THE GEOMETRIC DISTRIBUTION OF SELMER GROUPS

AARON LANDESMAN

1. Background on Elliptic curves

Theorem 1.1 (Mordell-Weil)**.** *Let E be an elliptic curve over a global field K (such as $\mathbb{Q}$ or $\mathbb{F}_q(t)$). Then the group of K-rational points $E(K)$ is a finitely generated abelian group.*

For E an elliptic curve over K, write $E(K) \simeq \mathbb{Z}^r \oplus T$ for T a finite group. Then, r is the **rank** of E.

Question 1.2 (Motivating Question)**.** What is the average rank of an elliptic curve?

Conjecture 1.3 (Minimalist Conjecture)**.** The average rank of elliptic curves is $1/2$, and moreover, 50% of curves have rank 0, 50% have rank 1, and 0% have rank more than 1.

2. Average sizes

We'd next like to make sense of a notion of the average size of Selmer groups. Unfortunately, there are infinitely many elliptic curves. To deal with this problem, we introduce a notion of height. There will be only finitely many elliptic curves of a given height, and so we can compute the average size by computing the average for a given height, and then taking a limit.

For the rest of the talk, we'll let $K = \mathbb{F}_q(t)$ and assume $\mathrm{char}(\mathbb{F}_q) > 3$, though everything can be modified to work in arbitrary characteristic, the modification to characteristic 3 being fairly routine, and the modification to characteristic 2 involving more work. All limits over q will be taken over prime powers q with q relatively prime to 6 (and also relatively prime to the index n of the relevant Selmer group).

Definition 2.1. An equation defining an elliptic curve E is in minimal Weierstrass form if it is of the form

$$y^2z = x^3 + A(s,t)xz^2 + B(s,t)z^3$$

where there exists d so that $\deg A(s,t) = 4d, \deg B(s,t) = 6d$ and d is minimal among all such equations cutting out E. The **height** of E is then

$$h(E) := d.$$

Definition 2.2. The **average size** of a function

$$X : \text{ isomorphism classes of elliptic curves over } \mathbb{F}_q(t) \to \mathbb{R}$$

of height up to d is

$$\text{Average}^{\leq d}(\#X/\mathbb{F}_q(t)) := \frac{\sum_{E/\mathbb{F}_q(t), h(E)\leq d} \#X(E)}{\#\{E/\mathbb{F}_q(t) \colon h(E) \leq d\}},$$

where the sum runs over isomorphism classes of elliptic curves $E/\mathbb{F}_q(t)$, having $h(E) \leq d$. We refer to this informally as "the average size of $\#X$ for elliptic curves of height $\leq d$ over $\mathbb{F}_q(t)$."

Conjecture 2.3 (A reformulated minimalist conjecture)**.** Let rk be the function assigning to an elliptic curve its rank. Then, $\lim_{q\to\infty}\lim_{d\to\infty}\text{Average}^{\leq d}\left(\text{rk}/\mathbb{F}_q(t)\right) = 1/2$ and moreover for $\text{rk} = i$ the indicator function assigning to an elliptic curve E the value 1 if $\text{rk}(E) = i$ and 0 otherwise,

- $\lim_{q\to\infty}\lim_{d\to\infty}\text{Average}^{\leq d}\left((\text{rk}=0)/\mathbb{F}_q(t)\right) = 1/2$
- $\lim_{q\to\infty}\lim_{d\to\infty}\text{Average}^{\leq d}\left((\text{rk}=1)/\mathbb{F}_q(t)\right) = 1/2$
- $\lim_{q\to\infty}\lim_{d\to\infty}\text{Average}^{\leq d}\left((\text{rk}\geq 2)/\mathbb{F}_q(t)\right) = 0$

2.1. **Selmer groups.** We'll explain, why, in a certain sense, 50% of elliptic curves have rank 0 and 50% have rank 1 more than 1, in accordance with this minimalist conjecture. The idea is to bound the rank in terms of a related object called the Selmer group.

Let $K = \mathbb{F}_q(t)$, and let v index the closed points of $\mathbb{P}^1_{\mathbb{F}_q}$. For E an elliptic curve over K, the multiplication by n exact sequence induces the sequences on étale cohomology

$$\begin{array}{ccccccccc} 0 & \longrightarrow & E(K)/nE(K) & \longrightarrow & H^1(\operatorname{Spec} K, E[n]) & \longrightarrow & H^1(\operatorname{Spec} K, E)[n] & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & \overset{\alpha}{\searrow} & \downarrow & & \\ 0 & \to & \prod_{v\in\mathbb{P}^1_{\mathbb{F}_q}} E(K_v)/nE_v(K_v) & \to & \prod_v H^1(\operatorname{Spec} K_v, E_v[n]) & \to & \prod_v H^1(\operatorname{Spec} K_v, E_v)[n] & \to & 0. \end{array}$$

Definition 2.4. The n**-Selmer** group of E is

$$\text{Sel}_n(E) := \ker\alpha$$

Conjecture 2.5 (Bhargava–Shankar [BS13, Conjecture 4] and Poonen–Rains [PR12, Conjecture 1.4(b)])**.** When all elliptic curves are ordered by height,

$$\lim_{q\to\infty}\lim_{d\to\infty}\text{Average}^{\leq d}(\text{Sel}_n/\mathbb{F}_q(t)) = \sum_{s|n} s.$$

Remark 2.6.

- An analogous statement over $\mathbb{Q}$ was shown for $n = 2, 3, 4, 5$ by Bhargava and Shankar.
- The upper bound was shown for $n = 3$ over $\mathbb{F}_q(t)$ by de Jong.
- This was shown for $n = 2$ more generally over function fields by Ho, Le Hung, and Ngo.

The above conjecture was generalized to a conjecture regarding all moments of prime order Selmer groups.

Conjecture 2.7 (Poonen–Rains)**.** For ℓ prime and $m \geq 1$,

$$\lim_{q\to\infty}\lim_{d\to\infty}\text{Average}^{\leq d}((\#\text{Sel}_\ell)^m/\mathbb{F}_q(t)) = (1+\ell)(1+\ell^2)\cdots(1+\ell^m).$$

3. The BKLPR distribution

We next want to discuss conjectures predicting the full distribution of Selmer groups, not just the moments.

Remark 3.1. It is worth noting that the moments in this case grow very quickly, and so do not determine the distribution. See Example 4.8 for more.

Let O_{2r} denote the orthogonal group associated to the quadratic form $q_{2r} = x_1x_2 + x_3x_4 + \cdots + x_{2r-1}x_{2r}$ over $\mathbb{F}_\ell$. Choose two random r-dimensional isotropic subspaces V and W. I.e., $q_{2r}|_V = q_{2r}|_W = 0$.

Model 3.1 (Poonen–Rains). The ℓ-Selmer group of an elliptic curve E is modeled as $V \cap W$. So,

$$\operatorname{Prob}(\dim \operatorname{Sel}_\ell^{BKLPR} = \alpha) = \lim_{r\to\infty} \operatorname{Prob}(\dim V \cap W = \alpha).$$

The distribution of the rank of an elliptic curve is modeled as

$$\mathrm{rk}^{BKLPR} = \dim V \cap W \bmod 2$$

Remark 3.2. Bhargava, Kane, Lenstra, Poonen, and Rains have a similar but more sophisticated model applying for composite n.

Conjecture 3.3 (Poonen–Rains)**.** For ℓ a prime, as E ranges over elliptic curves over $\mathbb{F}_q(t)$,

$$\lim_{q\to\infty}\lim_{d\to\infty} \operatorname{Prob}\left(\dim_{\mathbb{F}_\ell} \operatorname{Sel}_\ell(E) = v\right) = \operatorname{Prob}\left(\dim \operatorname{Sel}_\ell^{BKLPR} = v\right) = \left(\prod_{j\geq 0}\frac{1}{1+\ell^{-j}}\right)\left(\prod_{j=1}^{v}\frac{\ell}{\ell^j-1}\right).$$

The second equality is a not too difficult direct calculation of the distribution.

Remark 3.4. Why is it reasonable to model the Selmer group as an intersection of two maximal isotropic subspaces? Looking back at the definition of Selmer groups in terms of the exact sequences from étale cohomology, one can see that it is the intersection

$$H^1(\operatorname{Spec} K, E[n]) \cap \prod_v E(K_v)/nE_v(K_v) \subset \prod_v H^1\left(\operatorname{Spec} K_v, E_v[n]\right).$$

It turns out that these two spaces are isotropic with respect to a certain natural quadratic form on the infinite dimensional $\prod_v H^1\left(\operatorname{Spec} K_v, E_v[n]\right)$.

4. Main results

While counting $\lim_{d\to\infty} \text{Average}^{\leq d}(\#\operatorname{Sel}_n / \mathbb{F}_q(t))$ may be difficult, there is another way we can try to take this average. Instead of taking the large d limit, we can try to take the large q limit. First, we can take a large q limit. That is, we can try to show statements like

$$\lim_{q\to\infty} \text{Average}^{\leq d}(\#\operatorname{Sel}_n / \mathbb{F}_q(t)) = \sum_{s|n} s.$$

Indeed, our main result is that the prior conjectures all hold when we interchange the limits.

Theorem 4.1 (Feng-L-Rains). *For any integer $n \geq 1$, we have*

$$(\mathrm{rk}^{\mathrm{BKLPR}}, \mathrm{Sel}_n^{\mathrm{BKLPR}}) = \lim_{d\to\infty} \left(\limsup_{\substack{q\to\infty \\ \gcd(q,2n)=1}} (\mathrm{rk}, \mathrm{Sel}_n)^d_{\mathbb{F}_q} \right)$$
$$= \lim_{d\to\infty} \left(\liminf_{\substack{q\to\infty \\ \gcd(q,2n)=1}} (\mathrm{rk}, \mathrm{Sel}_n)^d_{\mathbb{F}_q} \right).$$

In other words, the BKLPR model is correct after first taking a large q limit!
Observe that we first take a large q limit, whereas BKLPR first takes a large height limit.

Original Conjecture:

$$\lim_{q\to\infty}\lim_{d\to\infty} \frac{\mathrm{Prob}_{E:h(E)\leq d}(\mathrm{rk}(E) = r, \mathrm{Sel}_n(E) \simeq G)}{\#\{E : h(E) \leq d\}} = \mathrm{Prob}((\mathrm{rk}^{\mathrm{BKLPR}}, \mathrm{Sel}_n^{\mathrm{BKLPR}}) = (r, G)).$$

Limits reversed:

$$\lim_{d\to\infty}\lim_{q\to\infty} \frac{\mathrm{Prob}_{E:h(E)\leq d}(\mathrm{rk}(E) = r, \mathrm{Sel}_n(E) \simeq G)}{\#\{E : h(E) \leq d\}} = \mathrm{Prob}((\mathrm{rk}^{\mathrm{BKLPR}}, \mathrm{Sel}_n^{\mathrm{BKLPR}}) = (r, G)).$$

Remark 4.2. Even though the limits are reversed, our results provide some of the first direct evidence for the connection between the arithmetic of elliptic curves and the complete conjectures of BKLPR.

Corollary 4.3. *In the large q limit, the minimalist conjecture holds. For example, the average rank of elliptic curves is $1/2$ in the large limit. That is,*

$$\lim_{d\to\infty}\lim_{q\to\infty} \mathrm{Average}^{\leq d}\left(\mathrm{rk}\,/\mathbb{F}_q(t)\right) = 1/2$$

Similarly, in the large q limit, the Poonen–Rains conjectures on average sizes, moments, and distributions of Selmer groups hold.

Remark 4.4. An analogous result holds for suitable families of quadratic twists of elliptic curves. Niudun Wang and Sun Woo Park are currently working out the details.

Remark 4.5. One can deduce a more precise version of Theorem 4.1 with estimates on the error terms in the above limits. One may also deduce the same result holds with algebraic rank replaced by analytic rank. Further, one may include the joint distribution of n-torsion in Tate-Shafarevich groups.

Remark 4.6 (The inverse Galois problem)**.** The proof shows that for even $d \geq 2$, and $\ell \nmid d-1$, the orthogonal groups $O_{12d-4}(\mathbb{F}_\ell)$ occurs as a Galois group over $\mathbb{Q}$, which gives some new instances of the inverse Galois problem.

Remark 4.7 (Homological stability)**.** Our computation of the moments in the large q limit can be interpreted as a stabilization in d of the number of components of certain moduli spaces parameterizing Selmer elements of height d elliptic curves. If one could show the

higher homologies stabilize, this would go a long way to proving in their original form, i.e., in the "large height, then large q limit."

In fact, an analogous problem was solved by showing a certain sequence of spaces have homological stability. Namely, in [EVW16], the authors verify the Cohen Lenstra heuristics for imaginary quadratic function fields by constructing a sequence of spaces whose points parameterize n-torsion in class groups, and they show these spaces satisfy homological stability and use this to deduce the Cohen Lenstra heuristics.

We conclude the comments with two fun probability "counterexamples" coming from arithmetic.

Example 4.8 (A distribution not determined by its moments)**.** Consider the distributions

$$\begin{aligned}&\mathrm{Sel}_\ell^{\mathrm{BKLPR}},\\ &(\mathrm{Sel}_\ell^{\mathrm{BKLPR}} | \dim \mathrm{Sel}_\ell^{\mathrm{BKLPR}} \equiv 0 \bmod 2)\end{aligned}$$

with the latter distributions conditioning upon whether the dimension is even.

These distributions are clearly vastly different, because the dimensions have different parities. However, their moments in fact agree! This shows that the distribution of the Selmer group of an elliptic curve is not determined by its moments.

The basic reason for the moments of the first two distributions agreeing is as follows. Via a monodromy perspective (which we will see later in the proof of the main result) the moments can be identified with the number of orbits of a certain group acting on a power of the underlying $\mathbb{F}_\ell$ vector space. Loosely speaking, the moments of the first distribution $\mathrm{Sel}_\ell^{\mathrm{BKLPR}}$ can be identified with the number of orbits for a certain action of an orthogonal group while the moments of $\mathrm{Sel}_\ell^{\mathrm{BKLPR}}$ conditioned on the rank being even can be identified with the number of orbits for the corresponding special orthogonal group. It is not too hard to show these have the same orbits. Essentially one has to produce a reflection which does not permute the orbits of the special orthogonal group.

Remark 4.9 (Ongoing work)**.** It seems that one can remove the restriction that q is prime to $2n$. It also seems that one can also obtain analogous results over higher genus function fields, though we have not yet checked all the details.

In the case that the characteristic of $\mathbb{F}_q$ divides n, an interesting phenomenon occurs in that the large q limit of the higher moments can differ from the moments of the distribution of Selmer groups the large q limit. That is,

$$\begin{aligned}&\lim_{k\to\infty} m\text{th moment of } [\#\mathrm{Sel}_\ell \text{ over } \mathbb{F}_{\ell^k}(t) \text{ of height } d] \neq\\ &m\text{th moment of } [\lim_{k\to\infty} \text{ distribution of } \#\mathrm{Sel}_\ell \text{ of height } d \text{ over } \mathbb{F}_{\ell^k}(t)].\end{aligned}$$

In particular, for $d = 2, m \geq 10$ both sides are finite but not equal, and when $m > 10$ the left hand side is infinite while the right hand side is finite.

That is, taking moments does not commute with taking limits. The reason for this discrepancy is that in characteristic dividing n, the geometric Selmer group can be correspond to points on a positive dimensional variety, (while in characteristic prime to n, it always corresponds to points on a 0 dimensional variety), along a codimension 10 locus of height d elliptic curves. This doesn't effect the distribution much because there are very few such curves, but can drastically effect large q limit of the higher moments.

5. Sketch of the proofs

We could prove the theorem regarding the distribution first, and deduce the theorem regarding average size, but I think it will be pleasing to see the geometry in the other order! So, we'll start by computing the average sizes of Selmer groups, and build up to computing the higher moments and then the full distribution.

5.1. **The Selmer Space.** To approach all of these problems, the key step will be to create a moduli space parameterizing Selmer elements.

For $\mathbb{F}_q$ a finite field, we will construct a space $\mathrm{Sel}^d_{n,\mathbb{F}_q}$ parameterizing pairs (E, X), where E is an elliptic curve over $\mathbb{F}_q(t)$ and X is an n-Selmer element of E with the following two properties

(1) Letting $\mathscr{W}^d_{\mathbb{F}_q}$ denote a parameter space for Weierstrass equations of elliptic curves $E/\mathbb{F}_q(t)$ of height d. There is a projection map

$$\begin{aligned}\pi : \mathrm{Sel}^d_{n,\mathbb{F}_q} &\to \mathscr{W}^d_{\mathbb{F}_q} \\ (E, X) &\mapsto [E].\end{aligned}$$

(2)

$$\mathrm{Sel}_n(E) = \pi^{-1}([E])(\mathbb{F}_q).$$

Now, here is the construction:

Definition 5.1. The moduli space of height d weierstrass models $\mathscr{W}^d_{\mathbb{F}_q}$ is an open in an affine space of dimension $10d + 2$ with coordinates $a_0, \dots, a_{4d}, b_0, \dots, b_{6d}$. The point $[a_0, \dots, a_{4d}, b_0, \dots, b_{6d}]$ corresponds to the elliptic surface defined by $y^2 = x^3 - (\sum_{i=0}^{4d} a_i t^i s^{4d-i})x + (\sum_{i=0}^{6d} b_i t^i s^{6d-i})$. This defines a universal elliptic surface

$$\mathcal{X} \xrightarrow{f} \mathbb{P}^1 \times \mathscr{W}^d_{\mathbb{F}_q} \xrightarrow{g} \mathscr{W}^d_{\mathbb{F}_q}.$$

We next define the Selmer space. As motivation, for most elliptic curves E over $\mathbb{F}_q(t)$, with associated Weierstrass model $f_E : X \to \mathbb{P}^1$, it turns out that $\mathrm{Sel}_n(E) \simeq H^1(\mathbb{P}^1, R^1(f_E)_*\mu_n)$. Therefore, we want to construct a space over $\mathscr{W}^d_{\mathbb{F}_q}$ whose fibers are given by $H^1(\mathbb{P}^1, R^1(f_E)_*\mu_n)$.

Definition 5.2. With notation as above, the n-Selmer space of height d elliptic curves over $\mathbb{F}_q$ is the algebraic space representing the sheaf $R^1g_*\left(R^1f_*\mu_n\right)$.

5.2. **The average sizes of Selmer groups.** Having constructed the Selmer space, we are next ready to prove:

Theorem 5.3. *Let n be a prime. The geometric average size of the n-Selmer group is $n + 1$, i.e.,*

$$\lim_{q\to\infty} \mathrm{Average}^{\leq d}(\#\,\mathrm{Sel}_n\, /\mathbb{F}_q(t)) = n + 1.$$

Similar arguments go through in the non-prime case.

Using our moduli space construction above, we find

$$\mathrm{Average}^{\leq d}(\#\,\mathrm{Sel}_n\, /\mathbb{F}_q(t)) = \frac{\sum_{E/\mathbb{F}_q(t), h(E)\leq d} \#X(E)}{\#\{E/\mathbb{F}_q(t) \colon h(E) \leq d\}} = \frac{\#\mathrm{Sel}^d_{n,\mathbb{F}_q}(\mathbb{F}_q)}{\#\mathscr{W}^d_k(\mathbb{F}_q)}.$$

The space of Weierstrass models is basically an affine space, so, we just want to compute the number of $\mathbb{F}_q$ points of the n-Selmer space. Here is where the large q limit comes in:

Theorem 5.4 (Lang-Weil)**.** *For X a finite type space over $\mathbb{F}_p$ with r geometrically irreducible components, $\lim_{q\to\infty} X(\mathbb{F}_q) = rq^{\dim X} + O(q^{\dim X - 1/2})$.*

By the Lang-Weil estimate, since $\mathscr{W}_k^d$ is geometrically irreducible, this ratio is the number of components of $\mathrm{Sel}_{n,k}^d$. So, it remains to show $\mathrm{Sel}_{n,\mathbb{F}_q}^d$ has $n+1$ components.

5.3. **The monodromy of the Selmer space cover.** It remains to compute the number of components. We do this by a monodromy argument. We have constructed a space $\mathrm{Sel}_{n,\mathbb{F}_q}^d$ whose $\mathbb{F}_q$ points parameterize Selmer elements. Let $\mathscr{W}^{\circ d}_{\mathbb{F}_q} \subset \mathscr{W}^d_{\mathbb{F}_q}$ be the dense open parameterizing smooth Weierstrass models. I.e., elliptic curves of height d whose Weierstrass model is smooth over $\mathbb{F}_q$. Set up the fiber square

$$\begin{array}{ccc} \mathrm{Sel}^{\circ d}_{n,\mathbb{F}_q} & \longrightarrow & \mathrm{Sel}^d_{n,\mathbb{F}_q} \\ \downarrow{\scriptstyle \pi^\circ} & & \downarrow{\scriptstyle \pi} \\ \mathscr{W}^{\circ d}_{\mathbb{F}_q} & \longrightarrow & \mathscr{W}^d_{\mathbb{F}_q}. \end{array}$$

Proposition 5.5. *The resulting map π° is finite étale.*

To compute the number of components of $\mathrm{Sel}_{n,k}^d$, we show that, over a dense open $\mathscr{W}^{\circ d}_k \subset \mathscr{W}^d_k$, $\mathrm{Sel}_{n,k}^d$ is a finite étale cover, with geometric fibers $V_{n,k}^d := (\mathbb{Z}/n\mathbb{Z})^{12d-4}$. Hence, we obtain a monodromy representation (or Galois representation)

$$\rho_{n,k}^d : \pi_1^{\text{ét}}(\mathscr{W}^{\circ d}_k) \to \mathrm{GL}(V_{n,k}^d).$$

Because this respects Poincaré duality, it lands in the orthogonal group $O(Q)$. As n is prime, we want to show this has $n+1$ components. The components are identified with the number of orbits of $\rho_{n,k}^d$ on the underlying vector space $V_{n,k}^d$

If the monodromy were the full orthogonal group, Chevalley's theorem says there are $n+1$ orbits: the 0 orbit and the n level sets of the quadratic form. In fact, some group theory shows index 2 subgroups of the orthogonal group have the same orbits. So, we have reduced to showing:

Theorem 5.6. *The image $\operatorname{im}\rho_{n,\mathbb{F}_q}^d \subset O(Q)$ up to index 2, is $O(Q)$.*

Over $k = \mathbb{C}$, it is shown in [dJF11] (by looking at small loops around a very singular elliptic surface, and examining a corresponding Dynkin diagram) that $\operatorname{im}\rho_{n,\mathbb{C}}^d$ has index at most 2 in $O(Q)$, when $d \geq 2$. Fortunately, this index 2 subgroup still has $n+1$ orbits, and so the space has $n+1$ components, over $\mathbb{C}$.

Transferring to positive characteristic: Since we are interested in components over $\mathbb{F}_p$, we need to show the components do not "come together" when one passes to finite fields. That is, we want to show the monodromy over $\overline{\mathbb{F}}_p$ agrees with that over $\mathbb{C}$. The theory of the tame fundamental group shows this is this case if, when one compactifies the map $\mathrm{Sel}^{\circ d}_{n,k} \to \mathscr{W}^{\circ d}_k$, the ramification orders along divisors in the complement of $\mathscr{W}^{\circ d}_k$ are prime

to p. There are only two divisors in the complement, one is an affine space, and the other corresponds to where two singular fibers come together (the elliptic curve has one place of type I_2 reduction). We want to check traveling around this divisor has order 2.

In fact, by a beautiful relative version of Abhyankar's lemma [R71, Exposé XIII, Prop 5.5], because our moduli space is actually defined over all of $\operatorname{Spec}\mathbb{Z}$, it's enough to show the ramification divisor in the moduli space of elliptic surfaces is generically reduced over $\mathbb{F}_p$. One has explicit equations for this divisor, and can check it is reduced whenever $p \neq 2$.

For a more geometric perspective, one can also directly prove the monodromy is tame by describing the effect of traveling once around in terms of combinatorial monodromy data.

5.4. **Higher moments.** Our next goal is to use the Selmer space and the above monodromy result to compute the higher moments. For example, let's use the above to show

Theorem 5.7. *Fix a prime ℓ and an integer m. For $d \geq 2$ and $> \frac{m+3}{6}$, the mth geometric moment of the ℓ-Selmer group is*

$$\lim_{d\to\infty}\lim_{q\to\infty} \text{Average}^{\leq d}((\#\text{Sel}_\ell)^m/\mathbb{F}_q(t)) = (1+\ell)(1+\ell^2)\cdots(1+\ell^m).$$

Let's concentrate on showing the second moment is $(1+\ell)(1+\ell^2)$. We want to rephrase this about counting points on a moduli space. If $\text{Sel}^d_{\ell,\mathbb{F}_q}$ parameterizes ℓ-Selmer elements, what is a moduli space parameterizing pairs of ℓ-Selmer elements? It is simply the fiber product

$$\text{Sel}^d_{\ell,\mathbb{F}_q} \times_{\mathscr{W}^d_{\mathbb{F}_q}} \text{Sel}^d_{\ell,\mathbb{F}_q}.$$

So, as before, using Lang Weil, the geometric second moment of the ℓ-Selmer group is simply the number of components of this space. Further, using our monodromy computation above, we in fact already know the monodromy of this cover of $\mathscr{W}^d_{\mathbb{F}_q}$ is essentially the orthogonal group. The number of components is then identified with the number of orbits of $O(Q)$ acting diagonally on $V \times V$, for $V \simeq (\mathbb{Z}/n\mathbb{Z})^{12d-4}$ the vector space underlying the geometric generic fiber of π. It is then a not too difficult group theory computation to see this has $(1+\ell)(1+\ell^2)$ orbits.

6. The full distribution

Our next goal will be to compute the full geometric distribution of Selmer groups. Since the moments actually grow very fast, they do not determine the distribution, and so additional work is needed.

Theorem 6.1. *The BKLPR model for distributions of Selmer groups holds in the large q limit. That is,*

$$\begin{aligned}(\text{rk}^{\text{BKLPR}}, \text{Sel}_n^{\text{BKLPR}}) &= \lim_{d\to\infty}\left(\limsup_{\substack{q\to\infty \\ \gcd(q,2n)=1}} (\text{rk}, \text{Sel}_n)^d_{\mathbb{F}_q}\right) \\ &= \lim_{d\to\infty}\left(\liminf_{\substack{q\to\infty \\ \gcd(q,2n)=1}} (\text{rk}, \text{Sel}_n)^d_{\mathbb{F}_q}\right).\end{aligned}$$

Recall that the BKLPR distribution modeled Selmer elements as intersections of maximal isotropic subspaces in a Grassmannian. We will compare this to another combinatorial model, which we call the random kernel model.

6.1. **The random kernel distribution.**

Model 6.2 (Random kernel model). Here is the random kernel model for elliptic curves of height d over $\mathbb{F}_q(t)$: Let $g \in O_{12d-4}(\mathbb{Z}/n\mathbb{Z})$ be a random element of the orthogonal group. Then, in the large q limit,

$$\mathrm{Sel}_n(E) \simeq \ker(g-1).$$

Further

$$\mathrm{rk}(E) = \begin{cases} 0 & \text{if } g \in SO_{12d-4}(\mathbb{Z}/n\mathbb{Z}) \\ 1 & \text{if } g \notin SO_{12d-4}(\mathbb{Z}/n\mathbb{Z}) \end{cases}$$

The general outline is then to show the BKLPR model agrees with this model in the large height limit, and then show that this model correctly models ranks and Selmer groups of height d. The first is essentially a probability theory proof, we have two random variables which we want to show agree. When n is prime, there are explicit established formulas for both which are known to be equal (though one can also derive equality more directly by computing enough moments). For composite n, we show both random variables satisfy the same Markov processes bootstrapping from the prime case. So, the more geometric issue is to compare this random kernel distribution to the actual distribution. We now explain how this is done.

6.3. **Sketch of distribution proof.** It remains to connect the actual distribution to the random kernel model. We can connect this to the random kernel model as follows.

$$\begin{aligned} \mathrm{Sel}_n(E) &= \left(\pi^{-1}([E])\right)(\mathbb{F}_q) \\ &= \ker(\rho^d_{n,\mathbb{F}_q}(\mathrm{frob}_{[E]}) - \mathrm{id})) \\ &= \ker(g_{[E]} - \mathrm{id}) \text{ for } g_{[E]} \text{ a random element of } O_{12d-4}(\mathbb{Z}/n\mathbb{Z}) \sim \mathrm{im}\,\rho^d_{n,\mathbb{F}_q} \end{aligned}$$

The first step is the key property of the Selmer space, the second is because $\mathbb{F}_q$ points are the kernel of Frobenius, and the third is a version of the Chebotarev density theorem.

6.4. **Average ranks.** Finally, we'd like to upgrade our main theorem to incorporate ranks of elliptic curves as well. Recall from our model that the rank of an elliptic curve is predicted to be 0 if the Frobenius element lies in SO and 1 otherwise.

Theorem 6.2. *For d a fixed integer at least* 2,

$$\lim_{d\to\infty}\lim_{q\to\infty} \mathrm{Average}^{\leq d}\left(\mathrm{rk}\,/\mathbb{F}_q(t)\right) = 1/2.$$

It may seem difficult at first to see how to relate the Selmer space to analytic ranks, until one realizes that the sheaf defining the Selmer space is actually the same one (!) appearing in the Grothendieck's definition of analytic rank.

In other words, we have the following lemma:

Lemma 6.3. *For E an elliptic curve over $\mathbb{F}_q$ corresponding to a point $x \in \mathscr{W}^d_{\mathbb{F}_q}$, and $L(T, E)$ the associated L-function, we have*

$$\det\left(\mathrm{id} - (\rho^{\mathbb{Z}_\ell}_{d,\mathbb{F}_q})(\mathrm{frob}_x)T\right) = L(T/q, E_x).$$

Corollary 6.4. *The analytic rank is equal to the rank of the generalized* 1*-eigenspace of* $(\rho^{\mathbb{Z}_\ell}_{d,\mathbb{F}_q})(\mathrm{frob}_x)$.

Proof. Recall that the analytic rank is the power of $T-1$ fully dividing the right hand side. Hence, it is the same as the power of $T-1$ fully dividing the left hand side, which is the same as the generalized 1-eigenspace of the matrix $(\rho^{\mathbb{Z}_\ell}_{d,\mathbb{F}_q})(\mathrm{frob}_x)$ (really it is that of $(\rho^{\mathbb{Z}_\ell}_{d,\mathbb{F}_q})(\mathrm{frob}_x)^{-1}$, but this has the same dimension as the generalized 1-eigenspace of the aforementioned element, since it lies in the orthogonal group). □

Now, using Chebotarev density again, we have reduced to understanding the dimension of generalized 1-eigenspaces for random elements of the orthogonal group. We want to show that half the time this dimension is 0 and half the time it is 1. Even stronger, we will show that when the Frobenius lies in SO it is almost always 0 and otherwise it is almost always 1. Essentially, the computation boils down to showing that

Lemma 6.5. *For most $g \in SO$, the element g has a trivial* 1*-eigenspace, while for most $g \in O - SO$, there eigenspace is* 1*-dimensional.*

Proof. Since the loci in O where the eigenspace jumps is a Zariski closed subscheme, it is not too difficult to deduce the above. The key is to observe that every element of $O - SO$ has an odd dimensional 1-eigenspace, since the complex eigenvalues all pair up for elements of the orthogonal group, leaving us with an even number of ± 1s multiplying to -1. This means there must be an odd number of -1s and hence also an odd number of $+1$s. □

Finally, above we established the claimed results for analytic rank in place of algebraic rank, but it is known that when the analytic rank is at most 1 over function fields, then analytic rank equals algebraic rank. So we can make the same conclusion for algebraic rank.

References

[BS13] Manjul Bhargava and Arul Shankar. The average number of elements in the 4-selmer groups of elliptic curves is 7. *arXiv preprint arXiv:1312.7333v1*, 2013.

[dJF11] A. J. de Jong and Robert Friedman. On the geometry of principal homogeneous spaces. *Amer. J. Math.*, 133(3):753–796, 2011.

[EVW16] Jordan S. Ellenberg, Akshay Venkatesh, and Craig Westerland. Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields. *Ann. of Math. (2)*, 183(3):729–786, 2016.

[PR12] Bjorn Poonen and Eric Rains. Random maximal isotropic subspaces and Selmer groups. *J. Amer. Math. Soc.*, 25(1):245–269, 2012.

[R71] A. Grothendieck and M. Raynaud. *Revêtements étales et groupe fondamental*. Springer-Verlag, Berlin-New York, 1971. Séminaire de Géométrie Algébrique du Bois Marie 1960–1961 (SGA 1).