\text{ét}}(\mathsf{C}_n/\overline{\mathbb{F}}_q;\mathbb{Q}_r)^\vee \circlearrowleft),$$

where r is an auxiliary prime number not dividing q. If there were a natural comparison isomorphism

$$H^i_{\text{ét}}(\mathsf{C}_n/\overline{\mathbb{F}}_q;\mathbb{Q}_r) \xrightarrow{\sim} H^i(\mathsf{C}_n(\mathbb{C})^{\mathrm{an}};\mathbb{Q}_r), \tag{2}$$

then by the isomorphism (1) and the fact that it is induced by the discriminant morphism, one would be able to calculate the étale cohomology of $\mathsf{C}_n/\overline{\mathbb{F}}_q$, as well as the action of Frob_q, to be

$$H^i_{\text{ét}}(\mathsf{C}_n/\overline{\mathbb{F}}_q;\mathbb{Q}_r) \cong \begin{cases} \mathbb{Q}_r(0) & \text{if } i = 0, \\ \mathbb{Q}_r(-1) & \text{if } i = 1, \\ 0 & \text{otherwise.} \end{cases}$$

(2) Its fundamental group is Artin's braid group on n strands, β_n, and in fact $\mathsf{C}_n(\mathbb{C})^{\mathrm{an}}$ is an Eilenberg-MacLane space for this group. The homology of this space therefore coincides with the group homology of β_n, and can also be studied from this perspective.