ARIZONA WINTER SCHOOL 2014 COURSE NOTES: GEOMETRIC ANALYTIC NUMBER THEORY

JORDAN S. ELLENBERG

1. What is geometric analytic number theory?

First of all:

1.1. **What is analytic number theory?** Already this means different things to different people. The sort of questions I mostly have in mind are typically questions about *asymptotic behavior of arithmetic objects*, for instance:

- How many pairs of coprime integers are there in $[1, N] \times [1, N]$?
- If X is a projective variety, how many points are there in $X(\mathbb{Q})$ with height at most N? This question may use some unfamiliar words, but to emphasize that it's down to earth, I'll comment that when $X = \mathbb{P}^1$ this is precisely the question above about coprime integers! (Note that most people would *not* call this a question of analytic number theory, except for rather special choices of X, but it fits into the framework we're discussing here.)
- How many prime numbers are there less than N?
- How many totally real cubic fields are there with discriminant less than N?
- You can combine the above two questions: how many totally real cubic fields are there with *prime* discriminant less than N?
- If h is a fixed integer, and $F(n)$ is the number of primes between n and $n + h$, what is the variance of $F(n)$ when n is chosen randomly in $[N, 2N]$? (A question of Goldston and Montgomery, recently considered in the geometric setting by Jon Keating and Zeev Rudnick.)
- Is the sign of $\mu(n)$ asymptotically uncorrelated with that of $\mu(n + 1)$, i.e. is it the case that $\sum_{n<N} \mu(n)\mu(n + 1) = o(N)$?
- What is the probability that a quadratic imaginary field $K = \mathbb{Q}(\sqrt{-d})$, where d is random in $[N, 2N]$, has class number prime to 7? More generally, what does the 7-primary part of the class group of K look like "on average"? (This is the subject of the Cohen-Lenstra heuristics and their many variants, some of which will be discussed in this conference.)
- How are the zeroes of the zeta function of a "typical" number field distributed along the line $\Re s = 1/2$?
- If n is a random squarefree integer in $[N, 2N]$, what is the probability that there exists a totally real quintic extension with discriminant N? (I think the limit as N gets large should be $1 - e^{-1/120}$, and I may or may not have time to explain why in these lectures.)

Date: 15-19 Mar 2014.

Remark: characteristic of the kinds of questions we're asking here is that the desired answer is typically not an *exact* formula, but rather an asymptotic. For instance, you probably already know that the probability that two "random" integers are coprime is $\zeta(2)^{-1} = \pi^2/6$. But this does not, of course, mean that

$$|\{(x, y) \in [1, N] \times [1, N]\}| = (\pi^2/6)N^2$$

but rather that

$$\lim_{N\to\infty} N^{-2}|\{(x, y) \in [1, N] \times [1, N]\}| = \pi^2/6.$$

More ambitiously still, we might ask for a *power-saving error term*; that is, an asymptotic of the form

$$|\{(x, y) \in [1, N] \times [1, N]\}| = (\pi^2/6)N^2 + O(N^{2-\delta})$$

for some $\delta > 0$. (In fact, the best known error term is $O(N^{11/54})$ – see Pappalardi's "Survey on k-Freeness.")

1.2. **What is geometric?** This question is, if anything, even more hard to answer precisely than the first one. The word "geometric" does the same work here that it does in the phrase "geometric Langlands program." In order to explain what we mean, it's probably best to work through an example. We'll do one which, from the point of view of classical analytic number theory, is very easy. But that won't stop us from recasting it as a computation in étale cohomology!

1.3. **Number fields and function fields.** The central idea of this course is the analogy between number fields and function fields. To be more precise:

Definition 1. A *global field* is either

- A number field, i.e. a finite extension of $\mathbb{Q}$; or
- The function field of a curve over a finite field $\mathbb{F}_q$, i.e. a field isomorphic to a finite extension of $\mathbb{F}_q(t)$.

These two classes of fields seem pretty different on the face of it; one is characteristic 0, one is characteristic p, one has to do with number theory, the other with algebraic geometry. But under the skin they're quite similar. The beautiful table in section 2.6 of Bjorn Poonen's lecture notes on curves provides a long list of similarities between number fields and function fields; the point of this section is to explicate just a few entries in that table.

For the moment, we will hold off on global fields in general and consider the analogy applied to just *two* global fields; the rational numbers $\mathbb{Q}$ and the rational function field $\mathbb{F}_q(t)$. (To be honest, the latter is really not just one global field but a family of such, parametrized by q – hold that thought!)

The rational numbers $\mathbb{Q}$ have a natural subring $\mathbb{Z}$, which we can think of as the set of rational numbers x such that $|x|_p \leqslant 1$ for all nonarchimedean absolute values $|\cdot|_p$. The function field $\mathbb{F}_q(t)$ also has a subring that jumps right out at us – the ring of polynomials $\mathbb{F}_q[t]$. Is this cut out by valuation bounds in the same way? Sort of, as we now explain.

A function $f \in \mathbb{F}_q(t)$ can (and should!) be thought of as a meromorphic function on $\mathbb{P}^1/\mathbb{F}_q$. For any extension $\mathbb{F}_{q^m}$ of $\mathbb{F}_q$ and any point $x \in \mathbb{P}^1(\mathbb{F}_{q^m})$, we can define a valuation ord_x on $\mathbb{F}_q(t)$ by setting $\text{ord}_x(f)$ to be the order of vanishing of f at x. In particular, $\mathrm{ord}_x(f) > 0$ precisely when f vanishes at x and $\mathrm{ord}_x(f) < 0$ precisely when f has a pole at x. We can turn this valuation into a nonarchimedean absolute value $|\cdot|_x$ by defining

$$|f|_x = q^{-m\mathrm{ord}_x(f)}$$

So the condition $|f|_x \leqslant 1$ says precisely that f doesn't have a pole at x.

As m and x vary, these are actually *all* the absolute values on $\mathbb{F}_q(t)$. So a function f with $|f|_x \leqslant 1$ for all non-archimedean absolute values x is a function with no poles. You probably know (or have heard, or can look up) that such a function must be constant. The constant functions do indeed give a subring $\mathbb{F}_q \subset \mathbb{F}_q(t)$, but not a very interesting one.

What went wrong with our analogy is that integers aren't small in *all* absolute values, only the nonarchimedean ones; in the archimedean absolute value, they tend to be kind of large. In the same way, in order to get an analogue for $\mathbb{Z}$, we might want to look at functions which are small in *all but one* absolute value on $\mathbb{F}_q(t)$. Which one should we single out? There's no canonical choice, but it is notationally convenient to choose the valuation given by ord_∞. One can check that, for a rational function $f = P/Q$ with P, Q polynomials, we have

$$\mathrm{ord}_\infty f = \deg Q - \deg P.$$

Note also that every root of Q is a pole of f at some point other than ∞; so the condition that $|f|_x \leqslant 1$ for all $x \neq \infty$ says that Q is constant. In other words

$$\{f \in \mathbb{F}_q(t) : |f|_x \leqslant 1 \text{ for all } x \neq \infty\} = \mathbb{F}_q[t].$$

Exercise 2. According to our definition of ord_∞, "$P(t)/Q(t)$ has a pole at ∞" should mean $\deg P > \deg Q$. On the other hand, it should also mean that $P(1/t)/Q(1/t)$ has a pole at 0. Check that these two definitions agree.

Remark 1. Our choice of ∞ as the "special point" of $\mathbb{P}^1$ was, of course, arbitrary; there is no reason but notational convenience to prefer $\mathbb{F}_q[t]$ to, say, $\mathbb{F}_q[1/(1-t)]$ as an analogue of $\mathbb{Z}$ in $\mathbb{F}_q(t)$.

With this analogy in place, it's easy to see that much of the apparatus of number theory carries over from $\mathbb{Z}$ to $\mathbb{F}_q[t]$. The positive integers can be thought of as a way of choosing a representative in $\mathbb{Z}$ for each orbit of $\mathbb{Z}^\times = \pm 1$; similarly, *monic* polynomials contain one representative of each orbit of multiplication by $\mathbb{F}_q[t]^\times = \mathbb{F}_q^\times$. Prime numbers correspond to irreducible monic polynomials.

Of all the absolute values on $\mathbb{Q}$, there's only one left unbounded on $\mathbb{Z}$, namely the archimedean valuation, which we usually denote simply by $|\cdot|$. Similarly, of the absolute values on $\mathbb{F}_q(t)$, the only one left unbounded is $|\cdot|_\infty$, and we denote this also by $|\cdot|$ to emphasize the analogy. For each polynomial $f \in \mathbb{F}_q(t)$

$$|f| = q^{-\mathrm{ord}_\infty(f)} = q^{\deg f}.$$

In many of the questions above, we talk about *intervals* in $\mathbb{Z}$. At first, there seems to be an obstacle to talking about intervals in $\mathbb{F}_q[t]$, since the polynomials $\mathbb{F}_q[t]$ do not come in any natural order. But an interval in $\mathbb{Z}$ can be written as

$$\{n : |n - n_0| < d\}$$

and in this language it is quite natural to define an interval in $\mathbb{F}_q(t)$ using the ∞-adic absolute value: an ∞-adic interval is a set of polynomials of the form

$$\{f : |f - f_0| < d\}.$$

For example, the set of polynomials of the form $t^9 + a_2q^2 + a_1q + a_0$ should be thought of as a small interval containing t^9.

1.4. **Squarefree integers and squarefree polynomials.** Let's test out this analogy on a simple example:

How many squarefree integers are there between N and $2N$?

One natural approach: to say N is squarefree is to say it is not divisible by 4, not divisible by 9, not divisible by 25, and so on. (Of course it is redundant to specify indivisibility by 16 when we've already locked in indivisibility by 4; to say N is squarefree is just to say it is not divisible by the square of any prime.) It seems reasonable to think of the conditions of indivisibility by p^2 and indivisibility by q^2 as independent events, when p and q are distinct primes, and this leads one to expect that the probability that a random n is squarefree is

$$(1 - 1/4)(1 - 1/9)(1 - 1/25)\ldots = \prod_p (1 - p^{-2}) = \zeta(2)^{-1}.$$

In other words, if $sf(N)$ is the number of squarefrees in $[N, 2N]$, we expect

$$\lim_{N\to\infty} N^{-1} sf(N) = \zeta(2)^{-1} \tag{1}$$

This is a heuristic, not a proof: it is easy to check that for any fixed P one has

$$\lim_{N\to\infty} N^{-1}\#\{n \text{ in } [N, 2N] \text{ divisible by no prime less than } P\} = \prod_{p<P} (1 - p^{-2})$$

but there is some delicacy in letting P and N go to infinity at the same time, as one must in order to count squarefrees. But never fear, this is easy before-breakfast stuff for classical analytic number theorists. Which we are not. So we turn our attention away from the issue for now, and merely report that (1) is correct. What's more, it's true with a power-saving error term:

$$sf(N) = \zeta(2)^{-1}N + O(N^{1/2}). \tag{2}$$

How does this problem look over $\mathbb{F}_q[t]$? The interval $[N, 2N]$ can be thought of as the set of positive integers whose absolute value is within a constant multiple of N; its analogue is thus the set of monic polynomials f such that $|f| = q^{\deg f}$ is within a constant multiple of N. The absolute value on $\mathbb{F}_q[t]$ is "lumpy," being supported on the highly sparse set $q^{\mathbb{Z}}$. In particular, to require $|f|$ to lie inside the range $[N, 2N]$ is actually to *fix* the value of $\deg f$ at some constant n. (If you're worried about the case $q = 2$, feel free to change $2N$ to $1.999N$, I don't care.)

What's the relationship between n and N? The number of integers in $[N, 2N]$ is about N, while the number of monic polynomials of degree exactly n is q^n, so we should think of N as q^n.

We have now produced analogous entities to everything on the left-hand side of (1): we want to understand

$$\lim_{n\to\infty} q^{-n} sf_q(n)$$

where $sf_q(n)$ is the number of monic squarefree polynomials in $\mathbb{F}_q[t]$ of degree n. It turns out that the limit exists, and

$$\lim_{n\to\infty} q^{-n} sf_q(n) = (1 - 1/q). \tag{3}$$

The analogy holds perfectly – because this is the same answer!

It doesn't *look* like the same answer, but the difference is purely cosmetic, as we now explain. The same heuristic we used to guess the asymptotic for $sf(n)$ makes sense for $sf_q(n)$, and can be proved in the same way:

$$\lim_{N\to\infty} q^{-n} sf_q(n) = \prod_p (1 - |p|^{-2})$$

where now the product is over all *monic irreducible polynomials* p, with $|p| = q^{\deg p}$. In other words,

$$\lim_{N\to\infty} q^{-n} sf_q(n) = \zeta_{\mathbb{F}_q[t]}(2)^{-1}.$$

The miracle here, of course, is that the infinite Euler product of rational functions defining the special value $\zeta_{\mathbb{F}_q[t]}(2)^{-1}$ actually simplifies to a rational function in q! This is part of a much bigger story, which some of you know, and which is orthogonal to the points we're trying to make here.

At this point, it would be reasonable to complain that I haven't proved anything – I've just asserted things! So let me give two proofs of (3): an easy one and a hard one. In fact, we will prove more:

Proposition 3. *$sf_q(n) = (1 - 1/q)q^n$ for all $n > 1$.*

Proof. (Easy) I learned this proof from Mike Zieve. Let $\Sigma_{n,e}$ be the set of monic polynomials of degree n of the form $a(t)b(t)^2$ with a monic squarefree of degree $n - 2e$, and b monic of degree e. Each of the q^n monic polynomials of degree n can be decomposed *uniquely* in this form, so we have

$$q^n = \sum_{e=1}^{\lfloor n/2 \rfloor} |\Sigma_{n,e}| \tag{4}$$

But it is clear that $|\Sigma_{n,e}| = q^e sf_q(n - 2e)$. In particular, the quantity we are trying to compute is $|\Sigma_{n,0}|$. The assertion now follows by induction from (4), given the base cases $sf_q(1) = q$ and $sf_q(0) = 1$. □

Error terms and the case of general function fields. Something striking about Proposition 3 is that it provides an *exact fomula* for the number of squarefree polynomials of degree n. The fact that there's no error term, by contrast with the number field case, seems to put some pressure on our governing analogy. But the apparent difference is slightly misleading.

What if we tried to do the same problem for a *different* function field $K/\mathbb{F}_q$, the function field of some smooth projective curve C? First one faces the question of what, exactly, the corresponding question really is. One could, as before, cut out some nice subring of K and try to count squarefree elements of that ring. But any such choice will force us to choose a point or points of C to play the role of ∞. Maybe it's better to avoid making this choice, and working directly with C. In this case it is much less natural to talk about elements of K, and much more natural to talk about *divisors*.

Definition 4. A *divisor* on C is a finite formal sum of points in $C(\overline{\mathbb{F}}_q)$:

$$D = \sum_{P \in C(\overline{\mathbb{F}}_q)} m_P[P]$$

where $m_P = 0$ for all but finitely many points of C. The *degree* of D is $\sum_P m_P$. We say a divisor is *squarefree* if $m_P \in \{-1, 0, 1\}$ for all P. We say a divisor is *effective* if $m_P \geqslant 0$ for all P.

If f is a function in K, its divisor $\operatorname{div}(f)$ is defined to be $\sum_P \operatorname{ord}_P(f)[P]$.

Note that not every divisor on C is the divisor of a function! And it turns out that problems about *divisors* are often more natural than problems about *functions.* (We will return to this theme in a later lecture, when we talk about the analogy between the class group of a number field and the Jacobian of a curve over a finite field.) Write $\operatorname{Eff}^n(C)$ for the set of effective divisors of C of degree n. There is an asymptotic

$$|\{\text{squarefree effective divisors on } C \text{ of degree } n\}| = \zeta_C(2)^{-1}|\operatorname{Eff}^n(C)| + o(|\operatorname{Div}^n(C)|) \tag{5}$$

but in this case the error term is typically *not* zero.

Exercise 5. (For people who know something, or want to learn something, about the zeta function of a curve over a finite field) Show that

$$\sum_{D \text{ squarefree}} q^{-s\deg(D)} = \frac{\zeta_C(s)}{\zeta_C(2s)}.$$

Use this to give an upper bound for the error term in (5), and show that your bound is sharp in some cases.

There is a lesson here about the analogy between number fields and function fields. It is tempting to think of $\mathbb{F}_q(t)$ as analogous to $\mathbb{Q}$. This is not quite right. Better to take the view that all global fields are in certain respects alike. If something is true for $\mathbb{Q}$ but not for other number fields, it is perhaps not so safe to expect it to be true for $\mathbb{F}_q(t)$; but if it is true for *all* number fields you are on firmer ground.

1.5. **Squarefree polynomials and configuration spaces.** Let us now return to the question of counting squarefree polynomials, and focus on a *difference* between $\mathbb{Z}$ and $\mathbb{F}_q[t]$. How can you tell whether an integer is squarefree? As far as I know, the only way is to trial-divide by squares! In particular, it is unknown, as far as I know, whether there is an algorithm that tests an n-digit integer for squarefreeness in time polynomial in n. (Update: a recent paper of Booker, Hiary, and Keating gives a faster algorithm, which is not polynomial in n, but which under some heuristics is expected to run in time $\exp(n^{2/3})$, so at least it beats trial divison!)

For a polynomial over $\mathbb{F}_q$ the story is quite different. For instance, the cubic polynomial

$$P(t) = t^3 + a_1t^2 + a_2t + a_3$$

is squarefree if and only if

$$a_2^2a_1^2 - 4a_3a_1^3 - 4a_2^3 + 18a_3a_2a_1 - 27a_3^2 \neq 0$$

The expression on the left-hand side above is called the *discriminant* of P. Where does it come from? Simple: if $\theta_1, \ldots, \theta_n$ are the roots of P, we can define

$$\Delta(P) = \prod_{i \neq j} (\theta_i - \theta_j).$$

This quantity is a polynomial in the θ_i which is fixed by the natural S_n-action; thus it is a polynomial in the elementary symmetric functions of the θ_i, which is to say it is a polynomial in the coefficients of P. Plainly, $\Delta(P) \neq 0$ precisely when the roots of P are distinct; equivalently, the squarefree polynomials P are precisely those which have $\Delta(P) \neq 0$.

This allows us to talk about the *moduli space of squarefree polynomials.* The space of all monic polynomials of degree n

$$x^n + a_1 x^{n-1} + \ldots + a_n$$

over a field k is naturally identified with the k-points of the affine space $\mathbb{A}^n$ with coordinates $a_1, \ldots, a_n$. The squarefree polynomials are parametrized by those points on $\mathbb{A}^n$ where $\Delta(P)$, which is a polynomial function in $a_1, \ldots, a_n$, doesn't vanish. In other words, the squarefree polynomials are parametrized by an open subvariety of $\mathbb{A}^n$, which we denote Conf^n.

Why do we call it Conf^n? Well, imagine that k is algebraically closed. Then the squarefree polynomials are naturally in bijection with the set of unordered n-tuples of *distinct* elements of k; the bijection sends a polynomial P to its set of roots, and an n-tuple $z_1, \ldots, z_n$ to the polynomial $\prod_i (t - z_i)$. Unordered n-tuples of distinct points are often called configurations out of some sense of loyalty to physics, where (e.g.) an unordered set of n distinct points in 3-space might be thought of as a configuration of particles (the distinctness hypothesis being there to match the physical constraint that two particles can't simultaneously occupy the same space.)

We can now say

$$sf_q(n) = |\,\mathrm{Conf}^n(\mathbb{F}_q)|$$

In other words, we have found that one kind of counting problem (counting squarefree polynomials) can be re-expressed as a different kind of counting problem (counting $\mathbb{F}_q$-points on a certain variety.) The benefit is that the problem of counting points on varieties over finite fields is one of the most studied problems in arithmetic geometry, and there are many techniques we can bring to bear. We turn to them now.

1.6. **The cohomology of configuration space (complex manifold story).** For the moment, let's turn our attention to $\mathrm{Conf}^n(\mathbb{C})$. We could think of this as a mere set, like $\mathrm{Conf}^n(\mathbb{F}_q)$, but that would be perverse; as the set of complex points of a smooth variety, it forms a complex manifold, and if $\mathrm{Conf}^n(\mathbb{C})$ comes to us endowed with that structure, we should use it.

Conf^1 is a simple space indeed – it is the space of 1-*tuples* of complex numbers, i.e. it is simply $\mathbb{C}$ itself.

Conf^2 is a little more interesting. You can think of it as the space of monic polynomials $t^2 + at + b$ where the discriminant $a^2 - 4b$ doesn't vanish, but that's not terribly enlightening. It's better to think about the space of unordered pairs $\{z_1, z_2\}$ of complex numbers with $z_1 \neq z_2$. There is a natural map

$$\phi : \mathrm{Conf}^2(\mathbb{C}) \to \mathbb{C}^*$$

defined by

$$\phi(\{z_1, z_2\}) = (z_1 - z_2)^2.$$

What do the fibers $\phi^{-1}(a)$ look like? The set of *ordered* pairs (z_1, z_2) with $(z_1 - z_2)^2 = a$ is the union of two disjoint lines: $z_1 - z_2 = \sqrt{a}$ and $z_1 - z_2 = -\sqrt{a}$, where $\sqrt(a)$ is a square root of a. Switching the coordinates switches these two lines; so $\phi^{-1}(a)$ is just a *single* line. From the point of view of homotopy theory, lines are points and $\mathbb{C}^*$ is a circle; so what we've shown here is that $\mathrm{Conf}^2(\mathbb{C})$ is homotopic to the circle, S^1. In particular, its cohomology agrees with that of the circle:

- $H^0(\mathrm{Conf}^2(\mathbb{C}); \mathbb{Q}) = \mathbb{Q}$
- $H^1(\mathrm{Conf}^2(\mathbb{C}); \mathbb{Q}) = \mathbb{Q}$
- $H^i(\mathrm{Conf}^2(\mathbb{C}); \mathbb{Q}) = 0, i > 1$

Now as n gets large, $\mathrm{Conf}^n(\mathbb{C})$ gets more complicated. One way to see this is by considering its fundamental group $\pi_1(\mathrm{Conf}^n(\mathbb{C}))$. What is a loop in $\mathrm{Conf}^n(C)$? We can think of it as a *moving* configuration of n points in the complex plane, such that no two points are ever allowed to collide. Such a path determines a *braid.* To make this more physical, we can think of a loop as a map $\gamma : [0, 1] \to \mathrm{Conf}^n(\mathbb{C})$. Then the map

$$\gamma \times \mathrm{id} : [0, 1] \to \mathrm{Conf}^n(\mathbb{C}) \times [0, 1]$$

has as its image n *strands*, which wind around each other but are never allowed to cross. Composition of loops corresponds to stacking one braid diagram atop another. To sum up, $\pi_1(\mathrm{Conf}^n(\mathbb{C}))$ is nothing more than the n-strand braid group B_n.

Note that the braid must begin and end at whatever basepoint in $\mathrm{Conf}^n(\mathbb{C})$ we've silently chosen, but because that basepoint is an *unordered* configuration space, the braid may permute the n points in some way; this gives us a homomorphism $B_n \to S_n$, whose kernel P_n is called the *pure braid group on n strands.*

It turns out, though we won't really need this here, that from the homotopy theorist's point of view, $\mathrm{Conf}^n(\mathbb{C})$ basically *is* the braid group. By this we mean that there is a contractible space K with a free action of B_n such that $\mathrm{Conf}^n(\mathbb{C})$ is homotopy equivalent to K/B_n. (We say that $\mathrm{Conf}^n(\mathbb{C})$ is a $K(\pi, 1)$ for the group $\pi = B_n$.) It follows for example, that the cohomology of $\mathrm{Conf}^n(\mathbb{C})$ is the same thing as the group cohomology of the discrete group B_n.

The braid group B_n gets more and more complicated as n gets larger; for instance, it cannot be generated by fewer than $n - 1$ elements. So in some sense the spaces Conf^n must be getting more and more complicated as topological spaces.

Which makes the following theorem of Arnol'd extremely surprising:

Theorem 6. *For all $n > 1$,*

- $H^0(\mathrm{Conf}^n(\mathbb{C}); \mathbb{Q}) = \mathbb{Q}$
- $H^1(\mathrm{Conf}^n(\mathbb{C}); \mathbb{Q}) = \mathbb{Q}$
- $H^i(\mathrm{Conf}^n(\mathbb{C}); \mathbb{Q}) = 0, i > 1$

In other words, while the space is getting more and more complicated, its cohomology with rational coefficients is not! Through the lens of cohomology with rational coefficients, $\mathrm{Conf}^n(\mathbb{C})$ looks just like a circle for all $n > 2$.

This is an example of the phenomenon of *homological stabilization*, a major theme in contemporary topology. (Another long story there's no room for here – but read anything about Harer's theorem and the Madsen-Weiss theorem settling the Mumford conjecture...)

Exercise 7. We know that $H^1(\mathrm{Conf}^n(\mathbb{C}), \mathbb{Q}) = H^1(B_n, \mathbb{Q})$. Using the standard presentation of the braid group by $n-1$ generators, show that $H^1(B_n, \mathbb{Q}) = \mathbb{Q}$ for all $n > 1$.

1.7. **The cohomology of configuration space (étale cohomology story).** What we will explain in this section is how Theorem 6, though it looks rather different from what we've proved so far, is yet another manifestation of the formula for the number of squarefree polynomials! It is, in a sense we shall make explicit, a *geometric* version of that formula.

Let X be a variety over $\mathbb{F}_q$. The *Grothendieck-Lefschetz trace formula* provides us with a means of counting points on $X(\mathbb{F}_q)$. It looks like this:

$$|X(\mathbb{F}_q)| = \sum_i (-1)^i \mathrm{Tr}\,\mathrm{Frob}\,|H^i_{et;c}(X/\overline{\mathbb{F}}_q; \mathbb{Q}_\ell) \tag{6}$$

Now comes a pedagogically difficult moment. If you are already familiar with the machinery of étale cohomology, this is familiar, and if not, it probably looks totally meaningless. In order to really prove theorems in the mode I'm about to describe, it's of course critical to be conversant with etale cohomology and its properties. But one of my goals in these notes is to make the case that you can get a good sense of what's going on by "pretending etale cohomology is topological cohomology" – and I will attempt to provide the properties of étale cohomology that are needed in black-box form.

Let us just say for now that etale cohomology is a cohomology theory that makes sense for arbitrary schemes, in particular varieties over finite fields, functorially associating vector spaces to schemes. What's more, when X is a scheme over a finite field $\mathbb{F}_q$, the cohomology group

$$H^i_{et}(X/\overline{\mathbb{F}}_q; \mathbb{Q}_\ell)$$

carries a natural invertible operator called Frob; this is the operator whose trace we refer to in the Grothendieck-Lefschetz trace formula above.

This cohomology theory enjoys many of the same properties as does singular cohomology of manifolds – for instance, when X is a smooth variety over a finite field, there is a Poincare duality between étale cohomology with compact support and étale cohomology, which alllows us in this case to write

$$|X(\mathbb{F}_q)| = q^{\dim X} \sum_i (-1)^i \mathrm{Tr}\,\mathrm{Frob}\,|H^i_{et}(X/\overline{\mathbb{F}}_q; \mathbb{Q}_\ell)^\vee \tag{7}$$

What's more, under favorable circumstances, étale cohomology not only behaves like singular cohomology of manifolds, it *agrees* with singular cohomology of a manifold. For this to make sense, we need X to be defined over a ring of mixed characteristic, like $\mathbb{Z}$, so that we can consider the basechange of X to $\mathbb{F}_q$, which is a variety over a finite field, and also the basechange of X to $\mathbb{C}$, which is a complex algebraic variety, and contemplate the relationship between these two a priori very different objects. This a much longer story than I can reasonably tell in this space. But the idea is that we would like to be able to say

$$\dim H^i_{et}(X/\overline{\mathbb{F}}_q; \mathbb{Q}_\ell) = \dim H^i(X(\mathbb{C}); \mathbb{Q}_\ell)$$

where the right-hand side is good old singular cohomology. If I'm going to say two $\mathbb{Q}_\ell$ vector spaces have the same dimension, why aren't I just saying they're isomorphic? It's because I don't want to give you the wrong impression that there's a *canonical* isomorphism between them. There is not. Indeed, this point is brought home very clearly from the fact that the left hand side admits a Frobenius action and the right hand side does not.

What does "favorable circumstances" mean? It turns out to be enough for X to be projective and smooth. But in practice, these conditions are often not satisfied, and indeed, the comparison tends to be valid somewhat more generally. For instance, the étale cohomology of configuration space in characteristic p is just what we might have expected, given Arnol'd's computation of the cohomology of $\mathrm{Conf}^n(\mathbb{C})$.

Proposition 8. *For all $n > 2$,*

- $H^0_{et}(\mathrm{Conf}^n/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell) = \mathbb{Q}_\ell$
- $H^1_{et}(\mathrm{Conf}^n/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell) = \mathbb{Q}_\ell$
- $H^i_{et}(\mathrm{Conf}^n/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell) = 0, i > 1$

Moreover, Frob *acts as* 1 *on* $H^0_{et}(\mathrm{Conf}^n/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell)$ *and as* q *on* $H^1_{et}(\mathrm{Conf}^n/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell)$.

Proof. It turns out to be very useful to consider another space, the *pure configuration space* PConf^n which parametrizes *ordered* n-tuples of distinct points. This space is nothing other than the complement in $\mathbb{A}^n$ of the union of hyperplanes $(z_i = z_j)_{i\neq j}$. It carries a natural fixed-point free action of S_n by permutation of coordinates, and $\mathrm{Conf}^n = \mathrm{PConf}^n/S_n$.

The topology of the complement of a hyperplane arrangement in affine space is a very rich topic, pioneered by the work of Orlik and Solomon, who show (among many other things) that the Betti numbers of such a complement has a very agreeable combinatorial description in terms of the intersection lattice[1] of the hyperplanes being removed. Results of Lehrer and Kim show that the same holds for étale cohomology, which extends the reach of the theorem from complex hyperplane arrangements to arrangements over more general bases. (Their results also tell you that the action of Frobenius on H^i_{et} of the complement of a hyperplane arrangement is always multiplication by q^i.) In the case of PConf^n, the intersection lattice is clearly the same whether we work in characteristic 0 or characteristic p; we thus have that

$$\dim H^i_{et}(\mathrm{PConf}^n/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell) = \dim H^i(\mathrm{PConf}^n(\mathbb{C});\mathbb{Q}_\ell)$$

In fact, the Orlik-Solomon construction is sufficiently functorial that the two cohomology groups are isomorphic not only as $\mathbb{Q}_\ell$-vector spaces but as $\mathbb{Q}_\ell[S_n]$-modules, the S_n-action on the cohomology groups being inherited from that on PConf^n. It follows that

$$\begin{aligned}\dim H^i_{et}(\mathrm{Conf}^n/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell) &= \dim H^i_{et}(\mathrm{PConf}^n/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell)^{S_n} \\ &= \dim H^i(\mathrm{PConf}^n(\mathbb{C});\mathbb{Q}_\ell)^{S_n} = \dim H^i(\mathrm{Conf}(\mathbb{C}),\mathbb{Q}_\ell)\end{aligned}$$

and the Proposition follows. □

Remark 2. The combinatorial description of Orlik and Solomon is certainly enough to recover Proposition 8 without invoking Arnol'd's result on the braid group.

[1] That is, the poset of all finite intersections of hyperplanes, ordered by inclusion.

When we put the information from Proposition 8 into (7) something very appealing happens; all the terms but the first two vanish, and we are left with

$$|\operatorname{Conf}^n(\mathbb{F}_q)| = q^n(\operatorname{Tr}\operatorname{Frob}|H^0_{et}(X/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell)^\vee - \operatorname{Tr}\operatorname{Frob}|H^1_{et}(X/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell)^\vee) = q^n(1-1/q)$$

So we recover the fact that the probability that a random polynomial is squarefree is $1-1/q$ as a fact about *topology.*

Remark 3. The absence of an error term in Proposition 3, in this context, follows from the fact that we have control of the contribution of H^i_{et} for *all i*. Typically, this is too much to ask for; there will be a *stable range* of degree i in which we can say something about the dimension of the étale cohomology and the Frobenius eigenvalues on it, and the size of this stable range will govern the size of the error term in our asymptotic.

1.8. **The three columns.** We have studied three problems;

- Computing the number of squarefree integers in a range;
- Computing the number of squarefree polynomials in $\mathbb{F}_q[t]$ of a given degree;
- Computing geometric facts about the space of squarefree polynomials in $\mathbb{C}[t]$ of a given degree.

And we have seen that all these problems are related to one another. There is a very tight analogy between the counting problems of the first and second kind. In some cases, counting squarefrees being one such, there are even arguments that work over an arbitrary global field, solving both cases at once. But usually the relation is "merely" analogistic; one expects or at least *hopes* to encounter analogous answers to analogous problems.

The passage between the third and second problems is more robust; by proving topological theorems about a certain moduli space over $\mathbb{C}$, one can deduce theorems about numbers of $\mathbb{F}_q$-points on "the same" moduli space over $\mathbb{F}_q$. This is the plan: use topology to *prove theorems* about arithmetic questions in the function field case, which in turn *gives us ideas* about the analogous arithmetic questions in the number field case.

This strategy is in no way new. Here's Andre Weil, writing to his sister Simone in 1948 (Max Krieger, trans.)

> The classical theory (that is, Riemannian) of algebraic functions over the field of constants of the complex numbers is infunitely richer; but on the one hand it is too much so, and in the mass of facts some real analogies become lost; and above all, it is too far from the theory of numbers. One would be totally obstructed if there were not a bridge between the two. And just as God defeats the devil: this bridge exists; it is the theory of the field of algebraic functions over a finite eld of constants.

Or:

> The mathematician who studies these problems has the impression of deciphering a trilingual inscription. In the first column one finds the classical Riemannian theory of algebraic functions. The third column is the arithmetic theory of algebraic numbers. The column in the middle is the most recently discovered one; it consists of the theory of algebraic functions over finite fields. These texts are the only source of knowledge about the languages in which they are written; in each column, we understand only fragments.

1.9. **What is geometric? (reprise).** We can now say what we mean by geometric analytic number theory.

- We are faced with a problem in analytic number theory or arithmetic statistics. We consider an analogous problem in which the number field is replaced by a function field, often a rational function field $\mathbb{F}_q(t)$.
- We interpret this problem as the problem of estimating $|X_n(\mathbb{F}_q)|$, for some sequence of varieties $X_1, X_2, \ldots$ which admit nice models over $\operatorname{Spec}\mathbb{Z}$.
- We formulate an assertion about the topology or geometry of $X_n(\mathbb{C})$ which would imply the desired estimate of the point-count over $\mathbb{F}_q$.

In the example we've seen, $X_n = \operatorname{Conf}^n$ and $X_n(\mathbb{F}_q)$ is the set of monic squarefree degree-n polynomials over $\mathbb{F}_q$. We briefly recount a few more examples, some of which we will return to in more depth later in these notes.

Linear factors of squarefree polynomials. Suppose we want to count pairs (f, ℓ) where f is a monic squarefree polynomial and ℓ is a linear factor of f. These are parametrized by the variety $\operatorname{PConf}^n / S_{n-1}$, and

$$\frac{|\operatorname{PConf}^n / S_{n-1}(\mathbb{F}_q)|}{|\operatorname{Conf}^n(\mathbb{F}_q)|} = \text{average number of linear factors of a squarefree polynomial.}$$

It is not hard to check (see e.g. Prop 4.4 and §4.3 of Church-Ellenberg-Farb, "Representation stability...") that this average is $(1+q^{-1})^{-1}$.

Distribution of the Möbius function. Suppose we want to study the distribution of the Möbius function on monic polynomials. One can check that $(-1)^{\deg f}\mu(f)$ is 1 if the discriminant $\Delta(f)$ is a quadratic residue in $\mathbb{F}_q^*$, -1 if $\Delta(f)$ is a non-residue, and 0 if $\Delta(f) = 0$. In other words, $(-1)^{\deg f}\mu(f) + 1$ is the number of square roots of $\Delta(f)$. Thus, it is natrual to define X_n to be the variety parametrizing pairs (f, x) where f is a monic polynomial and x is a square root of $\Delta(f)$. The variety X_n is a double cover of $\mathbb{A}^n$, ramified at the locus where the discriminant vanishes, and $q^{-n}|X_n(\mathbb{F}_q)| - 1$ (up to a sign) is the average value of the Möbius function over monic polynomials of degree n.

Autocorrelation of the Möbius function. The Chowla conjecture concerns autocorrelation between shifts of the Möbius function. For instance: what is the average of $\mu(n)\mu(n+1)$? When we ask this question in the function field case, we find we are counting points on the variety X_n parametrizing pairs (f, x) where x is a square root of the product $\Delta(f)\Delta(f+1)$. The study of this variety is the topic of the 2012 paper of Carmon and Rudnick, "The autocorrelation of the Mobius function and Chowla's conjecture" – more about this later.

Cohen-Lenstra heuristics. The Cohen-Lenstra heuristics address questions about the ℓ-primary parts of class groups of number fields. For example: what is the average number of ℓ-torsion elements in the class group of a quadratic imaginary field K? By class field theory, it's the same to ask: what is the average number of everywhere unramified $\mathbb{Z}/\ell\mathbb{Z}$-extensions of K? Over $\mathbb{F}_q(t)$, the analogue of a quadratic imaginary number field K is a hyperelliptic curve C, and the everywhere unramified $\mathbb{Z}/\ell\mathbb{Z}$ covers of C are parametrized by the $\mathbb{F}_q$-rational ℓ-torsion points of the Jacobian $\operatorname{Jac}(C)$. (I am being a little fast and loose here – I will pin this down more carefully in the section on the Cohen-Lenstra conjecture.) So the variety that computes the average size of the ℓ-torsion in the class group is the variety parametrizing pairs (C, P) where C is a hyperelliptic curve of some genus g and P is an ℓ-torsion point on $\mathrm{Jac}(C)$. In other words, X_g is a moduli space of hyperelliptic genus-g curves with ℓ-level structure.

Linnik/Malle/Bhargava conjectures. How many degree-d number fields are there with discriminant in $[N, 2N]$? An old conjecture of Linnik – now a theorem of Davenport and Heilbronn when $d = 3$ and of Bhargava when $d = 4, 5$ – holds that this quantity is $c_d N + o(N)$ (and Bhargava has a prediction for the constant for any n.) The corresponding function field question is: how many degree-d extensions $K/\mathbb{F}_q(t)$ are there whose discriminant is a divisor of degree n? Any such extension of function fields can be "spread out" uniquely to a branched covering of smooth curves $Y \to \mathbb{P}^1$ of degree d with n branch points. Branched covers of this kind are parametrized by moduli spaces called *Hurwitz spaces*, and the function-field analogue of Linnik's conjecture asks for an asymptotic formula for the number of $\mathbb{F}_q$-rational points on a Hurwitz space.

1.10. **Homological stability and asymptotic point-count.** Typically, the sort of assertion one seeks is a statement of homological stabilization. The governing idea is the following. We keep in mind that, for any variety X over $\mathbb{Q}$, the étale Betti number $\dim H^i_{et}(X_n/\bar{\mathbb{Q}}, \mathbb{Q}_\ell)$ agrees with the topological Betti number $\dim H^i(X(\mathbb{C}), \mathbb{Q})$.

Proposition 9. *Let $X_1, X_2, \ldots$ be a sequence of smooth algebraic varieties over* $\mathrm{Spec}\,\mathbb{Z}[1/N]$. *Suppose that*

- *(Homological stabilization) There is a constant $\alpha > 0$ such that, for all $i \leqslant \alpha n$, there is an isomorphism*
$$H^i_{et}(X_n/\bar{\mathbb{Q}}, \mathbb{Q}_\ell) \cong H^i_{et}(X_{n+1}/\bar{\mathbb{Q}}, \mathbb{Q}_\ell)$$
which commutes with the action of Galois on either side.
- *(Subexponential Betti numbers) There is a constant C such that* $\dim H^i_{et}(X_n/\bar{\mathbb{Q}}, \mathbb{Q}_\ell)$ *is at most C^n for all i.*
- *(Comparison) For all p not dividing N and all i,*
$$\dim H^i_{et}(X_n/\bar{\mathbb{Q}}, \mathbb{Q}_\ell) = \dim H^i_{et}(X_n/\bar{\mathbb{F}}_p, \mathbb{Q}_\ell)$$

For each i, write $H^i_{et}(X_\infty/\bar{\mathbb{F}}_p, \mathbb{Q}_\ell)$ for the direct limit in n of $H^i_{et}(X_n/\bar{F}_p, \mathbb{Q}_\ell)$. Then, for all $q > C^{2/\alpha}$,

$$\lim q^{-\dim X_n}|X_n(\mathbb{F}_q)| = \sum_{i=0}^{\infty}(-1)^i \mathrm{T}r\,\mathrm{Frob}\,|H^i_{et}(X_\infty/\overline{\mathbb{F}}_q; \mathbb{Q}_\ell)^\vee \tag{8}$$

Proof. From the homological stability, we know that
$$H^i_{et}(X_n/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell) \cong H^i_{et}(X_\infty/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell)$$
whenever $i \leqslant \alpha n$. Thus

$$q^{-\dim X_n}|X_n(\mathbb{F}_q)| - \sum_{i=0}^{\infty}(-1)^i \mathrm{Tr}\,\mathrm{Frob}\,|H^i_{et}(X_\infty/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell)^\vee \quad = \tag{9}$$

$$\sum_{i=\alpha n}^{\infty}(-1)^i(\mathrm{Tr}\,\mathrm{Frob}\,|H^i_{et}(X_n/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell)^\vee - \mathrm{Tr}\,\mathrm{Frob}\,|H^i_{et}(X_\infty/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell)^\vee) \tag{10}$$

By the subexponential Betti numbers and comparison, we have

$$\dim H^i_{et}(X_\infty/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell) = \dim H^i_{et}(X_{i/\alpha}/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell) < C^{i/\alpha}$$

Moreover,

$$\dim H^i_{et}(X_n/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell) < C^n < C^{i/\alpha}.$$

On the other hand, it follows from Deligne's proof of the Weil conjectures (you knew this had to be coming at some point, right?) that the eigenvalues of Frobenius acting on $\dim H^i_{et}(X_{i/\alpha}/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell)$ are at least $q^{i/2}$. So the contribution of the ith term in (10) is at most

$$2q^{-i/2}C^{i/\alpha} = 2(q^{-1/2}C^{1/\alpha})^i$$

The exponentiand is less than 1 by our hypothesis on q; thus, the error term goes to 0 as i grows and we are done. □

Exercise 10. Show that the sequence $X_n = \mathbb{P}^n$ satisfies the hypotheses (and thus the conclusion) of Proposition 9.

1.11. **Aside: motivic analytic number theory.** Suppose we have a sequence of varieties X_n such that $q^{-\dim X_n}|X_n(\mathbb{F}_q)|$ approaches a limit. How confident should we be that this regularity is explained geometrically by homological stabilization? Certainly this doesn't *have* to happen. For example, we could have X_n be the union of $\mathbb{A}^n$ with n^2 points; then $H^0(X_n)$ is badly nonstable, while the point-count $|X(\mathbb{F}_q)| = q^n + n^2$ behaves itself as nicely as you please.

But homological stabilization is not the only reason a variety could have a number of points that asymptotes to a constant multiple of q^n. We observed in exercise 10 that

$$q^{-n}|\mathbb{P}^n(\mathbb{F}_q)| \to 1 + 1/q + 1/q^2 + \dots$$

But another way to see this is that we can decompose $\mathbb{P}^n$ as a disjoint union of affine spaces; namely, the locus $x_0 \neq 0$ is a copy of $\mathbb{A}^n$, the locus $x_0 = 0, x_1 \neq 0$ is a copy of $\mathbb{A}^{n-1}$, and so on, and this decomposiiton yields a different argument for the formula

$$|\mathbb{P}^n(\mathbb{F}_q)| = |\mathbb{A}^n(\mathbb{F}_q)| + |\mathbb{A}^{n-1}(\mathbb{F}_q)| + \dots + |\mathbb{A}^0(\mathbb{F}_q)| = q^n + q^{n-1} + \dots + 1.$$

What we're doing here is establishing an identity in the *ring of motives.*

Definition 11. The ring of motives $K_0(\mathrm{Var}_R)$ is the quotient of the free abelian group on varieties defined over R by the relations

X = [Y] if there's an isomorphism between X and Y;

- $[X]-[Y] = [X\backslash Y]$ whenever Y is a closed subvariety of X. The product structure on $K_0(\mathrm{Var}_\mathbb{Q})$ is given by direct product of varieties.

The ring of motives admits a point-counting homomorphism $e_q : K_0(\mathrm{Var}_{\mathbb{Z}[1/N]}) \to \mathbb{F}_q$ for all q prime to N, defined by

$$e_q([X]) = |X(\mathbb{F}_q)|$$

This is evidently compatible with the relations in the ring of motives and with the product structure. Our argument above on $\mathbb{P}^n$ is really asserting a motivic identity

$$[\mathbb{P}^n] = [\mathbb{A}^n] + [\mathbb{A}^{n-1}] + \dots + [\mathbb{A}^0]$$

which implies the formula for $|\mathbb{P}^n(\mathbb{F}_q)|$ for all q at once, by application of the various e_q.

What's more, our "easy" proof of Proposition 3 is also actually a motivic argument, since it rests on the operation of "scissoring" A^n into pieces, the biggest one of which is Conf^n.

Exercise 12. (for people who like the ring of motives) Rewrite the proof of Proposition 3 to prove the identity

$$[\mathrm{Conf}^n] = [A^n] - [A^{n-1}]$$

in $K_0(\mathrm{Var}_{\mathbb{Z}})$.

Much, much more about this point of view can be found in the joint work of Melanie Matchett Wood and Ravi Vakil.

2. The Chowla conjecture, the large q regime, components

For the second part of the course, we will try to talk about some of the themes that frequently arise when you work on the geometric analogues of problems in analytic number theory. Our method will be to consider a series of examples, from each of which we can derive a lesson.

For this section, the lesson is

> Facts about arithmetic statistics in the "large q regime" correspond to geometric facts about irreducible components of moduli spaces.

To be more precise: suppose there is some arithmetic counting function whose asymptotic behavior we would like to understand. In the function field setting, this may be some kind of count or average over the monic polynomials of degree n in $\mathbb{F}_q[t]$; we denote this average by $F(q, n)$. For example, in the previous section, we might take $F(q, n)$ to be the number of squarefree polynomials among the monic ones; in that case, we derived an exact formula for $F(q, n)$, but more generally we might want to show the existence of, or better yet compute, the limit

$$\lim_{n\to\infty} q^{-n} F(q, n)$$

The "large q limit" version of this problem asks us to compute the value of

$$\lim_{n\to\infty} \lim_{q\to\infty} q^{-n} F(q, n)$$

This is of special interest in cases where we expect the value of $\lim_{n\to\infty} q^{-n} F(q, n)$ to be independent of q; if this is the case, than one may well expect its limiting value to be equal to the large q limit above. (Of course, this is not a formal statement about limits; consider the function $q/(n + q)$, which goes to 1 in the large q limit (identically in n) but to 0 in the large n limit (identically in q.)

We will illustrate this point with a discussion of the geometric analogue of the Chowla conjecture.

(Good references for this section: Terry Tao's 2012 blog post "The Chowla conjecture and the Sarnak conjecture" and Carmon and Rudnick's 2012 paper "The autocorrelation of the Mobius function and Chowla's conjecture for the rational function field.")

The Möbius function $\mu : \mathbb{Z} \to \{-1, 0, 1\}$ is defined by

$$\mu(n) = (-1)^k$$

When n is the product of k distinct primes, and 0 otherwise. Note that $(\mu(n))^2$ is the characteristic function of the set of squarefree integers.

It is customary to think of μ as a random sign attached to a squarefree integer. Of course, it is not really random at all – $\mu(n)$ reflects the multiplicative structure of n. But from the point of view of the *additive* structure of integers, it should look essentially random. For instance, one might expect $\mu(n)$ and $\mu(n+1)$ to be in some sense probabilistically independent. This would mean, at the very least, that one should expect

$$\sum_{n \leqslant N} \mu(n)\mu(n+1)$$

to be small, because there should be a lot of cancellation in the sum. If it were truly a sum of random signs, it would be expected to be about $\sqrt{N}$.

But that is too much to ask for – note that even in the case of the average of Möbius itself, that

$$\sum_{n \leqslant N} \mu(n) = O(N)$$

is already about as hard as the prime number theorem, and the conjectural assertion that

$$\sum_{n \leqslant N} \mu(n) = O(N^{1/2} + \epsilon) \tag{11}$$

is equivalent to the Riemann hypothesis!

What happens when we ask about this question over function fields? As always, the analogue of summing over the interval $[1, N]$ – or, more precisely, of summing over an interval $[N, cN]$ – is summing over all divisors on a curve C of a given degree n. In this setting, the summatory function of Möbius has a very clean analytic intepretation, which we leave it as an exercise to check:

$$\sum \mu(D)|D|^{-s} = \zeta_C(s)^{-1}$$

Here, $|D|$ denotes $q^{\deg D}$, the natural notion of the "size" of a divisor. We can then recover

$$\sum_{\deg D = n} \mu(D)$$

as the q^{-ns} term in the power series expansion of $\zeta_C(s)^{-1}$. In particular, when $C = \mathbb{A}^1/\mathbb{F}_q$, i.e. when we replace $\mathbb{Z}$ with $\mathbb{F}_q[T]$, we have

$$\zeta_{\mathbb{A}^1}(s)^{-1} = (1 - q^{1-s})$$

which tells us that

$$\sum_{\deg f = n} \mu(D) = 0$$

where the sum ranges over monic polynomials in $\mathbb{F}_q[t]$ of degree n. For a general projective curve C, we have

$$\sum_D \mu(D)|D|^{-s} = \zeta_C(s)^{-1} = (1 - q^s)(1 - q^{1-s})P(q^{-s})^{-1}$$

where P is the characteristic polynomial of Frobenius acting on $H^1(C/\overline{\mathbb{F}}_q, \mathbb{Q}_\ell)$; the Weil bounds on the eigenvalues of Frobenius then show that

$$\sum_{\deg f=n} \mu(D) = O(q^{n/2+\epsilon}).$$

Indeed, when the eigenvalues of Frobenius are distinct, the ϵ can be removed; all this and more is discussed in Byungchul Cha's 2011 preprint "The summatory function of the Möbius function in function fields."

What about combinations of shifts of Möbius? The governing conjecture is due to Chowla:

Conjecture 13. (Chowla) Fix a positive integer m and fix integers $a_1, \ldots, a_m$ and integers $e_1, \ldots e_m$, not all even. Then

$$\sum_{n \leqslant N} \mu(n+a_1)^{e_1} \ldots \mu(n+a_m)^{e_m} = O(N).$$

Note that we might as well take the e_i to be 1 or 2 since $\mu^3 = \mu$.

Exercise 14. Why did we impose the condition that not all the e_i are even?

This conjecture plays a central role in contemporary questions in number theory and dynamics. An argument of Sarnak (explained in Tao's blog post) shows that the Chowla conjecture would imply that a wide variety of interesting arithmetic sequences are "uncorrelated with Möbius" – see Sarnak's lecture notes "Three lectures on the Möbius function." So Chowla can be seen as a sort of "master conjecture" asserting that it is roughly safe to think of the Möbius function as a random sign.

2.1. Function field Chowla and geometric Chowla.

Almost nothing is known about the Chowla conjecture. It would be a major advance if one could show

$$\sum_{n \leqslant N} \mu(n)\mu(n+1) = o(N)$$

With this in mind, let us try to understand how to translate this special case of Chowla from $\mathbb{Z}$ to $\mathbb{F}_q[T]$, and from there to geometry. For the present discussion, q is an odd prime power.

Problem 15. How does all this play out if $\mathbb{F}_q$ is a finite field of characteristic 2?

Consider the sum

$$\sum_{\deg(f)=n} \mu(f)\mu(f+1)$$

where f ranges over monic polynomials of degree n in $\mathbb{F}_q[t]$.

As mentioned above, $\mu(f)$ has a very nice interpretation in terms of the discriminant $\Delta(f)$; namely, $(-1)^n\mu(f)$ is 1 if $\Delta(f)$ is a square in $\mathbb{F}_q^\times$, and -1 if $\Delta(f)$ is a non-square in $\mathbb{F}_q^\times$. (By definition, $\mu(f) = 0$ if and only if $\Delta(f) = 0$.) It follows that

$$\mu(f)\mu(f+1) + 1 = \text{number of square roots of } \Delta(f)\Delta(f+1) \text{ in } \mathbb{F}_q$$

Now define X_n to be the variety cut out of the afffine $(n+1)$-space with coordinates $y, a_1, \ldots, a_n$ by the equation

$$y^2 = \Delta(x^n + a_1x^{n-1} + \ldots + a_n)\Delta(x^n + a_1x^{n-1} + \ldots + a_n + 1).$$

The map sending $(y, a_1, \ldots, a_n)$ to $(a_1, \ldots, a_n)$ expresses Y_n as a double cover of $\mathbb{A}^n$; it is branched precisely at the locus where $\Delta(f)\Delta(f+1) = 0$, i.e. the locus where either f or $f+1$ has a double root. From the above discussion it's clear that

$$|Y_n(\mathbb{F}_q)| - q^n = \sum_{\deg(f)=n} \mu(f)\mu(f+1)$$

So the statement to be proved becomes

$$|Y_n(\mathbb{F}_q)| = q^n + o(q^n).$$

or

$$\lim_{n\to\infty} q^{-n} Y_n(\mathbb{F}_q) = 1.$$

This is very promising – this seems the same sort of thing that we worked out for Conf^n in the previous section! And I've even given you a tool – show that the cohomology of Y_n stabilizes as n grows. (We would also need to obtain subexponential Betti numbers in order to control the contribution of the low-degree cohomology.) If all that worked, you would find that

$$\lim_{n\to\infty} q^{-n}|Y_n(\mathbb{F}_q)| = \sum_{i=0}^{\infty} (-1)^i \mathrm{Tr}\,\mathrm{Frob}\,|H^i_{et;c}(Y_\infty/\overline{\mathbb{F}}_q; \mathbb{Q}_\ell)$$

and you would want the infinite sum on the right-hand side to be equal to 1. What would make this so? The most natural guess is that the top cohomology contributes the 1, and all the other cohomology vanishes. So one might formulate a *geometric Chowla conjecture* as follows:

Conjecture 16 (Geometric Chowla)**.** Let k be an algebraically closed field. For all sufficiently large n, the variety Y_n/k is irreducible, and there is a constant $\alpha > 0$ such that

$$H^{2n-i}_{et,c}(Y_n, \mathbb{Q}_\ell) = 0$$

for all $i < \alpha n$.

Remark 4. This should not really be called a "conjecture," since we have no good reason to believe it.

Remark 5. The condition that Y_n is irreducible, when $k = q^n$, implies that $H^{2i}_{et,c}$ is a 1-dimensional space on which Frobenius acts as q^n; thus the stable H^0 contributes 1 to the limit, while all the other cohomology groups contribute 0, yielding the desired value of 1.

Remark 6. Conjecture 16 as written would not imply the Chowla conjecture over $\mathbb{F}_q[t]$; a subexponential Betti bound in the sense of Proposition 9 would also be required.

Back in the real world, we don't know that the cohomology groups $H_i(Y_n)$ stabilize as $n \to \infty$, let alone what they stabilize to.

However, there is still something that can be done! Remember, to talk about *the* rational function field version of Chowla's conjecture is misleading; there are infiniitely many different rational function fields, one for each finite field $\mathbb{F}_q$. So q is another knob we can turn. Instead of asking about

$$\lim_{n\to\infty} q^{-n}|Y_n(\mathbb{F}_q)|$$

we can ask about

$$\lim_{q\to\infty} q^{-n}|Y_n(\mathbb{F}_q)|$$

and *only then* let n get large. In other words, we can make a weaker, but still very relevant conjecture of Chowla type:

Question 17 (Large q Chowla)**.** Is $\lim_{n\to\infty}\lim_{q\to\infty} q^{-n}|Y_n(\mathbb{F}_q)| = 1$?

Why is the case of large q interesting? Because letting q get large while everything else stays fixed puts a very large hammer in our hands. It follows from the Weil bounds, proven by Deligne (see e.g. Theorem 1 of "La conjecture de Weil: II") that the eigenvalues of Frobenius acting on $H^i_{et;c}(Y_\infty/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell)$ are algebraic numbers with complex absolute value at most $q^{i/2}$. So the contribution of the trace of Frobenius on all the cohomology groups *other* than $H^{2i}_{et;c}$ is bounded above by

$$Bq^{n-1/2}$$

where B, a constant, is the sum of the ith Betti number as i ranges from 0 to $2n-1$.

On the other hand, the top cohomology group $H^{2n}_{et;c}(Y_\infty/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell)$ is the $\mathbb{Q}_\ell$-vector space spanned by the irreducible components of $Y_n/\overline{\mathbb{F}}_q$, with the Frobenius action being the permutation representation induced by the action of Frobenius on the irreducible components, composed with multiplication by q^n. It follows that

$$\lim_{q\to\infty} q^{-n}|Y_n(\mathbb{F}_q)| = \text{number of irreducible components of } Y_n/\overline{\mathbb{F}}_q.$$

Remark 7. There is a slight subtlety here. The trace of Frob_q on $H^{2n}_{et;c}(Y_\infty/\overline{\mathbb{F}}_q;\mathbb{Q}_\ell)$ isn't the number of irreducible geometric components, but rather the number of those components which are rational over $\mathbb{F}_q$. So when we say $q\to\infty$ above, we have in mind that q is getting large and at the same time $q-1$ is getting sufficiently divisible that all geometric components of Y_n are $\mathbb{F}_q$-rational for all sufficiently large q.

This instantiates the slogan at the beginning of this section: if we want to prove the function field version of the Chowla conjecture, we need to know something about all the cohomology groups of the relevant moduli spaces Y_n. But if we aim at the more modest goal of verifying that the Chowla conjecture is true "in the large q limit", it suffices to compute the irreducible components of Y_n.

In particular, in order to answer Question 17 in the affirmative, it suffices to show that Y_n is geometrically irreducible. Equivalently, we need to know that the function $\Delta(f)\Delta(f+1)$, considered as a function in the ring $\mathbb{F}_q[a_1,\ldots,a_n]$, is *not a perfect square.*

This is precisely what Carmon and Rudnick prove. In fact, by being a little more careful, they are able to prove the analogous statement for the general Chowla conjecture, and they obtain explicit bounds for the Chowla sum. They prove:

Proposition 18 (Carmon-Rudnick)**.** *Let $\mathbb{F}_q$ be a finite field of odd characteristc and let $a_1,\ldots,a_m$ be distinct polynomials in $\mathbb{F}_q[t]$. Then*

$$\sum_{\deg f=n} \mu(f+a_1)^{e_1}\ldots\mu(f+a_m)^{e_m} \leqslant 2mnq^{n-1/2} + 3rn^2q^{n-1}.$$

This gives the desired asymptotic for the product of shifted Möbius functions as long as q is large relative to n.

Problem 19. Compute the value of the Chowla sum $\sum_{\deg(f)=n} \mu(f)\mu(f+1)$ for $n = 3$ and various values of q (this is not hard in Sage.) From this, can we make good guesses about the cohomology groups of Y_n – the Betti numbers of this variety over $\bar{\mathbb{Q}}$, and even the Galois representations appearing in its etale cohomology? This would be a good problem for the working groups to think about.

2.2. **Geometric twin primes.** There is a very similar story one could tell about the twin prime conjecture. One version of the conjecture holds that if $T(q, n)$ is the number of monic polynomials f in $\mathbb{F}_q(t)$ such that f and $f + 1$ are both irreducible, then we should have an asymptotic

$$\lim_{n\to\infty} n^2 q^{-n} T(q, n) \to 1.$$

Just as in the case of Chowla, one can interpret this as a question about the etale cohomology groups of a certain variety; but now the variety changes. Let Z_n be the variety parametrizing n-tuples $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ subject to the relation

$$\prod_i (t - x_i) - \prod_i (t - y_j) = 1$$

Then Z_n carries an action of $S_n \times S_n$ by permutation of the x_i and the y_i, and the quotient $Z_n/(S_n \times S_n)$ is naturally identified with $\mathbb{A}^n$ (use as coordinates the elementary symmetric functions in the x_i.) In this case, the Grothendieck-Lefschetz formula expresses $T(q, n)$ in terms of the étale cohomology of Z_n, considered as $\mathbb{Q}_\ell[S_n \times S_n]$-module, along the lines described in Church-Ellenberg-Farb, "Representation stability in cohomology and asymptotics for families of varieties over finite fields."

In particular, the large q limit case of the twin prime conjecture, which follows from a more refined 2008 theorem by Paul Pollock ("A polynomial analogue of the twin prime conjecture") amounts to the assertion that Z_n is geometrically irreducible.

Problem 20. Write down an assertion about the cohomology of Z_n that deserves to be called "the geometric twin primes conjecture."

In fact, one can go further: Lior Barry-Soroker, in a 2012 paper, "Hardy-Littlewood tuple conjecture over large finite fields," proves the much more general statement that for any set of polynomials $a_1, \ldots, a_r$ of degree less than n, the set of $f \in \mathbb{F}_q[t]$ of degree n such that $f + a_1, \ldots, f + a_r$ are all irreducible has approximately the expected density, n^{-r} – as long as q is allowed to go to ∞ with n fixed. Note that in Barry-Soroker's result, the error term has implicit constants depending only on n, r; in particular, the bound is uniform in the a_i and indeed the a_i can be chosen separately for each q!

2.3. **Some problems are boring when q is large.** How many monic degree-n polynomials f are there such that f and $f + 1$ are both squarefree? To answer this question is to count the number of points on the complement of the vanishing locus $Z \subset \mathbb{A}^n$ of $\Delta(f)\Delta(f + 1)$. But $|Z(\mathbb{F}_q)|$ is bounded above by a constant multiple of q^{n-1}, so

$$\lim_{q\to\infty} q^{-n}(\mathbb{A}^n \backslash Z)(\mathbb{F}_q) = 1.$$

All this tells us is that "over $\overline{\mathbb{F}}_q$, both f and $f + 1$ are squarefree 100% of the time," which is more or less obvious given that each one of them is squarefree 100% of the time. This is a case where $\lim_{n\to\infty} q^{-n}(\mathbb{A}^n \backslash Z)(\mathbb{F}_q)$ is *not* independent of q, but rather carries some interesting information in the terms which are lower-order in q – this information, of course, is lost in the large q limit. There are standard analytic number theory proofs which show that, over $\mathbb{Z}$, the proportion of integers n such that both n and $n+1$ are squarefree is

$$\prod_p (1 - 2/p^2)$$

and I presume, though I have not checked, that these proofs work just as well over $\mathbb{F}_q(t)$, giving an asymptotic of the same form but with the product over irreducible monic polynomials instead of prime numbers.

3. Cohen-Lenstra heuristics, big monodromy, random matrices

The lesson of this section:

> In many problems of interest, the computation of irreducible components necessary to answer an "arithmetic statistics in the large q limit" problem is naturally cast as a problem of computing a monodromy group. What's more, computing these monodromy groups is not just a way of ratifying existing conjectures, but provides a machine for producing *new* conjectures which are to some extent backed by geometry.

In our discussion of the Chowla conjecture, we constructed a variety Y_n which was a double cover of $\mathbb{A}^n$ branched along a closed subvariety: namely, the vanishing locus of $\Delta(f)\Delta(f+1)$. Denote the complement of this closed subvariety by U_n, and the restriction $Y_n \times_{\mathbb{A}^n} U_n$ by V_n. Then $V_n \to U_n$ is an etale double cover, which is to say it is described by a homomorphism

$$\phi : \pi_1(U_n) \to \mathbb{Z}/2\mathbb{Z}$$

The question raised in the previous section is

- whether Y_n is irreducible, or equivalently;
- whether V_n is connected, or equivalenty;
- whether the image of ϕ – the *monodromy* group of the cover – is the whole group $\mathbb{Z}/2\mathbb{Z}$.

In other words, what we are proving when we prove that $\Delta(f)\Delta(f+1)$ is not a perfect square can be thought of as a *big monodromy* result. It might be better to say an *as-big-as-possible monodromy* result, since $\mathbb{Z}/2\mathbb{Z}$ is not a very big group.

We now turn our attention to the Cohen-Lenstra heuristics.

Conjecture 21 (Cohen-Lenstra)**.** Let p an odd prime and $E_{r,\ell,N}$ be the expected value, as d ranges over squarefree integers in $[N, 2N]$, of

$$\mathrm{Surj}(\mathrm{Cl}(\mathbb{Q}(\sqrt{-d})), (\mathbb{Z}/\ell\mathbb{Z})^r).$$

Then $E_{r,\ell,N}$ approaches a limit as $N \to \infty$, and this limit is 1. For instance, this means

$$\mathbb{E}(\mathrm{Cl}(\mathbb{Q}(\sqrt{-d}))[\ell]) = 2.$$

Remark 8. These conjectures are part of a rather general family of heuristics, some due to Cohen and Lenstra, some due to Cohen-Lenstra-Martinet. There are Cohen-Lenstra heuristics about real quadratic fields, too, and about the class groups of fields of any degree, not just quadratic fields, and about the whole p-primary part of the class group, not just the p-torsion, and so on. As with Chowla, we will restrict our attention to the subcase of the problem above, which is already rich enough to support the geometric point I'm trying to make

Almost nothing is known about Conjecture 21, apart from a theorem of Davenport and Heilbronn which says the conjecture holds in case $r = 1, p = 3$.

We note in passing that the conjecture of Poonen and Rains, which will also be discussed in this workshop, has a similar fom:

Heuristic: (Poonen-Rains) Let $E/\mathbb{Q}$ be an elliptic curve, let p be an odd prime, and let $E_{r,p,X}$ be the expected value, as d ranges over squarefree integers in $[-X, -2X]$ of

$$\mathrm{Surj}(\mathrm{Sel}_p(E_d), (\mathbb{Z}/pZ)^r).$$

Then $E_{r,p,X}$ approaches a limit as $X \to \infty$, and this limit is $p^{(1/2)k(k+1)}$.

In particular, this means that the average size of Sel_p is $p + 1$.

3.1. **Function-field Cohen-Lenstra.** Let $\mathbb{F}_q$ be a finite field of characteristic prime to ℓ. When we pass from $\mathbb{Q}$ to $\mathbb{F}_q(t)$, the analogue of a quadratic field is a hyperelliptic curve. To be *imaginary* is to be ramified at ∞; so it is natural to declare the quadratic imaginary function fields to be those which are ramified over $\infty \in \mathbb{P}^1$, which is to say, those with affine model $y^2 = f(x)$ with $f(x)$ of *odd* degree. Call this curve C_f. We write ∞_f for the unique point of C_f over ∞. Then we have analogies

$\mathcal{O}$, the ring of integers of $\mathbb{Q}(\sqrt{-d})$	the affine curve $U = C_f - \infty_f$
ideal of $\mathcal{O}$	divisor on U
principal ideal of $\mathcal{O}$	principal divisor on U
class group of $\mathcal{O}$	$\mathrm{Jac}(C_f)(\mathbb{F}_q)$

The last line requires a little explanation. The group of divisors mod principal divisors is the Picard group $\mathrm{Pic}(C_f)(\mathbb{F}_q)$, which sits in an exact sequence

$$0 \to \mathrm{Jac}(C_f)(\mathbb{F}_q) \to \mathrm{Pic}(C_f)(\mathbb{F}_q) \to \mathbb{Z} \to 0$$

where the last map is the degree map. So the map $D \mapsto D - (\deg D)[\infty_f]$ induces a bijection from $\mathrm{Div}(U)$ to $\mathrm{Div}^0(C)$, the group of degree-0 divisors on C, which descends to an isomorphism from the class group of U to $\mathrm{Jac}(C_f)$.

Remark 9. In the case of a hyperelliptic curve which was split at ∞, which is to say, $y^2 = f(x)$ with f of *even* degree, there are two points ∞_f and ∞'_f of C_f over ∞. In this case

$$\mathrm{Pic}(U) = \mathrm{Pic}(C_f)/\langle \infty_f, \infty'_f \rangle = \mathrm{Jac}(C_f)/\langle \infty_f - \infty'_f \rangle$$

So our model for the class group of a real quadratic field is not quite described by an abelian group A, but rather by an abelian group A with a specified element a, the class group being the quotient A/a . The Cohen-Lenstra heuristics say that the p-part in the class group of a real quadratic field should be modelled by A/a where A and a are chosen *uniformly at random* in a suitable sense.

Given the above, it is natural to define a space $\mathrm{Conf}^n(\ell)$ to be the moduli space of configurations *with ℓ-level structure*, i.e. pairs (f, P) where f is a monic squarefree polynomial of degree n and P is a nonzero ℓ-torsion point on the Jacobian of C_f, the curve with equation $y^2 = f(x)$

Now define $E_{q,\ell,r,n}$ to be the expected value

$$\mathbb{E}_f |\operatorname{Surj}(\operatorname{Jac}(C_f)(\mathbb{F}_q), (\mathbb{Z}/\ell\mathbb{Z})^r)|$$

as f ranges over monic squarefree polynomials of degree n.

Then $E_{q,\ell,1,n}$ is the expected number of nonzero $\mathbb{F}_q$-rational ℓ-torsion points on the Jacobian of C_f, where f is chosen randomly in $\operatorname{Conf}^n(\mathbb{F}_q)$; in other words,

$$E_{q,\ell,1,n} = \frac{|\operatorname{Conf}^n(\ell)(\mathbb{F}_q)|}{|\operatorname{Conf}^n(\mathbb{F}_q)|}.$$

And the function field Cohen-Lenstra conjecture then demands that

$$\lim_{n\to\infty} \frac{|\operatorname{Conf}^n(\ell)(\mathbb{F}_q)|}{|\operatorname{Conf}^n(\mathbb{F}_q)|} = 1.$$

3.2. **Function-field Cohen-Lenstra in the large q limit and monodromy.** The natural map $\operatorname{Conf}^n(\ell) \to \operatorname{Conf}^n$ which sends (f, P) to f is an etale cover; the fiber over f is $\operatorname{Jac}(f)[\ell]$, and the cover is described by specifying an action of $\pi_1(\operatorname{Conf}^n)$ on $\operatorname{Jac}(f)[\ell]$. But note that by π_1 here I mean the étale fundamental group; this fits into an exact sequence

$$1 \to \pi_1(\operatorname{Conf}^n_{\overline{\mathbb{F}}_q}) \to \pi_1(\operatorname{Conf}^n) \to \operatorname{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q) \to 1.$$

(You should complain at this point that I have not specified a basepoint for my fundamental groups; I just ask you to trust me that this is one of those times when we can get away with suppressing that choice.)

The action of $\pi_1(\operatorname{Conf}^n)$ on $\operatorname{Jac}(f)[\ell]$ is linear, so its image in the permutation group of the fiber actually lies in $\operatorname{GL}(\operatorname{Jac}(f)[\ell])$; what's more, it preserves the (symplectic) Weil pairing on $\operatorname{Jac}(f)[\ell]$ up to scaling, so the action of $\pi_1(\operatorname{Conf}^n)$ actually takes image in the generalized symplectic group $\operatorname{GSp}(\operatorname{Jac}(f)[\ell])$. What's more, the geometric fundamental group $\pi_1(\operatorname{Conf}^n_{\overline{\mathbb{F}}_q})$ preserves the Weil pairing on the nose, so its image lies in $\operatorname{Sp}(\operatorname{Jac}(f)[\ell])$.

Write $\Gamma \subset \operatorname{GSp}(\operatorname{Jac}(f)[\ell])$ for the image of $\pi_1(\operatorname{Conf}^n)$, and $\Gamma_0 \subset \operatorname{Sp}(\operatorname{Jac}(f)[\ell])$ for the image of $\pi_1(\operatorname{Conf}^n/\overline{\mathbb{F}}_q)$. The former group is called the *monodromy group* of the cover, and the latter the *geometric monodromy group.*

At this point, there is no harm in choosing a basis for $\operatorname{Jac}(C_f)[\ell]$ and thinking of Γ as a subgroup of the standard generalized symplectic grouip $\operatorname{GSp}_{2g}(\mathbb{Z}/\ell\mathbb{Z})$. (Here g is the genus of C_f, so $g = (1/2)(n-1)$.) Note that $\operatorname{GSp}_{2g}(\mathbb{Z}/\ell\mathbb{Z})/\operatorname{Sp}_{2g}(\mathbb{Z}/\ell\mathbb{Z}) \cong (\mathbb{Z}/\ell\mathbb{Z})^\times$; we denote by $[q]$ the class in $\operatorname{GSp}_{2g}(\mathbb{Z}/\ell\mathbb{Z})/\operatorname{Sp}_{2g}(\mathbb{Z}/\ell\mathbb{Z})$ corresponding to $q \in (\mathbb{Z}/\ell\mathbb{Z})^\times$. The fact that $\operatorname{Frob} \in \operatorname{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$ scales the Weil pairing by q implies that the image of Γ in $\operatorname{GSp}_{2g}(\mathbb{Z}/\ell\mathbb{Z})/\operatorname{Sp}_{2g}(\mathbb{Z}/\ell\mathbb{Z})$ is $[q]^{\mathbb{Z}}$. So we have an exact sequence

$$1 \to \Gamma_0 \to \Gamma \to [q]^{\mathbb{Z}} \to 1$$

where the third term is a finite cyclic group whose order is the order of q in $(Z/\ell\mathbb{Z})^\times$.

Now standard facts about étale covers and the étale fundamental group tell us:

- The geometric components of $\operatorname{Conf}^n(\ell)$ are in bijection with the orbits of Γ_0 on $(\mathbb{Z}/\ell\mathbb{Z})^{2g}$.
- The action of Frob on the geometric components is given by the action of $[q]$ on the orbits of Γ.

Write Γ^q for the coset of Γ_0 in Γ which maps to $[q]$ in $[q]^{\mathbb{Z}}$.

Proposition 22. *The following six numbers are the same:*

1. *The average size* $|\operatorname{coker}(g-1)| - 1$*, where g is a random element of* Γ^q*;*
2. *The average number of nonzero fixed points of a random element of* Γ^q *acting on* $(\mathbb{Z}/\ell\mathbb{Z})^{2g}$*;*
3. *The number of* Γ_0*-orbits on nonzero elements of* $(\mathbb{Z}/\ell\mathbb{Z})^{2g}$ *which are fixed by* $[q]$*;*
4. *The number of geometrically irreducible components of* $\operatorname{Conf}^n(\ell)$ *which are defined over* $\mathbb{F}_q$*;*
5. $\lim_{q\to\infty} \frac{|\operatorname{Conf}^n(\ell)(\mathbb{F}_q)|}{|\operatorname{Conf}^n(\mathbb{F}_q)|}$*;*
6. $\lim_{q\to\infty} E_{q,\ell,1,n}$*.*

Proof. Most of the implications are straightforward from definitions. The passage from the second number to the third is a slightly tricked-out version of Burnside's Lemma. The passage from the fourth number to the fifth follows from the Weil bounds, as in our discussion of large q Chowla. □

Proposition 23. *When ℓ is an odd prime all six numbers above are equal to* 1*.*

Proof. It suffices to prove this for the number of Γ_0-orbits on $(\mathbb{Z}/\ell\mathbb{Z})^{2g}$ which are fixed by $[q]$. In order to compute this number, we must of course know what Γ_0 is. In other words, we need to know the monodromy group of the cover. Fortunately for us, this is known! It is a theorem of Jiu-Kang Yu – proved by him expressly with this application to "large q Cohen-Lenstra" in mind – that Γ_0 is the whole of $\operatorname{Sp}_{2g}(\mathbb{Z}/\ell\mathbb{Z})$. (Yu never published his paper, and proofs later appeared in papers of Hall and of Achter-Pries.)

Given this big monodromy theorem, we know that Γ_0 acts transitively on the nonzero elements of $(\mathbb{Z}/\ell\mathbb{Z})^{2g}$. This proves the desired statement. □

We conclude that

$$\lim_{q\to\infty} E_{q,\ell,1,n}$$

which is to say that the Cohen-Lenstra conjecture holds in the large q limit when $r = 1$.

Exercise 24. When $\ell = 2$, show that the geometric monodromy group Γ_0 of $\operatorname{Conf}_n(2) \to \operatorname{Conf}_n$ is much smaller than $\operatorname{Sp}_{2g}(\mathbb{Z}/2\mathbb{Z})$. What is it? What is Γ? What does this computation tell us about $\lim_{q\to\infty} \frac{|\operatorname{Conf}^n(\ell)(\mathbb{F}_q)|}{|\operatorname{Conf}^n(\mathbb{F}_q)|}$? Does this match what you know about the 2-part of the class group of imaginary quadratic fields?

3.3. **Monodromy tells us when Cohen-Lenstra needs modification.**

In fact, the same argument easily handles general r. We can define $\operatorname{Conf}^n(\ell, r)$ to be the space parametrizing pairs (f, ϕ) where $\phi : (Z/\ell\mathbb{Z}) \hookrightarrow \operatorname{Jac}(C_f)[\ell]$ is an injection. This, too, is an etale cover of Conf^n, corresponding to the action of $\pi_1(\operatorname{Conf}^n)$ on the injections from $(\mathbb{Z}/\ell\mathbb{Z})^r$ to $(\mathbb{Z}/\ell\mathbb{Z})^{2g}$. Since the action of $\pi_1(\operatorname{Conf}^n)$ on $\operatorname{Jac}(C_f)[\ell]$ factors through Γ, one has by the same argument as above that

$$\lim_{q\to\infty} E_{q,\ell,r,n} = \text{the number of } [q]\text{-fixed orbits of } \Gamma_0 \text{ on } \operatorname{Inj}((\mathbb{Z}/\ell\mathbb{Z})^r, (\mathbb{Z}/\ell\mathbb{Z})^{2g}).$$

Now the symplectic group does not act transitively on the injections from $(\mathbb{Z}/\ell\mathbb{Z})^r$ to $(\mathbb{Z}/\ell\mathbb{Z})^{2g}$ once $r > 1$. Suppose $\phi : (\mathbb{Z}/\ell\mathbb{Z})^r \to (\mathbb{Z}/\ell\mathbb{Z})^{2g}$ is an injection and ω the symplectic form on $(\mathbb{Z}/\ell\mathbb{Z})^r$ obtained by pulling back the Weil pairing along ϕ. Then

ω is an invariant of the Γ_0-orbit of ϕ; in fact, by the symplectic Witt theorem, it is the *only* invariant. Moreover, the action of $[q]$ on these orbits multiplies ω by $[q]$.

Suppose q is not congruent to 1 mod ℓ. Then if the orbit of ϕ is fixed by $[q]$, we must have $\omega = 0$; now Γ_0 acts transitively on the isotropic subspaces. In other words, there are $\ell^{\binom{r}{2}}$ components of $\mathrm{Conf}^n(\ell, r)$, one for each antisymmetric quadratic form, but only one of them is defined over $\mathbb{F}_q$, and this is the only one that can contribute $\mathbb{F}_q$-rational points.

If q is congruent to 1 mod ℓ, the situation is totally different. Then $[q]$ is the identity, and all $\ell^{\binom{r}{2}}$ components are defined over $\mathbb{F}_q$.

So we must be very careful about the way we let q go to infinity, as discussed in Remark 7. If we let q get large in a sequence of prime powers which are never 1 mod ℓ, we have

$$\lim_{q\to\infty} E_{q,\ell,r,n} = 1 \tag{12}$$

which agrees with the Cohen-Lenstra heuristic.

But if we let q get large in a sequence of prime powers which are 1 mod ℓ, we get instead

$$\lim_{q\to\infty} E_{q,\ell,r,n} = \ell^{\binom{r}{2}}. \tag{13}$$

The two answers agree for $r = 1$ but are very different for large r.

One way to think of this is that, when $q = 1 \pmod{\ell}$, the monodromy group Γ is *smaller*; namely, it is equal to Γ_0.

So what's going on? One answer agrees with Cohen-Lenstra, the other is quite different. Is this evidence *for* Cohen-Lenstra, or *against* it? In a sense, it is both. We are somehow thinking of the large q limit of $\mathbb{F}_q(t)$ as "modeling" a number field. What does it mean for q to be 1 mod ℓ? It means precisely that $\mathbb{F}_q(t)$ contains an ℓ-th root of unity. And q not congruent to 1 mod ℓ means $\mathbb{F}_q(t)$ does not have an ℓth root of unity. Since $\mathbb{Q}$ doesn't have an ℓ-th root of unity (remember, ℓ is odd!) the first model is a more appropriate one for $\mathbb{Q}$; so with a sigh of relief we can say that the monodromy computation supports Cohen-Lenstra.

But what about the ℓ-torsion in class groups of $K(\sqrt{-d})$, where K is another number field? This computation suggests that the behavior may be quite different depending on whether K contains ζ_ℓ. And this indeed appears to be the case! More than twenty years after the initial publication of the heuristics (in their general form due to Cohen, Lenstra, and Martinet) Günter Malle noticed computationally that the conjectures seemed to be way off in some cases: for example, with respect to the 3-torsion in the class group of quadratic extensions of $\mathbb{Q}(\zeta_3)$. Malle recognized that the presence of extra roots of unity was the deciding factor, and proposed a modified conjecture that fit the data much better.

Another way to make predictions about the behavior of the ℓ-part of the class group would be to declare that the average number of injections from $(\mathbb{Z}/\ell\mathbb{Z})^r$ to $\mathrm{Cl}(K(\sqrt{d}))$ should be $\ell^{\binom{r}{2}}$ when K is a field containing ζ_ℓ. Derek Garton, in his 2012 thesis, shows that this reproduces Malle's modified Cohen-Lenstra heuristic exactly! What's more, he explains in many cases how the full suite of Cohen-Lenstra predictions (which involve the full ℓ-primary part of the class group, not just the ℓ-torsion) ought to be modified in the presence of ℓ-power roots of unity, at least if we expect the arithmetic to obey the laws suggested by the geometry.

3.4. Geometric Cohen-Lenstra.

The connectedness of $\mathrm{Conf}^n(\ell)$ told us that, in the large q limit, the average size of the ℓ-part of the class group is 2. (As we put it above, the average number of nonzero elements is 1.)

But in the end, we don't want to take the large q limit; in order to prove what should rightfully be called the Cohen-Lenstra conjecture over $\mathbb{F}_q(t)$, we need to prove (among other things!) that

$$\lim_{n\to\infty} \frac{|\mathrm{Conf}^n(\ell)(\mathbb{F}_q)|}{|\mathrm{Conf}^n(\mathbb{F}_q)|} = 1. \tag{14}$$

What would we mean by "geometric Cohen-Lenstra"? It's not enough to compute $H^0(\mathrm{Conf}^n(\ell))$, which, as we've seen, amounts to computing a monodromy group. We need to go deeper into the cohomology of $\mathrm{Conf}^n(\ell)$. In order to show the existence of the limit above, it suffices, by Proposition 9 to prove:

Theorem 25. *It is the case that:*

- *(Homological stabilization) There is a constant $\alpha > 0$ (depending on ℓ) such that, for all $i \leqslant \alpha n$, there is an isomorphism*
$$H^i_{et}(\mathrm{Conf}^n(\ell)/\bar{\mathbb{Q}}, \mathbb{Q}_\ell) \cong H^i_{et}(\mathrm{Conf}^{n+1}(\ell)/\bar{\mathbb{Q}}, \mathbb{Q}_\ell)$$
which commutes with the action of Galois on either side.
- *(Subexponential Betti numbers) There is a constant C such that* $\dim H^i_{et}(\mathrm{Conf}_n(\ell)/\bar{\mathbb{Q}}, \mathbb{Q}_\ell)$ *is at most C^n for all i.*
- *(Comparison) For all p not dividing N and all i,*
$$\dim\dim H^i_{et}(\mathrm{Conf}^n(\ell)/\bar{\mathbb{Q}}, \mathbb{Q}_\ell) = \dim H^i_{et}(\mathrm{Conf}^n(\ell)/\bar{\mathbb{F}}_p, \mathbb{Q}_\ell)$$

But we believe that the limit is 1! And we've already shown in the previous section, when we computed the large q limit, that the contribution of the stable H^0 is 1. It is thus natural to hope that all the other terms contribute 0.

Conjecture 26 (Geometric Cohen-Lenstra)**.** There is a constant $\alpha > 0$ (depending on ℓ) such that, for all $1 < i \leqslant \alpha n$,

$$H^i_{et}(\mathrm{Conf}^n(\ell)/\bar{\mathbb{Q}}, \mathbb{Q}_\ell) = 0$$

Theorem 25 is proved in the paper "Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields," by Ellenberg, Venkatesh, and Westerland. (To be more precise: all of this theorem is proved there except for the Galois equivariance in the homological stabilization; that we figured out how to do later and is not written down yet.)

We presented a proof of the geometric Cohen-Lenstra conjecture – which implies the function field Cohen-Lenstra conjecture (14) in a sequel, but the argument turned out to have a fatal flaw. As I write this , we are still trying to work around it!

The geometric Cohen-Lenstra conjecture asserts that stable cohomology of a certain family of moduli spaces *vanishes* in degree greater than 0. Another way to say this is that the fixed-q limit $\lim_{n\to\infty} \frac{|\mathrm{Conf}^n(\ell)(\mathbb{F}_q)|}{|\mathrm{Conf}^n(\mathbb{F}_q)|}$ agrees with the large q limit, which in turn, as we saw in the previous section, could be computed from statistics of random matrices. Here we find another lesson.

> In many problems, stable cohomology in positive degree can be thought of as measuring a deviation from agreement with a random matrix model. As q grows, this deviation approaches 0 and the statistics being studied approach those of random matrices. If we conjecture that the statistics being studied agree with those of random matrices *without* letting q grow, then one might well predict a vanishing of stable cohomology on the geometric side.

An easier example of the lesson above can be found in the observation made earlier in the notes, that

$$\lim_{n\to\infty} \frac{|\operatorname{PConf}^n/S_{n-1}(\mathbb{F}_q)|}{|\operatorname{Conf}^n(\mathbb{F}_q)|}$$

which is to say, the average number of linear factors of a squarefree polynomial is $(1+q^{-1})^{-1}$. The random matrix model for the action of Frobenius on the roots of f is to take this action to be a random permutation of the roots. The number of linear factors is the number of fixed points of Frobenius in this action, so the random matrix model would predict that a random squarefree polynomial would have 1 linear factor. This is true in the large q limit, but not true for fixed q; and the "reason" for the deviation is that the higher cohomology of $\operatorname{PConf}^n/S_{n-1}$, while it does stabilize, does not stabilize to 0. (What it *does* stabilize to is described very explicitly in Church-Ellenberg-Farb, "Representation stability in cohomology...")

4. More material

The above represents more than enough material to cover in my lectures at the Arizona Winter School, but there is much more geometric analytic number theory one could talk about. Here are some brief ideas about further topics, some of which I may expand into new sections for a later version of these notes; and of course I would be happy to talk about any of these topics at the AWS!

Geometric Malle-Bhargava. Let $G \subset S_d$ be a group and let $N_G(N)$ be the number of degree-d extensions of $\mathbb{Q}$ with Galois group G and discriminant at most N. Malle predicts an asymptotic

$$N_G(N) \sim c(G)N^{a(G)}(\log N)^{b(G)}$$

for specified values of $a(G), b(G)$ and an unspecified constant $c(G)$. Bhargava conjectures a value of $c(G)$ for $G = S_d$ (we remark that $a(S_d) = 1$ and $b(S_d) = 0$.)

The geometric version of this conjecture involves the moduli space of branched G-covers of $\mathbb{P}^1$, which is called a *Hurwitz space* – then the geometric Malle-Bhargava conjecture says that these spaces have the same stable cohomology as Conf^n. Loosely speaking, there is a map from Hurwitz space to configuration space (forget the G-cover, remember only the branch points) which we expect stably to induce an isomorphism on cohomology with rational coefficients; thus, the number of rational points on Hurwitz space is approximately the same as the number of points on configuration space, which in turn say that there is on average 1 G-extension per discriminant. It turns out that for some choices of G this should *not* be expected to hold; there is an extra factor, coming from the Schur multiplier of G. This story is described in Venkatesh-Ellenberg's 2010 paper "Statistics of Number Fields and Function Fields."

Geometric Poonen-Rains. The geometric version of their conjecture on the mod ℓ Selmer group would have to do with the cohomology of a moduli space parametrizing classes (S, α), where S is an elliptic surface and α is a class in $H^2(S, \mathbb{Z}/\ell\mathbb{Z})$. The pairing on $H^2(S, \mathbb{Z}/\ell\mathbb{Z})$ is symmetric, not antisymmetric like the Weil pairing, so the verification of Poonen-Rains in the large q limit case should be obtainable via a suitable big monodromy theorem in the H^2 of a family of elliptic surfaces. A group of us at AIM figured out how to do this in principle and are writing it up.

Geometric Linnik-Duke. Vivek Shende and Jacob Tsimerman have a beautiful 2013 paper, "Equidistribution on the spae of rank two vector bundles," where they study function field analogues of conjectures about equidistribution of Heegner points. The geometric analogues of these conjectures turn out to be conjectures about the cohomology of generalized theta divisors on Jacobians! One interesting feature of their work is that they can prove the homological stabilization and subexponential Betti bounds they need, but for some of the most general theorems they'd like to prove, the comparison step – moving from cohomology in characteristic 0 to cohomology in characteristic p – has not yet been established.

Geometric Batyrev-Manin. The Batyrev-Manin conjectures give asymptotics for the number of rational points on a variety X of height at most B. The geometric analogue concerns the moduli space of maps from $\mathbb{P}^1$ to X, i.e. the space of rational curves on X; more generally, it considers the moduli space of *sections* of a fibration $\mathcal{X}/\mathbb{P}^1$. It turns out that the geometric conjectures which naturally correspond to the Batyrev-Manin heuristics are quite reasonable from the algebro-geometric point of view; they are conjectures of the form "the space of holomorphic rational curves on X of very large degree is topologically very similar to the space of smooth maps from S^2 to $X(\mathbb{C})$." In many cases where the Batyrev-Manin conjecture is known to hold (e.g. when X is a homogenous variety, or a toric variety) the corresponding geometric statement about the space of rational curves on X is also known to hold. A very interesting exception is the case of very low degree hypersurfaces in very high dimension, in which case the number of rational points of bounded height can be computed by the Hardy-Littlewood circle method. Is there a geometric circle method? Is there a motivic circle method? (In connection with the last question, see the work of David Bourqui.)