NOTES ON THE MOMENTS PREDICTED BY THE COHEN-LENSTRA HEURISTICS

AARON LANDESMAN

1. COHEN-LENSTRA OVER $\mathbb{Q}$

The Cohen-Lenstra heuristics, introduced in 1984, predict the average size, and more generally, the distribution of ℓ-torsion in class groups of quadratic fields. Here is a consequence of them. Let $IQ_{\leq X}$ denote the set of imaginary quadratic fields (of the form $\mathbb{Q}(\sqrt{-d})$ with discriminant at most X.

Conjecture 1.1. As K ranges over imaginary quadratic fields of discriminant $\leq X$, and ℓ an odd prime number, the average size of ℓ-torsion in the class group of $\mathscr{O}_K$ is 2. That is,

$$\lim_{X\to\infty} \mathbb{E}_{K\in IQ_{\leq X}}(\#\mathrm{Cl}(\mathscr{O}_K)[\ell]) = 2.$$

The Cohen-Lenstra heuristics not only predict the average size of ℓ-torsion in the class group, but more generally predict the full distribution of ℓ-torsion in the class group. In particular, if $H = \mathbb{Z}/\ell\mathbb{Z}$, the above can be phrased as saying that the average number of surjections from $\mathrm{Cl}(\mathscr{O}_K)$ onto H is 1. More generally:

Conjecture 1.2. For any odd order abelian group H,

$$\lim_{X\to\infty} \mathbb{E}_{K\in IQ_{\leq X}}(\#\mathrm{Surj}(\mathrm{Cl}(\mathscr{O}_K), H)) = 1.$$

This has also be generalized to composite odd torsion orders and real quadratic extensions in place of imaginary quadratic extensions. Additionally, it has been generalized to other global fields and to a non-abelian setting.

Remark 1.3. Very little is known about the Cohen-Lenstra heuristics. For example, the only odd prime ℓ for which Conjecture 1.1 is known is $\ell = 3$, which was essentially proven by Davenport and Heilbronn in 1971, before the Cohen-Lenstra heuristics were even formulated. Even proving the case $\ell = 5$ would be a major breakthrough.

2. COHEN-LENSTRA OVER FUNCTION FIELDS

We now pivot to algebraic geometry by working over function fields. Here is a relevant quote by David Mumford. "Algebraic geometry seems to have

acquired the reputation of being esoteric, exclusive, and very abstract, with adherents who are secretly plotting to take over all the rest of mathematics. In one respect this last point is accurate."

Replacing the $\mathbb{Q}$ with $\mathbb{F}_q(t)$, we can make very similar conjectures. In this case, the analog of quadratic fields are hyperelliptic curves, and the ℓ-torsion in their class group corresponds to ℓ-torsion line bundles on the hyperelliptic curve.

Definition 2.1. Let $\mathcal{MH}_{n,q}$ denote the set of function fields K of monic hyperelliptic curves $y^2 = f(t)$ for $f \in \mathbb{F}_q[t]$ squarefree of degree n. Let $\mathscr{O}_K$ denote $\mathbb{F}_q[t,y]/(y^2 - f(t))$, (the normalization of $\mathbb{F}_q[t]$ in the quadratic extension $K/\mathbb{F}_q(t)$,) and let $\mathrm{Cl}(\mathscr{O}_K)$ denote the set of line bundles on $\mathrm{Spec}\ \mathscr{O}_K$ (or ideal classes in $\mathscr{O}_K$).

Remark 2.2. Over function fields, more about Cohen-Lenstra is known. Namely, it is known to hold if one first takes the discriminant tending to ∞, and then takes $q \to \infty$ over odd prime powers so that $q(q-1)$ is relatively prime to the order of H. Namely,

$$\lim_{q\to\infty} \lim_{\substack{n\to\infty \\ n\equiv 1 \bmod 2}} \mathbb{E}_{K\in\mathcal{MH}_{n,q}}(\#\mathrm{Surj}(\mathrm{Cl}(\mathscr{O}_K), H)) = 1.$$

This was shown by Ellenberg-Venkatesh-Westerland in a 2009 preprint, published in 2016 in Annals (part of Venkatesh's fields medal citation).

In an follow up preprint in 2012, EVW claimed to be able to upgrade their result to a computation for fixed q. However, their follow up had a serious error, which they were unable to fix, and has remained open for the last decade. We resolve this error and complete the program.

Theorem 2.3 (L-Levy)**.** *For any odd order abelian group H, and q an odd prime power with $\gcd(H, q(q-1)) = 1$, there is an integer C depending on H so that for $q > C$,*

$$\lim_{n\to\infty, n\equiv 1 \bmod 2} \mathbb{E}_{K\in\mathcal{MH}_{n,q}}(\#\mathrm{Surj}(\mathrm{Cl}(\mathscr{O}_K), H)) = 1.$$

Remark 2.4. The assumption that $q-1$ is prime to H is needed to assume there are no extraneous roots of unity dividing the order of H in $\mathbb{F}_q(t)$. This is necessary because the predicted average size changes there are such roots of unity. We prove that the average size is $\wedge^2 H[h]$ if $h := \gcd(|H|, q-1)$ in general.

Remark 2.5. More generally still, one might want to formulate a non-abelian Cohen-Lenstra conjecture, which replaces the class group (the maximal *abelian* unramified extension) with the maximal unramified extension. We also prove a suitable version of these conjectures over function fields.

3. THE STABLE COHOMOLOGY

We next introduce these Hurwitz spaces so we can precisely state the result we are aiming to prove.

Definition 3.1. Fix a group G and a conjugacy class c. The Hurwitz space $\mathrm{Hur}_n^{G,c}$ is the space of branched Galois G-covers $X \to \mathbb{A}^1_{\mathbb{C}}$, branched over n points with monodromy around those branch points lying in c.

Remark 3.2. Two of the central types of spaces in algebraic geometry are $\mathscr{M}_g$, the moduli of curves, and Hurwitz spaces. For the moduli of curves, the cohomology is known to stabilize as g grows (Harer) and its stable value has also been computed (Madsen Weiss). These computations are both topological in nature, and have far-reaching consequences.

For certain Hurwitz spaces, the cohomology is also known to stabilize (EVW) but its stable value remained open. We compute its stable value.

The general strategy for proving Theorem 2.3 is to construct appropriate moduli spaces, (Hurwitz spaces) whose $\mathbb{F}_q$ points parameterize quadratic extensions together with a surjection of the class group onto H. One is then reduced to counting $\mathbb{F}_q$ points on these spaces. The Grothendieck-Lefschetz trace formula then relates the $\mathbb{F}_q$ points on these spaces to the trace of Frobenius on the cohomology of these spaces.

$$\frac{\#\,\mathrm{Hur}_n^{G,c}}{q^n} = \sum_{i=0}^{2n} (-1)^i \,\mathrm{tr}\left(\mathrm{Frob}_q^{-1} \,|H^i(\mathrm{Hur}_{n,\overline{\mathbb{F}}_q}^{G,c}, \mathbb{Q}_{\ell'})\right)$$

where Frob_q denotes geometric Frobenius.

Therefore, one is reduced to understanding the cohomology sufficiently well. In EVW, they showed these cohomology groups stabilize, but did not compute the stable value. By computing the stable value, we were able to prove this stronger result.

Let Conf_n denote the configuration space over the complex numbers parameterizing degree n divisors in $\mathbb{A}^1_{\mathbb{C}}$. There is a branch locus map $\mathrm{br} : \mathrm{Hur}_n^{G,c} \to \mathrm{Conf}_n$ sending a cover to its branch locus.

Theorem 3.3. *If H is an odd order abelian group, $\mathbb{Z}/2\mathbb{Z}$ acts on H via inversion, and $c \subset H \rtimes \mathbb{Z}/2\mathbb{Z}$ is the conjugacy class of order 2 elements, then there are constants I, J so that for $n > iI + J$, and any component $Z \subset \mathrm{Hur}_n^{G,c}$, the map induced by the branch locus map $H_i(Z;\mathbb{Q}) \to H_i(\mathrm{Conf}_n;\mathbb{Q})$ is an isomorphism.*

Example 3.4. If one takes $G = \mathbb{Z}/3\mathbb{Z}$, the Hurwitz space parameterizes $\mathbb{Z}/3\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z} = S_3$ covers which are simply branched (the monodromy is a transposition). Then, it turns out there is a unique component Z_n of the Hurwitz space parameterizing connected covers branched at n points

(a classical result of Hurwitz). Because the cohomology in low degrees of Z_n agrees with that of Conf_n, the number of $\mathbb{F}_q$ points of these two spaces is nearly the same. As one takes $n \to \infty$, $\frac{\#Z_n(\mathbb{F}_q)}{\#\mathrm{Conf}_n(\mathbb{F}_q)} \to 1$. And this ratio can be interpreted as the same ratio appearing in Theorem 2.3.

We note that it is not even known whether these cohomology groups stabilize for $G = S_4$ and c the conjugacy class of transpositions.

4. Sketch of the proof for $G = S_3$

We now sketch the proof of Theorem 3.3 when $G = S_3$ and $c = \{x, y, z\}$ with $x = (12), y = (23), z = (13)$ is the conjugacy class of transpositions.

It was shown in EVW that the map $[(12)]^2 + [(13)]^2 + [(23)]^2$ (where $[g]^2$ is the map on cohomology obtained by gluing a cover branched at two additional points with monodromy g) induces an isomorphism on cohomology. Moreover, it follows from the group completion theorem and work of Etingof-Grana (essentially the content of the retracted EVW paper) that the composite of two distinct transpositions, say $[x]$ and $[y]$ on homology vanishes on the kernel $V_n := \mathrm{coker}(H^i(\mathrm{Conf}_n) \to H^i(Z_n))$. They are indexed by $g \in c$ and then we have that for all $h \neq g$, we have $[g]^k[h]$ acts by 0 for all $k \geq 0$. We only need to show this subspace is 0.

We then take some cocycle with this property. we want to show it is a sum of boundary terms. To do this, one can then write down a completely explicit cochain complex (the Fox-Neuwirth complex) which computes this cohomology.

Here is how to describe the cochains: To describe the cochains for $C^0(\mathrm{Hur}_n^c)$ consist of tuples of n boxes, each labeled by a transposition in c. To describe C^1, the cochains are similar, except two of the boxes are considered to be joined, so we have a single length 2 rectangle. For C^2 there are either two boxes of length 2 or one of length 3. In general, for C^i, we join i of the boxes.

The boundary maps on these complexes are given by choosing two adjacent rectangles and shuffling. For example, the boundary applied to $(12) \otimes (13)(23)$ is $(12)(13)(23) - (13)(12)(23) + (23)(13)(12)$, where (12) traverses right and acts on the other elements.

We now start with an element of C^1 which is killed by all $[x]$ and $[y]$ We show we are able to modify the given element by a sequence of boundaries until it becomes 0. The first step is to simplify the form of our cocycle. Say it starts as $x_1 + x_2 + x_3 + x_4$ as in the picture. The definition of the boundary map $[h]$ is that it takes all the elements with last place equal to h. Since we are assuming this is 0, we can put g in the last slot of x_2, x_3, and x_4. Similarly, since $[g][h]$ acts by 0, we can put g, g in the last two slots of x_3

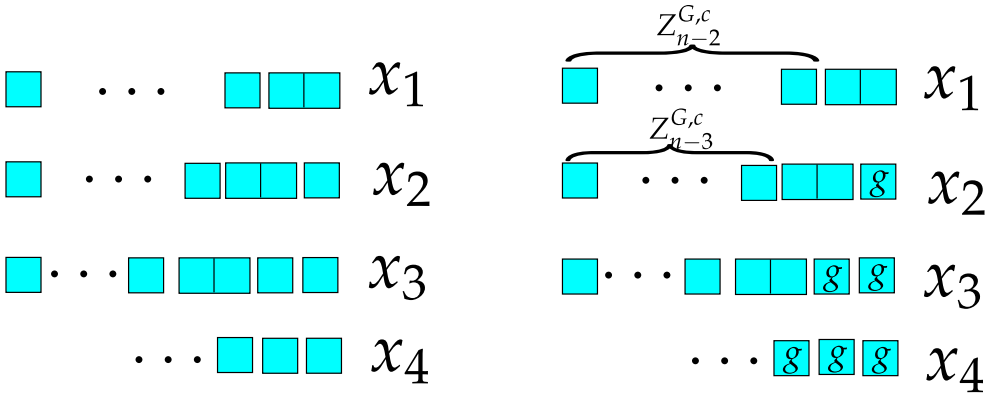

FIGURE 1. The left is a picture of the Fox-Neuwirth/Fuks cell structure of an element of the first stable cohomology. The right hand side pictures some additional constraints we may impose on the cell structure.

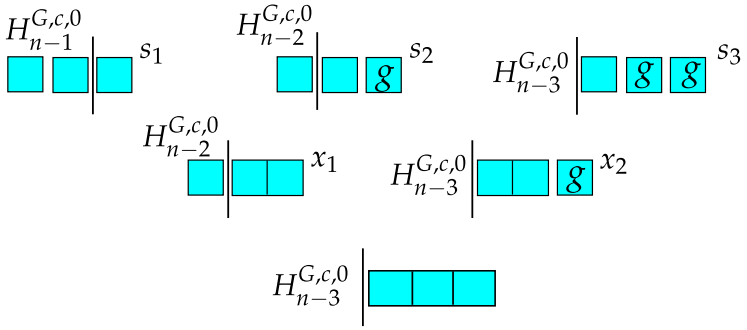

FIGURE 2. This is a picture of part of the two-sided K complex used to compute the stable first cohomology.

and x_4. (Technically, we only know it acts by 0 on cohomology, but we can modify this by cochains so it actually acts as 0.)

Next, because we know these elements are cocycles, we can argue that the parts of x_1 and x_2 to the left of the length 2 rectangles are cocycles, basically because their boundaries can't cancel out with boundaries of any of the other x_i. We next claim we can modify x by a coboundary to remove the x_1 term. This uses the two-sided K-complex, which gives an exact sequence as in the picture. We prove exactness of this complex using an explicit chain homotopy.

Finally, we have arranged that there is a g at the end of all the terms. Applying a version of this 2-sided K-complex again, one can modify terms to make there be two g's on the right, and then three g's and repeat until the

element only consists of g's, at which point we obtain it must be 0 since it doesn't have any boxes of length 2.

For the explicit nullhomotopy describing the two sided K-complex, the space can be described as follows: One draws a square and puts points in the square at different vertical heights, one of which is not (12), differing from each other by some $\varepsilon > 0$. One can create (12)'s from the left and absorb them into the right, but when a (13) or (23) gets absorbed into the right, it becomes a base point. When one pulls (12) above some element of c it gets conjugated by that element.

The nullhomotopy is described as follows. One pulls (12) above the lowest element which is not (12). It gets conjugated to something else, hits the boundary, and becomes the base point.

5. The higher algebra proof

We now give a sketch of the above proof using higher algebra.

We again sketch the proof of Theorem 3.3 when $G = S_3$ and $c = \{x, y, z\}$ with $x = (12), y = (23), z = (13)$ is the conjugacy class of transpositions.

We will use $\mathrm{Hur} = \coprod_{n \geq 0} \mathrm{Hur}_n^{G,c}$ and $\mathrm{Conf} := \coprod_{n \geq 0} \mathrm{Conf}_n$. We can make a space $\mathrm{Conf}^c := \pi_0 \mathrm{Hur} \times_{\pi_0 \mathrm{Conf}} \mathrm{Conf}$ which is constructed by taking each component of Hur_n and replacing it with a copy of Conf_n. The point of doing this is that it now suffices to prove the map $f : C_* \mathrm{Hur} \to C_* \mathrm{Conf}^c$ induces an isomorphism stably.

Recall, it was shown in EVW that the map $U := x^2 + y^2 + z^2$ induces an isomorphism on $C_* \mathrm{Hur}$, and so what we really want to show is that $f[U^{-1}]$ induces an isomorphism on homology.

To get us started, we will need the following lemma:

Lemma 5.1. *The map $f[U^{-1}]$ is an equivalence if and only if $f[U^{-1}][x^{-1}]$ and $(f[U^{-1}])^\wedge_x$ are both equivalences.*

Corollary 5.2. *It suffices to show $f[S^{-1}]^\wedge_{c-S}$ is an equivalence for each $S \subset c$. (This means that we invert each element in S and complete at each element outside of S.)*

We now have to check eight cases for the eight subsets $S \subset c$. But really there are some symmetries, and so it suffices to check the 4 cases depending on whether the size of S is $0, 1, 2$, or 3.

5.3. **Case 1: $S = \emptyset$.** When $S = \emptyset$, we know $U = x^2 + y^2 + z^2$ is invertible, but each of x, y, z also act topologically nilpotently, which means U is both invertible and topologically nilpotent, hence both the source and target of the map are 0.

5.4. **Case 2:** $|S| = 2$**.** Say $S = x, y$. Then, z acts topologically nilpotently and x acts invertibly. This means y acts topologically nilpotently because $y = x^{-1}zx$. Hence, y acts both topologically nilpotently and invertibly, so the source and target are again 0.

5.5. **Case 3:** $S = c$**.** Here we want to show $f[x^{-1}, y^{-1}, z^{-1}]$ is an equivalence. This was shown in EVW's retracted paper, see also Oscar Randal-Williams Bourbaki article.

5.6. **Case 4:** $|S| = 1$**.** Let's say $S = x$. This is the trickiest case, and essentially the whole point of our proof. Now we want to compute

$$f[x^{-1}]^{\wedge}_{y,z} : (C_* \operatorname{Hur}[x^{-1}])^{\wedge}_{y,z} \to (C_* \operatorname{Conf}^c[x^{-1}])^{\wedge}_{y,z}.$$

This map turns out to be induced from a map of pointed spaces

$$\operatorname{res} : \operatorname{Hur}_+ \to \operatorname{Conf}_+$$

which sends all points containing a y or z to the basepoint. The map $f[x^{-1}]^{\wedge}_{y,z}$ is $C_*(\operatorname{res})[x^{-1}]^{\wedge}_{y,z}$.

Now, the key lemma (which is a fairly standard higher algebra lemma that can be proved via descent via taking a resolution of R and S via iterated tensor products of $\pi_0 S$ over them) is the following:

Lemma 5.7. *If $g : R \to S$ is a map of $\mathbb{E}_1$ rings which is a bijection on π_0 and $\pi_0 S \otimes_R \pi_0 S \to \pi_0 S \otimes_S \pi_0 S$ is an equivalence, then g is an equivalence.*

Applying the lemma, and using $\pi_0 \operatorname{Conf}[x^{-1}] \simeq \mathbb{Z}$ it ends up being enough to show

$$\mathbb{Z}_+ \otimes_{\operatorname{Hur}_+} \mathbb{Z}_+ \simeq \mathbb{Z}_+ \otimes_{\operatorname{Conf}_+} \mathbb{Z}_+$$

is an equivalence.

These are essentially bar constructions, and via scanning, we can give very explicit topological models for these spaces. We suggest the reader consult Figure 3 for a pictorial description of the relevant nullhomotopy. We next describe the nullhomotopy in words.

The first space $\mathbb{Z}_+ \otimes_{\operatorname{Hur}_+} \mathbb{Z}_+$ has a model where we put points x, y, z in a configuration space with no two in the same horizontal row. When one point passes above another it gets conjugated by the other, and when x collides into the left or right boundaries in increments the $\mathbb{Z}$ coordinate. When y or z collide, the configuration gets sent to the base point.

The space $\mathbb{Z}_+ \otimes_{\operatorname{Conf}_+} \mathbb{Z}_+$ is similar but there can only be x's in the center.

The difference between these spaces is that the component of the basepoint in the first space gets contracted to the base point in the second space. So, to show this is an equivalence, it suffices to show the component of the base

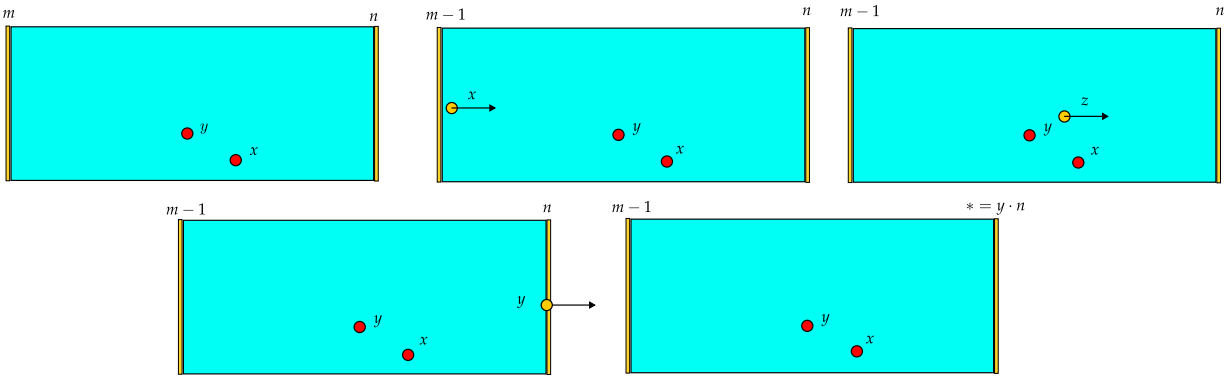

FIGURE 3. This is a pictorial description of the nullhomotopy. We pull a copy of the element x out from the left and let it traverse to the right. When it passes the y it gets replaced with $z = y^{-1}xy$ and then when it passes the x it gets replaced with $y = x^{-1}zx$. When it meets the right boundary, it acts on the boundary by y. Since $y \neq x$, this becomes the base point, and this whole configuration is then identified with the base point.

point is contractible. (For technical reasons, this isn't strictly correct. Rather, we exhaust the space by a sequence of contractible subspaces, depending on the vertical distance between two points of the configuration.)

We then have an explicit contracting homotopy which is given by pulling a point x out from the right just above the lowest point which is y or z. When x passes past that point, it gets conjugated to y or z. We picture this in Figure 3. This then hits the right boundary and gets sent to the base point. This describes a path from any such point to the base point, and gives a contracting homotopy.

6. MALLE'S CONJECTURE

Recall above that we constructed a map $\mathrm{Hur}_n^{G,c} \to \mathrm{Conf}_n$ so that for each component $Z \subset \mathrm{Hur}_n^c$, with n sufficiently large the map $H_i(Z) \to H_i(\mathrm{Conf}_n)$ induced an isomorphism for large n relative to i. Let $\mathrm{CHur}_n^{G,c} \subset \mathrm{Hur}_n^{G,c}$ denote the connected G-covers. More generally we conjecture:

Conjecture 6.1. Let $\mathrm{CHur}_n^{G,c} \subset \mathrm{Hur}_n^{G,c}$ denote the union of connected components parameterizing connected covers. For any group G and, any conjugacy class c generating G and any component $Z \subset \mathrm{CHur}_n^{G,c}$ the induced map $H_i(Z) \to H_i(\mathrm{Conf}_n)$ is an isomorphism for n sufficiently large relative to i.

One can make a more general conjecture with multiple conjugacy classes, where the map to multicolored configuration space is an isomorphism.

Remark 6.2. This is closely related to the Malle-Bhargava heuristic, which roughly says that one should be able to count G-covers by counting the possible discriminants.

Remark 6.3. If one could prove Conjecture 6.1, one would be well on their way to proving a function field version of Malle's conjecture. This is something Ishan and I are currently working on.