# Analysis, Spectra, and Number Theory: A Conference in Honor of Peter Sarnak's 60th Birthday
December 15-19, 2014

Notes by Tony Feng

# Conference Program

## Disclaimer

These are *very rough* and informal notes that I live-TEXed at the conference. I emphasize that they are my personal notes and *may not accurately reflect the actual contents* of the talks. (In particular, I was unable to scribe for those of Brian Conrey and Alex Eskin. Several other notes are incomplete.) Their faithfulness to the originals has suffered from my insufficiently fast typing, lack of understanding, mental exhaustion, and in some cases shortage of computer battery. For some of the talks, especially the Beamer presentations, only the first portion was captured. Of course, I take full responsibility and apologize for all omissions and inaccuracies.

My intention in writing these notes was for private use, but I have made them public in case they turn out to be useful for anybody. If something should pique your interest, then you can find recordings of the talks at the conference website https://sites.google.com/site/asnt2014/videos.

# THE AVERAGE RANK OF ELLIPTIC CURVES: DATA, CONJECTURES, AND THEOREMS

MANJUL BHARGAVA

## 1. Introduction

Our guiding question is easy to state:

> What is the rank of elliptic curves over $\mathbb{Q}$ *on average*?

Now there are infinitely many different elliptic curves over $\mathbb{Q}$, so in order to formulate this question more precisely, we need a natural way to *order* elliptic curves to have a meaningful notion of "average."

The simplest such measure is the "naïve height," which basically measures the size of the coefficients defining the elliptic curve. That is, if we write

$$E_{A,B}\colon y^2 = x^3 + Ax + B$$

then any $E/\mathbb{Q}$ is isomorphic to a unique $E_{A,B}$ such that $p^4 \mid A \implies p^6 \nmid B$. (Reason: one can perform a simple change of variables $x \mapsto x/p^2$, $y \mapsto y/p^3$.) So there is a *canonical* representation of any $E/\mathbb{Q}$, and we define the height simply by

$$H(E_{A,B}) := \max\{|A|^3, B^2\}.$$

There are other candidates: the *Faltings height* (essentially the logarithm of the naïve height), the discriminant $-4A^3 - 27B^2$ (which should be about the same as the naïve height in general), and the *conductor*.

These are all conjecturally comparable, but even counting isomorphisms classes (let alone rank) with respect to the other invariants is difficult, so we focus on the naïve height.

Now we can formulate our "averaging" question as follows:

$$\boxed{\text{What is } \lim_{B\to\infty} \frac{\sum_{H(E)\le B} \operatorname{rank}(E)}{\sum_{H(E)\le B} 1}\,?}$$

**Conjecture 1.1** (Goldfeld, Katz-Sarnak)**.** *This quantity is* $1/2$*, and more precisely "50%" of elliptic curves should have rank* 0 *and* 50% *should have rank* 1.

The basic heuristic underlying the conjectures is that the parity of the rank (which corresponds to the $\varepsilon$-factor of the elliptic curve) should be equally distributed, and as small as possible in consideration of this.

Attempted computations of the average rank do not seem to lend much support to this conjecture.

- Brumer and McGuinness (1990) observed that the proportion of rank 2 curves seems to be *increasing* with the naïve height. This was confirmed and extended in more recent computations.
- The average rank empirically starts around 0.7 and goes up to 0.87....

It is not even clear from the empirical data that the average rank is bounded. However, there is some theoretical evidence towards this conclusion.

- In 1992, Brumer showed that GRH and BSD together imply that the average rank is bounded (by 2.3).
- In 2004, Heath-Brown showed that GRH and BSD imply that the average rank is at most 2, and this was improved in 2009 to 1.79.

Recently, Bhargava and Shankar proved an *unconditional* boundedness result on the average rank of elliptic curves, which is the subject of our talk.

## 2. Unconditional Boundedness

We study the rank by studying the $n$-Selmer group $S^{(n)}(E)$. We won't go into the definition of what this is now, but it fits into an exact sequence

$$0 \to E(\mathbb{Q})/nE(\mathbb{Q}) \to S^{(n)}(E) \to \text{Ш}_E[n] \to 0.$$

Since the $n$-torsion subgroup of $E$ tends to be trivial, the left term tends to be $n^{\operatorname{rank}(E)}$. In particular, $n^{\operatorname{rank}(E)} \leq |S^{(n)}(E)|$.

Thus to prove the boundedness of average rank, it suffices to bound the average size of $|S^{(n)}(E)|$ for any $n > 1$.

**Theorem 2.1** (Bhargava-Shankar)**.** *Let $n = 1,2,3,4$, or $5$. Then the average size of $S^{(n)}(E)$ is $\sigma(n) := \sum_{d|n} d$.*

*Remark* 2.2. We shall see later a more natural formulation of the result.

**Outline of Proof.**

(1) For each $n \leq 5$, construct a representation $V$ of an algebraic group $G$ over $\mathbb{Z}$ such the $n$-Selmer group injects into its (integral) orbits. More precisely:
   (a) The action of $G(\mathbb{C})$ on $V(\mathbb{C})$ has a free ring of invariants generated by two elements, say $A$ and $B$.
   (b) There is an injective map

$$S^{(n)}(E_{A,B}) \hookrightarrow G(\mathbb{Z})/V(\mathbb{Z})_{A,B}.$$

*Remark* 2.3. By a theorem of Borel and Harish-Chandra, this already implies the finiteness of $S^{(n)}(E_{A,B})$, giving a new proof of Mordell's Theorem.

(2) Next count elements in $G(\mathbb{Z})/V(\mathbb{Z})_{A,B}$ having bounded $A, B$.
(3) Perform a sieve to obtain only the count of elements in the image of the injection, which are defined by an infinite set of congruence conditions.

2.1. **Constructing the representations.** The representations that we require were all constructed classically. We tabulate them below.

| $n$ | $G$ | $V$ | Due to: |
|---|---|---|---|
| 2 | $\mathrm{PGL}_2$ | $\mathrm{Sym}^4(2)$ (binary quartics) | Birch-Swinnerton-Dyer |
| 3 | $\mathrm{PGL}_3$ | $\mathrm{Sym}^3(3)$ (ternary cubics) | Cassels/Cremona-Fisher-Stoll |
| 4 | $< \mathrm{GL}_2 \times \mathrm{GL}_4$ | $2 \otimes \mathrm{Sym}^2(4)$ (two quadrics in $\mathbb{P}^3$) | Cassels/Cremona-Fisher-Stoll |
| 5 | $< \mathrm{GL}_5 \times \mathrm{GL}_5$ | $5 \otimes \bigwedge^2(5)$ | Buchsbaum-Eisenbud/Fisher |

How were these obtained? An element of $S^{(n)}(E_{A,B})$ may be viewed as a map $C \to \mathbb{P}^{n-1}$ (via a complete linear system of degree $n$), where $C$ is a genus 1 curve with Jacobian $E_{A,B}$ and $C$ has points locally at every place.

(1) If you map to $\mathbb{P}^1$, you get 4 ramified points, hence a binary quartic. If you map to $\mathbb{P}^2$, you get a plane cubic (ternary cubic).
(2) If you map to $\mathbb{P}^3$, you get a complete intersection of two quadrics.
(3) If you map to $\mathbb{P}^4$, you no longer get a complete intersection. However, it was classically explored how to paramtrize genus one curves in $\mathbb{P}^4$. Given 5 skew-symmetric matrices $A, B, C, D, E$, you get $Ax + By + Cz + Ds + Et$. The determinant is 0, because it's an odd symplectic matrix, but its $4 \times 4$ minors are (quartic) squares, so their Pfaffians are quadrics. Thus one can obtain the curve as a (not necessarily complete) intersection of 5 quadrics (construction due to Cayley-Sylvester).

Amazingly, this theory works even over $\mathbb{Z}$.

2.2. **Counting lattice points.** We next wish to count the total number of elements of $G(\mathbb{Z})\backslash V(\mathbb{Z})_{A,B}$ with $H(A,B) := \max\{|A|^3, B^2\} < X$.

(1) First construct a fundamental set $L$ for the action of $G(\mathbb{R})$ on $V(\mathbb{R})$ that is absolutely bounded in $V(\mathbb{R})$.
(2) Second, construct a fundamental domain $\mathscr{F}$ for the action of $G(\mathbb{Z})$ on $G(\mathbb{R})$ that is contained in a "Siegel set" (as described by Harish-Chandra) i.e. $\mathscr{F} = N'A'K$ where $N'$ is a bounded set of lower triangular matrices, $A'$ is an unbounded set of elements in a torus, and $K$ is a maximal comacpt subgroup of $G(\mathbb{R})$.

Then for any $g$, the set $\mathscr{F}gL \subset V(\mathbb{R})$ is a fundamental domain for the action of $G(\mathbb{Z})$ on $V(\mathbb{R})$.

Thus we wish to obtain the count of integral points in $\mathscr{F}gL$ having $H(A,B) < X$. This region has finite volume but is *unbounded*, with many cusps going off to infinity. Morally the number of lattice points should be well approximated by the volume, but it turns out that the infinite tentacles may have lots of integral points. This is in general a difficult problem, but here we can exploit the fact that the number of points in this region $\mathscr{F}gL$ is independent of the element $g \in G(\mathbb{R})$.

The solution is to *average* over a compact continuum of the fundamental domains, which has the effect of "thickening" the cusps, allowing a better estimate for the number of integral points in the cusps.

At this point algebro-geomeric techniques come into play. Let $N(V;X)$ be the number of generic $G(\mathbb{Z})$-orbits on $V(\mathbb{Z})$ with height $< X$. Here "generic" means that the corresponding map $C \to \mathbb{P}^{n-1}$ does not correspond to an $n$-Selmer element of order $< n$. The point is that it turns out to be easier to count "generic" elements after averaging.

Let $G_0$ be the closure of a bounded open set in $G(\mathbb{R})$. Then (tautologically)

$$N(V;X) = \frac{\int_{g\in G_0} \#\{v \in \mathscr{F}gL \cap V(\mathbb{Z})^{\text{gen}} \colon H(v) < X\}\, dg}{\int_{g\in G_0} dg}.$$

Now we can "switch the order of integration," so this is

$$\frac{\int_{h\in\mathscr{F}=N'A'K} \#\{v \in hG_0L \cap V(\mathbb{Z})^{\text{gen}} \colon H(v) < X\}\, dh}{\int_{g\in G_0} dg}.$$

Now the integrand is always nice (because $G_0$ is compact). However, the badness has been pushed off to the noncompact parts of $\mathscr{F}$, and this in fact has lots of lattice points. That is explained by an *algebraic condition*: the presence of a variety with many integral points, which turns out to correspond to non-genericness.

So one then partitions $\mathscr{F}$ into a main body and a cuspidal region depending on $A'$. This allows one to show that

$$N(V;X) = \text{Vol}(\mathscr{F}gL \cap \{v \in V(\mathbb{R}) \colon H(v) < X\}) + o(X^{5/6}).$$

This would *not be true* if we were not restricting our attention to generic points!

2.3. **Sieving.** Finally, one has to sieve to count only the elements of $G(\mathbb{Z})\backslash V(\mathbb{Z})$ in the image of the map. To carry this out, one must essentially determine the density of squarefree values taken by the irreducible polynomial

$\Delta(v) = -4A(v)^3 - 27B(v)^2$ (which in the case of $n = 5$, is a polynomial of degree 60 in 50 variables!)..

This is the most technical part of the strategy. One has to be able to bound points that lie on a small (say codimension 2) subvariety over $\mathbb{Z}$ for some large prime $p$.

**Theorem 2.4** (based on work of Ekedahl/Sarnak)**.** *Let $B$ be a compact region in $\mathbb{R}^n$ of finite measure and let $Y$ be any subvariety of $\mathbb{A}^n_{\mathbb{Z}}$ of codimension $k \geq 2$. Let $r$ and $M$ be positive real numbers. Then*

$$\#\{a \in rB \cap \mathbb{Z}^n : a \mod p \in Y(\mathbb{F}_p)) \text{ for some } p > M\} = O(\frac{r^n}{M^{k-1}\log M} + r^{n-k+1}).$$

Once this theorem is proved, this takes care of sieving away those $v$ that have squarefull discriminant at $p > M$ for $\pmod p$ reasons: one gets two conditions, hence a codimension 2 subvariety.

To handle those $v$ that have squarefull discriminant at $p$ for $\pmod{p^2}$ reasons, we use the symmetry group $G$ of $V$ to transform the $\pmod{p^2}$ condition on $v$ to a $\pmod p$ condition. (This is a useful general principle whenever there is a "large enough" symmetry group.) For other applications his hasn't always worked in a straightforward way, in which case an ad hoc solution is to map into a bigger space.

So we have reduced the problem to one of computing the volume of the fundamental domain.

**Lemma 2.5.**

$$\int_{v \in V} f(v)\, dv = |\mathscr{J}| \int_{g \in G(\mathbb{R})} \int_{w \in L} f(g \cdot w) dAdBdg.$$

The remarkable thing here is that the Jacobian is constant: $dv$ is pretty much on the nose $dAdBdg$. Applying this, one gets a product over the real place and all the finite ones which cancel by the product formula. The remaining parts combine to the Tamagawa number, which is known.

2.4. **Conclusion.** We have shown that for $n \leq 5$, the average number of order $n$ elements in the $n$-Selmer group is $n$. It is natural to conjecture that this holds for all $n$.

**Conjecture 2.6.** *For all $n > 0$, the average number of order $n$ elements in the $n$-Selmer group is $n$.*

Therefore, the average number of order $n$ elements in the $n$-Selmer group is $n$. This is established for $n \leq 5$, but it is natural to conjecture for all $n$.

**Proposition 2.7.** *Suppose the conjecture holds. Then 100% of elliptic curves have rank 0 or 1. Furthermore if the equidistribution of root numbers is true, then the strong Goldfeld-Katz-Sarnak conjecture is true.*

Using what is known, one already gets that the average rank is at most 0.885. This implies, for instance, that 80% of elliptic curves have rank 0 or 1.

Combining this with work of Skinner, one can conclude that a positive proportion have rank 1. In particular, the average rank is strictly positive.

Further combining this stuff with work of Zhang on Kolyvagin systems, one can conlude: most curves (at least 66%) have analytic rank 0 or 1, and thus satisfy BSD (by Gross-Zagier). As mentioned above, if the average size of the $n$-Selmer group is $\sigma(n)$ for all $n$, then 100% of curves have analytic rank 0 or 1.

This is just barely too little to deduce anything about the average rank graph turning around, according to the old data. In fact, the average rank graph (recomputed) goes *past* the proven bound, so it *must* turn around! Happy birthday Sarnak!

# THE CRITICAL ZEROS - 100% SOMETIMES

HENRY IWANIEC

## 1. Zeros of $L$-functions

We consider an $L$-function with Euler product

$$L(s) = \prod_p (1 - \lambda(p)p^{-s} + \nu(p)p^{-2s})^{-1} = \sum_n \lambda(n)n^{-s}$$

such that $|\lambda(p)| \leq 2$ and $|\nu(p)| \leq 1$.

**Proposition 1.1.** *$L(s)$ converges for* $\operatorname{Re} s > 1$ *and admits a functional equation of the form*

$$\Lambda(s) = Q^{s/2}G(s)L(s) = \eta\overline{\Lambda}(t - \bar{s}).$$

We denote a zero $\rho$ of $L$ by $\rho = \beta + i\gamma$ for $0 < \beta < 1$.

*Definition* 1.2. We set

$$N(T) = \#\{\rho : |\gamma| \leq T\}$$

and

$$N_0(T) := \#\{\rho : \beta = \frac{1}{2}, |\gamma| \leq T\}$$

**Proposition 1.3.** *We have*

$$N(T) = \frac{2T}{\pi} \log QT + O(T).$$

It's important to remember that this estimate comes from the Gamma factor, showing that every place is important!

Conjecturally (i.e. according to the Generalized Riemann Hypothesis) we have $N(T) = N_0(T)$.

*Definition* 1.4. We set

$$N(\alpha, T) := \#\{\rho : \beta \geq \alpha, |\gamma| \leq T\}$$

**Proposition 1.5.**

$$N(\alpha, T) \ll T^{4\alpha(1-\alpha)}(\log T)^A \text{ if } \alpha \geq \frac{1}{2}, T \geq Q^A.$$

This shows that almost all zeros are very close to the critical line $\mathrm{Re}\, s = \frac{1}{2}$.

**History.** Concerning the critical zeros of $\zeta(s)$,

(1) Hardy-Littlewood showed that $N_0(T) \gg T$.
(2) Selberg showed that $N_0(T) > \kappa N(T)$ for some $\kappa > 0$ (1942).
(3) Levinson showed that one can take $\kappa = 0.342$ (1974).
(4) Conrey improved this to 0.4088 (1989).
(5) Feng further improved this to $\kappa = 0.4128$ (2012).

## 2. Selberg's sign changes

Let

$$f(t) = \frac{G(1/2+it)}{|G(1/2+it)}\zeta(1/2+it)$$

for some function $G$ to be determined later. Selberg's basic idea was to estimate $N_0(T)$ by estimating the sign changes of $f(t)$ in short segments which is derived by comparison of various integral mean values.

We introduce the mollifier

$$M(s) = \sum_{m<M} a(m)m^{-s}$$

Here $M(s)$ "pretends to be" $1/\zeta(s)$ and $a(m)$ pretends to be $\mu(m)$.

We will argue by instead studying the sign changes of $f(t) = \frac{G(1/2+it)}{|G(1/2+it)|}\zeta(1/2+it)M(1/2+it)$. Now, we must take care that $M$ not screw up the sign changes, so Selberg proposed to take $M(s) = |N(s)|^2$ where

$$N(s) = \sum_{n\le N} b(n)n^{-s}, \quad N = \sqrt{M}.$$

Here $N(s)$ pretends to be $1/\sqrt{\zeta(s)}$ and $b(n)$ pretends to be $\mu(n)/\sqrt{\tau(n)}$. Then one throws in some smooth cropping factor $\left(1 - \frac{\log n}{\log N}\right)$ to make the sums nice. ♠♠♠ TONY: [like the Selberg sieve!]

### 2.1. Levinson's work.

Levinson's method is different. He starts with a linear combination of derivatives:

$$G(s) = L(s) + L'(s)/\log, \quad \log N \asymp \log T$$

and

$$F(s) = G(s)M(s), \quad M(s) \text{ a mollifier.}$$

We consider not the critical line but slightly to the *left*: $s = \frac{1}{2} - a + it$, $a = 1 - \alpha/\log TQ$ for some $\alpha > 0$. By a functional equation for $G(s)$, the argument variations and the Littlewood rectangle lemma one gets the inequality

$$N_{00}(T) \geq N(T) - \frac{1}{\pi a}\int_{-T}^{T} \log|F(s)|\,dt + O(T) \geq N(T)(1-\alpha\log I(T)) + O(T)$$

♠♠♠ TONY: [What's $N_{00}$? The number of zeros on the critical strip with imaginary part bounded] where

$$I(T) = \frac{1}{2T}\int_{-T}^{T} |F(t)|\,dt \leq c + o(1), \quad c \geq 1.$$

The methods of Selberg and Levinson are diametrically opposite to each other. Levinson's approach is risky, because it may produce a negative result. But it has the great advantage of opening up the possibility of getting 100% of critical zeros if the mollification is nearly perfect, and one gets $F(s) + 1 + o(1)$.

So that raises the question: does there exist a perfect mollifier? No such thing has been constructed, of course, but the question has been studied by Conrey, Farmer, Goldston, Gonek, Ghosh, etc. For *families* of $L$-functions possessing some structure (e.g. extra orthogonality).

*Remark* 2.1. Amusingly, if one *assumes* the Riemann Hypothesis then one can prove by this method that 100% of the zeros are on the critical line.

## 3. Lacunary $L$-functions

Suppose that we are considering an $L$-function

$$L(s) = \sum_n \lambda(n) n^{-s}$$

where $\lambda(n)$ vanishes or is quite small frequently: quantitatively,

$$\sum_{Q^2<n<N} |\lambda(n)| n^{-s} \leq \varepsilon \frac{\log N}{\log Q}.$$

HJence $\lambda$ is "sparser" than the prime numbers in segments $Q^2 < n < Q^A$ for any $A > 2$ if $s$ is sufficiently small. We want a mollifier close to

$$L(s)^{-1} = \sum_m \rho(m) m^{-s}.$$

Perfect mollification can be achieved with short sums of the form

$$M(s) = \sum_{m \leq M} \rho(m) g(m)$$

where

$$g(m) = \left(1 - \frac{\log m}{\log M}\right)^e \text{ (cropping factor).}$$

Here $M$ is quite small, about $T^{1/2000}$.

3.1. **Exceptional Discriminants.** Let $K = \mathbb{Q}(\sqrt{D})$ be a quadratic number field. Let $\psi\colon \mathrm{Cl}(K) \to \mathbb{C}$ be a character of the ideal class group. Set

$$L(s,\psi) = \sum_{\mathfrak{a}} \psi(\mathfrak{a})\,\mathrm{Nm}(\mathfrak{a})^{-s}.$$

♠♠♠ TONY: [missed some discussion]

**Theorem 3.1.** *Let $N_{00}(T)$ denote the number of simple zeros $\rho = \frac{1}{2} + i\gamma$ of $L(s,\psi)$ with $|\gamma| \leq T$ and let $N(T)$ be the number of all zeros $\rho = \beta + i\gamma$ with $0 < \beta < 1$ and $|\gamma| \leq T$ counted with multiplicity. Put $\varepsilon = \varepsilon(D) = L(1,\chi)\log|D|$. Then we have*

$$N_{00}(T) = \{1 + O(s(D)^{1/2})\}N(T) + O(T)$$

*for $|D|^{1/2} < T < |D|^{A/\varepsilon}$ with any constant $A > 1$, where the implied constant depends only on $A$.*

*Definition* 3.2. An infinite sequence of discriminants $D$ with $\varepsilon(D) \to 0$ is called *exceptional*.

**Corollary 3.3.** *As $D$ runs over an exceptional sequence, the number of critical simple zeros of $L(s,\psi)$ of height $\leq T$ approaches a 100% of all its zeros of height $\leq T$ (counted with multiplicity) provided $|D|^{1/2} \leq T \leq |D|^{A/\varepsilon}$.*

# THE DISTRIBUTION OF MODULAR CLOSED GEODESICS REVISITED

BILL DUKE

## 1. BINARY QUADRATIC FORMS

Let $D$ be a fundamental discriminant (i.e. a discriminant of a quadratic extension of $\mathbb{Q}$). Let

$$\Lambda_D = \{Q(x,y) = Ax^2 + Bxy + Cy^2 \colon D = B^2 - 4AC\}/\Gamma = \mathrm{PSL}(2,\mathbb{Z}).$$

This is a *finite* set (but it's non-trivial to prove that).

**Proposition 1.1.** *If $h(D)$ is the class number of $Q(\sqrt{D})$, then*

$$\#\Lambda_D = h(D).$$

If $D < 0$, then associated to $Q$ is the CM-point

$$z_Q = \frac{-B + \sqrt{D}}{2A} \in \mathbb{H}.$$

If $D > 0$, then we instead get a *geodesic*. If $t^2 - Du^2 = 4$ and $t, u \geq 1$ are minimal (so $\varepsilon = \frac{t+u\sqrt{D}}{2}$ is a fundamental unit greater than 1) then the geodesic corresponds to

$$g_Q = \pm \begin{pmatrix} \frac{t+Bu}{2} & Cu \\ -Au & \frac{t-Bu}{2} \end{pmatrix} \in \Gamma.$$

♠♠♠ TONY: [so the geodesic is $g_Q$ acting on $\begin{pmatrix} e^{s/2} & \\ & e^{-s/2} \end{pmatrix}$]?

## 2. THE CLASS NUMBER FORMULA

Let

$$\zeta(z,s) = \frac{1}{2} \sum_{m,n}{}' |mz+n|^{-2s} = \zeta(2s) E(z,s)$$

where

$$E(z,s) = \sum_{\gamma \in G_\infty \backslash \Gamma} (\mathrm{Im}\, \gamma)^s.$$

Then we have $\zeta_K(s) = L(s,\chi_D)\zeta(s)$ which is

$$\zeta_K(s) = (\frac{2}{\sqrt{D}})^s \zeta(2s) \sum_{[Q]} E(z,s) \quad \text{if } D < 0$$

and (perhaps more unfamiliar!)

$$\zeta_K(s) = D^{-s/2} \frac{P(s)}{P(s/2)^2} \sum_{[Q]} \int_{g_Q} E(z_Q,s) \frac{\sqrt{D}\,dz}{Q(z,1)} \quad \text{if } D > 0.$$

**Theorem 2.1** (Class Number Formula)**.** *If $D < 0<$ then*

$$L(1,\chi_D) = \frac{2\pi}{\omega} |D|^{-1/2} h(D).$$

*If $D > 0$, then*

$$L(1,\chi_D) = D^{-1/2}(\log \varepsilon) h(D).$$

The first formula clearly counts the number of CM points. One can think of the second expression as the total length of the geodesics, as each one has length $\log \varepsilon$. That gives a uniform interpretation of the class number formula as relating the special value of $L(1,\chi_D)$ and a "volume."

Siegel had the remarkable insight that even if Siegel zeros exist, one can still establish an (ineffective) lower bound:

$$L(1,\chi_D) \gg_\varepsilon |D|^{-\varepsilon}.$$

$E(z,s)$ is part of the spectral resolution of $\Delta = y^2(\partial_x^2 + \partial_y^2)$. If $f$ is "nice," then it will have a decomposition o fthe form

$$f(z) = \int_0^\infty g(t) E(z, \frac{1}{2} + it)\, dt + \sum c_n \varphi_n(z)$$

## 3. Vertical Distribution

Suppose $D < 0$. Let $\varphi_0$ be a smooth, compactly supported function and

$$\varphi(z) = \sum_{\gamma \in \Gamma_0 \backslash \Gamma} \varphi_0((\mathrm{Im}\, \gamma z)^{-1}).$$

We want to study the sum

$$\sum_{z_Q \in \Lambda_D} \varphi(z_Q).$$

To digest this, we introduce the Mellin transform

$$\widetilde{\varphi_0}(s) = \int_0^\infty \varphi_0(t) t^s \frac{dt}{t}$$

which satisfies an inverse Mellin transform property:

$$\varphi_0(s) = \frac{1}{2\pi i} \int_{\mathrm{Re}\, s = c} \widetilde{\varphi}_0(s) y^{-s}\, ds$$

Applying this term by term to the sum defining $\varphi$, we obtain

$$\begin{aligned}\sum_{z\in\Lambda_D}\varphi(z_Q) &= \frac{1}{2\pi i}\int_{\mathrm{Re}\, s=c}\widetilde{\varphi}(s)\sum_Q E(z,s)\,ds\\ &= c\cdot h(D)+O(|D|^{1/4}\int_{-\infty}^{\infty}|L(\frac{1}{2}+it,\chi_D)||\varphi_0(\frac{1}{2}+it)|\,dt.\end{aligned}$$

♠♠♠ TONY: [not exactly sure what happened here - some shift in contour?] So we want a bound on the $L$-function for large $|D|$.

**Theorem 3.1** (Burgess)**.** *We have*

$$|L(1/2+it)| \ll D^{\frac{1}{4}-\delta} \quad \text{for } \delta < 1/16.$$

## 4. The Katok-Sarnak Formula

Let $D<0$ and $\varphi$ be a Maass cusp form (even, to avoid trivialities). Katok and Sarnak consider

$$\frac{1}{\langle\varphi,\varphi\rangle}\sum_Q\varphi(z_Q) = 24\pi|D|^{3/4}\sum_{F_j\to\varphi}\rho_j(D)\overline{\rho_j}(1)$$

Here $\rho_j(D)$ is a Fourier coefficient of $F_j$, having weight $1/2$. Iwaniec found a way of estimating this nontrivially. Duke and Iwaniec used this idea to prove the following:

**Theorem 4.1** (Duke-Iwaniec)**.** $|\rho_j(D)| \ll |D|^{-1/4-\delta}$ *for some* $\delta>0$.

This leads to equidistribution results.

## 5. Genus Character

Factorizations $D=D_1D_2$ such that $(D_1,D_2)=1$ parametrize characters of the class group of $K=\mathbb{Q}(\sqrt{D})$. ♠♠♠ TONY: [oh?] If you have a factorization where $D_1<0$ and $D_2<0$, corresponding to the character $\chi$, then

$$L_K(s,\chi) = iD^{-5/2}\frac{\Gamma(s)\zeta(2s)}{\Gamma(\frac{s+1}{2})^2}\sum_Q\chi(Q)\int_{g_Q}\partial_z E(z,s)\,ds$$

and

$$\begin{aligned}L(0,\chi_D) &= L(0,\chi_{D_1})L(0,\chi_{D_2})\\ &= \frac{4}{\omega_1\omega_2}h(D_1)h(D_2)\\ &= \frac{1}{12}\sum_{Q\in\Lambda_D}\chi(Q)\Psi(Q)\end{aligned}$$

♠♠♠ TONY: [what is $\Psi$?]

There is a (signed) measure $\mu_\chi$ which basically corresponds to adding up the area weighted by an integer multiplicity, and the theorem is that it becomes equidistributed.

**Theorem 5.1.**
$$\frac{\int \varphi(z)\, d\mu_\chi}{\int \varphi(z)\, d\mu} \to \frac{3}{\pi} \int_{\Gamma\backslash\mathbb{H}} \varphi\, d\mu.$$

# MULTIPLICATIVE RELATIONS AMONG SINGULAR MODULI

JONATHAN PILA

## 1. SINGULAR MODULI

Singular moduli are "special values" of the $j$-function.

*Definition* 1.1. A *singular modulus* is $j(\tau)$ where $[\mathbb{Q}(\tau)\colon \mathbb{Q}] = 2$.

**Theorem 1.2** (Oort)**.** *The values $\tau$ and $j(\tau)$ are both algebraic if and only if $j(\tau)$ is a singular modulus.*

Let $\Sigma$ be the set of singular moduli. André Oort's conjecture deals with the existence of "special" subvarieties of $\mathbb{C}^n$ (regarded as a Shimura variety), which are those having some coordinates in $\Sigma$.

♠♠♠ TONY: [So view Oort's theorem as saying something about $\tau, j(\tau)$ being in "special position" with respect to each other.]

*Definition* 1.3. A *multi-modular $n$-tuple* is an $n$-tuple of distinct elements of $\Sigma$ which satisfy a non-trivial multiplicative relation, but such that no proper subset of them does.

**Theorem 1.4** (Pila, Tsimerman)**.** *There are only finitely many multi-modular $n$-tuples.*

*Remark* 1.5. This is ineffective.

This is all part of a more general framework of Zilber-Pink.

Let $X := X_n := \mathbb{C}^n \times (\mathbb{C}^\times)^n$. *Special subvarieties* in $\mathbb{C}^n$ are modular subvarieties $M$ with a condition like some factor has coordinates all being singular moduli ♠♠♠ TONY: [actual condition was a little more technical]. Special subvarieties in $(\mathbb{C}^\times)^n$ are "torsion cosets," i.e. cosets $T$ of subtori by torsion points.

The special subvarieties in $X \times (\mathbb{C}^\times)^n$ are those of the form $M \times T$.

There is also a weak version, where you allow $x_i =$ constant (not necessarily in $M$).

Anyway, let $V \subset X$ be defined by $x_i = t_i$. If $P = (\sigma_1, \dots, \sigma_n)$ is a multiplicative tuple, then $(P,P)$ lies on $V$. Note that this is an atypical situation: $(P,P)$ lies on the intersection of a codimension $n$ and codimension $n+1$ variety.

**Conjecture 1.6** (Zilber-Pink)**.** *Let $X$ be a variety of "mixed Shimura" type and $V \subset X$. There is a finite subset $\mathscr{S}_V$ of proper special subvarieties such that if $\mathscr{S}$ is a special subvariety and $A \subset V \cap S$ is atypical then $A \subset B$ for some atypical $B \subset V \cap T$, where $T \in \mathscr{S}_V$.*

This encompasses many of the other conjectures: André-Oort, Mordell-Lang, etc.

## PROBLEM SESSION

(1) (Matt Emerton) This is a question about talks that will come later. Let $\Gamma_j$ be a decreasing sequence of Bianchi groups. It's known that that the $\mathbb{R}$-Betti numbers of Bianchi groups grow sublinearly with the volume, i.e.

$$\frac{b_i(\Gamma_j, \mathbb{Q})}{\mathrm{Vol}(\Gamma_j)} \to 0$$

Is this same true for Betti numbers with $\mathbb{F}_p$-coefficients?

(2) Here's an analog of the preceding questionfor class groups of imaginary quadratic fields. Can we put good bounds on the $p$-part of class groups?

Sound says that using GRH one should be able to prove something, e.g. "$|D|^{1/2-1/p}$."

(3) Continuing on this theme, can we prove nontrivial bounds for averages and moments of $p$-parts of class groups (along the lines of work of Heath-Brown)?

(4) Can we perhaps use computation to find examples of elliptic curves with large rank (larger than Elkies's record of 28), or examples with small conductor and (relatively) large rank? This might shed insight into the growth of the rank. Also, for all ranks up to 28 can we find examples with maximal (or at least smaller) conductor? There is good empirical "evidence" for this.

(5) (Frank Calegari) This is a question about thin groups. Let $\Gamma \leq \mathrm{SL}_3(\mathbb{Z})$ be a thin, Zariski-dense subgroup. Must $\Gamma$ map to a finite group that is not a quotient of $\mathrm{SL}_3(\mathbb{Z})$?

(6) (Sarnak) A theorem of Vinberg says that if $F$ is a rational quadratic form, $G$ the orthogonal group of $F$, and $R_2$ is the subgroup generated by hyperbolic reflections in $G(\mathbb{Z})$, then $R_2$ is normal and $G(\mathbb{Z})/R_2$ is infinite if $\dim F > 30$.

Follow-up question: is this still true if you replace $R_2$ with the group generated by *all* reflections? The motivation is that one wants to allow elements of the full Weyl group, not just the hyperbolic reflections.

(7) Here is a question about central values of $L$-functions of $\mathrm{GL}_n \times \mathrm{GL}_r$. Fix a representation $\pi$ of $\mathrm{GL}_n$ a (self-dual) representation $\pi'$ for $\mathrm{GL}_r$ such that

$$L(\frac{1}{2}, \pi \times \pi') \neq 0.$$

For $r = n-1$, this is often possible. For $r = 1$, it is open ("more difficult than the Riemann hypothesis"). For which pairs $(n,r)$ can we prove that it is possible?

Sarnak: some insight into these types of questions should come out of an improved understanding of the trace formula. Can we make the trace formula for $\mathrm{GL}_n$ more "analytically flexible" (say, as flexible as it is for $\mathrm{GL}_2$)?

A special case where one might try this is $\mathrm{GL}_n \times \mathrm{GL}_n \subset \mathrm{GL}_{2n}$ (work of Friedberg and Jacquet).

(8) (Alex Lubotzky) Does $\mathrm{SL}_3(\mathbb{Z})$ have a thin subgroup with the super-rigidity property? More generally, to what extent do thin groups behave like arithmetic ones?

Motivation: we should be able to construct more "complicated" examples of thin subgroups.

(9) (Nick Katz) Are there thin groups of exceptional type (e.g. $G_2$) coming from monodromy?

(10) Definition of a Landau Siegel zero: a sequence of characters with

$$L(1,\chi) \leq \frac{\varepsilon(q)}{\log q}$$

and $\varepsilon(q) \to 0$ as $q \to \infty$. How could you recognize a Siegel zero "in nature"?

(11) (Henrik Iwaniec) How can you exploit Sato-Tate in conjunction with sieve methods to get new results about prime numbers? The motivation is analogies with results exploiting exceptional characters.

# SUMS OF THREE SQUARES AND SPATIAL STATISTICS ON THE SPHERE

ZEEV RUDNICK

## 1. Introduction

1.1. **Classical results.** Lagrange showed in 1770 that every positive integer is a sum of 4 squares. Jacobi refined this in 1834 by showing that the number of such representations is

$$N_4(n) = 8 \sum_{\substack{d|n \\ d \not\equiv 0 \pmod 4}} d \implies N_4(n) \approx n^{1+o(1)}.$$

What about sums of two squares? Fermat showed that a prime $p$ is a sum of two squares if and only if $p \equiv 1 \pmod 4$. More generally, an integer $n$ is a sum of two squares if and only if each prime factor congruent to 3 $(\mathrm{mod}\ 4)$ appears with *even* multiplicity. Therefore, asymptotically 0% of integers can be expressed as a sum of two squares.

So most integers are sums of 4 squares in many ways, and most integers are not sums of 2 squares. What about 3 squares?

Legendre/Gauss showed that $n$ is a sum of 3 squares if and only if $n \neq 4^a(8n+7)$. If we define

$$N_n = \#\{(x,y,z) \mid x^2+y^2+z^2 = n\}$$

then if $n$ is primitively representable as a sum of three squares, we have $N_n \approx n^{1/2 \pm \varepsilon}$. The ingredients of proof are Gauss' formula

$$N_n \approx \sqrt{n} L(1, \chi_{-n})$$

plus Siegel's theorem on lower bounds for class numbers. The GRH would give a good, effective lower bound.

1.2. **Spatial distribution of solutions.** We can project integer solutions of

$$x^2+y^2+z^2 = n$$

onto the unit sphere (by scaling). We call the images *Linnik points*. Our question concerns the *distribution* of these points. For instance, are they "random" or "rigid"?

By random, we mean the distribution obtained by drawing independently from a uniform distribution. By rigid, we basically mean something that looks like a lattice on the sphere.

*Definition* 1.1. A collection of subsets $E_n$ in $S^2$ become *uniformly distributed* if for any nice set $B$ in $S^2$,

$$\frac{\#(E_n \cap B)}{\#E_n} \xrightarrow{n\to\infty} \mu(B).$$

Linnik conjectured that the Linnik points $L(n)$ become uniformly distributed on $S^2$. This was proved by Linnik "partially" assuming GRH (1940). Much later, it was proved *unconditionally* by Duke, and by Golubeva-Fomenko (1988). Pommerenke (1959) showed that a similar result holds in higher dimension.

**Failures on the circle.** Kátai-Környei (1977) and Erdös-Hall (1999) establish that for "almost all" $n$ that *are* sums of two squares, the projected lattice points $(x,y)/\sqrt{n} \in S^1$ are uniformly distributed. (Recall that most $n$ are not sums of two squares, so one is conditioning on a strong condition.)

However, examples of *specific sequences* have been found such that the limit distribution is not uniform.

## 2. Electrostatic Energy

The electrostatic energy of $N$ points $P_1,\ldots,P_N$ on the sphere $S^2$ is

$$E(P_1,\ldots,P_N) := \sum_{i=1}^{N}\sum_{j\neq i}\frac{1}{|P_i - P_j|}.$$

**Problem.** (Thompson, 1904) Find the configurations of charges on the sphere which *minimize* energy.

These should be "stable configurations." Thompson was motivated by an attempt to model the atom as a "plum pudding" of electrons among a "pudding" of positively charged matter. (Unfortunately, this model was shortly disproved.)

This turns out to be a very difficult problem. They have been identified numerically for $N < 112$. Rigorously, they are known only for 6 cases: $N = 2,3,4,5,6,12$.

However, Wagner (1992) established that the energy of stable configurations is $\approx N^2$. The basic idea is that

$$E(P_1,\ldots,P_N) \approx N^2 \int_x \int_{|x-y|\le 1} \frac{dy}{|x-y|}\, dx$$

Peeled (2010) showed that for $N$ *random* points, the energy is about $N^2$. That suggests that random points should give an almost stable configuration.

**Theorem 2.1** (Bourgain, Rudnick, Sarnak)**.** *The energy of the Linnik points $L(n)$ is close to minimal:*

$$E(L(n)) = N^2 + O(N^{2-\delta}).$$

*Proof.* We would like to use uniform distribution to claim that for each $P \in L(N)$,

$$\frac{1}{N}\sum_{Q\neq P}\frac{1}{P-Q} \sim \int_{S^2}\frac{dx}{|P-x|} \approx 1.$$

Therefore,

$$E(L(n)) = \sum_{P\in L(n)}\sum_{Q\neq P}\frac{1}{|P-Q|} \approx N^2.$$

There are some problems with this heuristic argument. One is that the function we are averaging is actually not continuous, so equidistribution does not apply. Another is that there can be two points whose projection onto the sphere is *much* closer than expected. For instance, if $n = k^2+(k+1)^2$, then the projections of $((k+1),k,0)$ and $(k,(k+1),0)$ are about $\frac{\sqrt{2}}{\sqrt{n}}$ apart.

So we need to control the number of "unexpectedly close" pairs of points. Introduce the counting function

$$A(n,h) := \#\{(X,Y)\in\mathbb{Z}^3\times\mathbb{Z}^3 \mid |X|^2=|Y|^2=n, |X-Y|^2=h\}.$$

Siegel's mass formula gives a formula for this number as a product of local densities. Eventually, after explicit computation of the the local factors, one can show that the number of close pairs grows slowly, and that wraps up the proof. $\square$

## 3. Nearest Neighbor Distances

Define the nearest neighbor distance

$$d(x) = \min_{y\neq x}|x-y|.$$

Heuristically, we guess that among a set of $N$ points, the nearest distances should be about $1/\sqrt{N}$.

This follows from a simple packing argument. We claim that for any set $X$ of $N$ points on the sphere,

$$\sum_{x\in X} d(x)^2 \leq 16.$$

To prove this, estimate the area of a ball of radius $A$ about each point, supposing that no two such balls intersect, and bound this above by the total surface area of the sphere.

Dahlberg (1978) showed that in any stable configuration, the nearest neighbor distances are *all* commensurable to $1/\sqrt{N}$. On the other hand, for random points the distribution of minimal distances versus $1/\sqrt{N}$ is "unbounded."

Based on empirical evidence, we conjecture:

**Conjecture 3.1.** *For the Linnik ponts $L(n)$, the distribution of nearest neighbor distances is like those of random points.*

**Theorem 3.2.** *Assuming GRH, any possible limit of the distribution is absolutely continuous.*

This means that one has to show some upper bound on the number of points with a certain nearest neighbor distance. This is a quantity $A(n,h)$ of the form mentioned earlier. One finds that this is essentially bounded by a multiplicative function along a quadratic progression:

$$A(n,h) \leq 24F_n(h(2n-h)).$$

There is general machinery to evaluate sums of this form. The result is some exponential sum. We don't know how to bound it unconditionally, so we resort to GRH.

3.1. **Least Spacing Statistic.** What if we study not the minimum neighbor distance for a given point, but the least such minimum over all points? For "rigid" systems this value is about $1/\sqrt{N}$, as discussed already, but for "random" systems it is about $1/N$.

**Theorem 3.3.** *For almost all $n$,*

$$d_{\min} L(n) \approx N^{-1+o(1)}.$$

This is random-like behavior (not rigid!).

This follows from the assertion that almost all $n$ can be expressed as the sum of two squares and a "mini-square,"

$$n = x^2 + y^2 + z^2 \quad , |z| \leq n^{\varepsilon}.$$

Linnik conjectured this for all $n$, but it is open. However, Wooley (2013) proved that for almost all $n$, one does get a good enough upper bound to deduce the theorem.

# ARITHMETIC STATISTICS AND FUNCTION FIELDS

JON KEATING

## 1. Introduction

Let $a(n)$ denote an arithmetic function (e.g. $\Lambda(n), \mu(n), \ldots$). The overarching questions is:

| |
|---|
| What are the statistical characteristics of the fluctuations in $a(n)$? |

More specifically, one could ask about the sum of $a(n)$ over "short" intervals, or "correlation sums" of the form

$$\sum_{n} a(n)a(n+h).$$

Here we are concerned with arithmetic statistics in *function fields*. Let $\mathscr{P}_n$ be the set of poynomials of degree $n$ over $\mathbb{F}_q$ and $\mathscr{M}_n \subset \mathscr{P}_n$ the subset of monic polynomials. We define the *norm* of a polynomial $0 \neq f \in \mathbb{F}_q[t]$ to be $q^{\deg f}$.

## 2. Prime Number Theorems

### 2.1. The von Mangoldt function.

The Prime Number Theorem says that

$$\sum_{n \leq X} \Lambda(n) \sim X.$$

The Hardy-Littlewood conjecture predicts the correlations of the von Mangoldt function:

$$\sum_{n \leq X} \Lambda(n)\Lambda(n+h) \sim C_{HL}(h) \cdot X.$$

There is also a conjecture due to Goldston and Montgomery (1987) and Montgomery and Soundararajan (2004) conerning the fluctuations in small intervals. This says that for $X^{\delta} < H < X^{1-\delta}$,

$$\frac{1}{X} \int_2^X \left| \sum_{n \in [x-H/2, x+H/2]} \Lambda(n) - H \right|^2 dx \sim H(\log X - \log H - \gamma_E - \log 2\pi).$$

It is related to the Hardy-Littlewood conjecture, capturing the correlations among prime numbers.

2.2. **Function Fields.** The"Prime Number Theorem" for function fields says that

$$\sum_{f \in \mathscr{M}_n} \Lambda(f) = q^n.$$

Interestingly, this is an *exact* formula. That might lead you to believe that considering monic polynomials is too rigid. However, Keating and Rudnick showed the exact analogue of the Goldston-Montgomery conjecture for variance of $\Lambda$ in "intervals," and this is quite subtle analytically.

*Proof Sketch.* Averaging over *all* monic polynomials is simple, using analogous techniques to those from analytic number theory, and the fact that the zeta function has *no* zeros.

To evaluate over special subsets (e.g. intervals and arithmetic progressions) one uses characters to "project" the sums. This leads to expressions in terms of the associated $L$-functions, which have zeros and functional equation, etc. In particular, they can be expressed in terms of unitary matrices. In the limit as $q \to \infty$ one uses equidistribution results due to Katz to express the sums in terms of matrix integrals. $\square$

*Remark* 2.1. These results apply in the "$q$ limit" ($q \to \infty$), but we can't show anything in the "$n$ limit" although they probably should be true.

2.3. **The Möbius function.** The Prime Number Theorem (reformulated) is

$$\sum_{n \le X} \mu(n) = o(X).$$

The Chowla conjecture (1965) predicts that there is "no correlation" in the Möbius function: for any distinct $\alpha_1, \ldots, \alpha_m$ and exponents $a_1, \ldots, a_m$ at least one of which is odd,

$$\sum_{n \le X} \mu(n+\alpha_1)^{a_1} \ldots \mu(n+\alpha_m)^{a_m} = o(X).$$

If one defines an analogue for function fields, then one can prove analogues of arithmetic conjectures for sums of Möbius in intervals.

Curiously, the variance of $\mu(f)$ over small intervals is *more* subtle than the number field setting.

# SEARCHING FOR THIN GROUPS

NICK KATZ

## 1. Introduction

Sarnak interested me in the following problem:

> Find (over $\mathbb{C}$) a one-parameter family of curves of genus $g \geq 2$ whose integer monodromy group $\Gamma \subset \mathrm{Sp}(2g,\mathbb{Z})$ is Zariski dense but *thin* (of infinite index).

I have not been able to find any such families, and I am beginning to wonder if any exist at all.

Now, there *are* thin, Zariski-dense subgroups of $\mathrm{Sp}(4,\mathbb{Z})$ occurring in the monodromy of one-parameter groups of *three-dimensional* varieties. One might hope to "transport" these to a family of genus two curves, but this turns out to be impossible (in any family of curves the Jordan blocks have size at most two, but in the desired families the Jordan blocks have size 4).

### 1.1. Attempts.

Here are three examples of families that I studied.

*Example* 1.1. The family

$$y^2 = (x^{2g} - 1)(x - t)$$

or more generally

$$y^2 = f_{2g}(x)(x - t).$$

Here the (finitely many) bad values of $t$ are the roots of $f$, and at each the local monodromy is a transvection (unipotent pseudoreflection).

*Example* 1.2. The family

$$y^2 = f_{2g+1}(x) - t$$

for $f_{2g+1}(x)$ any chosen Morse polynomial of degree $2g+1$ (i.e. $f'$ has $2g$ distinct zeros, and $f$ separates them).

*Example* 1.3. The family

$$y^2 = x^{2g+1} + ax + b$$

over the parameter space which is the curve of discriminant 1:

$$(n-1)^{n-1}a^n + n^n b^{n-1} = 1.$$

Unfortunately, for all three families the monodromy group is of finite index in $\mathrm{Sp}(2g,\mathbb{Z})$. Why?

There aren't so many tools available to prove this kind of result, so we can basically list all of them .

(1) A'Campo (1979): over the parameter space $\mathrm{Config}_{2g+1}$ the family of curves

$$y^2 = \prod_{i=1}^{2g+1}(x-a_i)$$

has an (explicitly known) monodromy of finite index in $\mathrm{Sp}(2g,\mathbb{Z})$.

(2) *Margulis normal subgroup theorem* (special case): if $\Gamma \subset \mathrm{Sp}(2g,\mathbb{Z})$ for $g \geq 2$ is a subgroup of finite index and $\Gamma_1 \subset \Gamma$ is a normal subgroup, then $\Gamma_1$ is either itself of finite index in $\mathrm{Sp}(2g,\mathbb{Z})$ or trivial or $\pm I$.

How do we construct normal subgroups? We use the low end of a long exact homotopy sequence of a (Serre) fibration $F \to E \to B$, ending in (if the fiber is connected)

$$\ldots \to \pi_1(B) \to \pi_1(F) \to \pi_1(E) \to \pi_1(B) \to 1$$

so the image of $\pi_1(E)$ is normal.

Let's try to apply these two tools to the first family. We consider a map

$$E = \mathrm{Config}_{2g+1} \to B = \mathrm{Config}_{2g}$$

by forgetting the last coordinate. The fiber over $(b_1,\ldots,b_{2g}) \in B$ is the subset of $\mathbb{C}\setminus\{b_1,b_2,\ldots,b_{2g}\}$ over which we have the one parameter family $y^2 = f_{2g}(x)(x-t)$ for $f_{2g}(x) := \prod_{i=1}^{2g}(x-b_i)$. By Margulis's theorem (since we already know that the monodromy for $E$ is Zariski-dense a priori) we must be in the finite index case.

A variant concerns the family

$$\prod(x-a_i) = x^{2g+1} + \sum s_k x^k$$

Recall that we wanted to restrict our attention to Morse polynomials. So we need one more tool:

**Lemma 1.4.** *If $U$ is a smooth connected quasiprojective variety over $\mathbb{C}$ of positive dimension and $V \subset U$ is non-empty Zariski open, then $\pi_1(V) \twoheadrightarrow \pi_1(U)$.*

*Proof Sketch.* This is clear if $d = 1$. If $d \geq 2$, we reduce to the $d = 1$ case by successively using the Lefschetz hyperplane theorem, which says that the fundamental group of a general hyperplane section maps onto the original.) $\square$

The point is that we can use this to apply our tools to the open Zariski dense locus of Morse polynomials inside the locus of polynomials with non-vanishing discriminant. That shows the failure of the second example.

Finally, we consider the third example. For the two parameter family $y^2 = x^{2g+1} + ax + b$, the parameter space is $\mathbb{A}^2[\Delta^{-1}]$. This is a fibration with fibers the curves $\Delta =$ constant, and you use Margulis plus the known Zariski density to win.

1.2. **Conclusion.** Given the difficulty of constructing the desired example, we ask:

> Is there any conceptual reason to think that we cannot get Zariski dense thin subgroups of $\mathrm{Sp}(2g,\mathbb{Z})$ as monodromy of families of curves (or of abelian varieties)?

# A STANDARD ZERO FREE REGION FOR RANKIN-SELBERG L-FUNCTIONS ON $\mathrm{GL}(n)$

XIAOQING LI

## 1. Review of Poussin's method

Vallée Poussin (1989) proved a zero-free region for $\zeta(s)$ using an auxiliary $L$-function satisfying certain conditions. His method works for all automorphic $L$-functions. It *also* works for the Rankin-Selberg $L$-functions $L(s,\pi\times\pi')$ if one of $\pi$ and $\pi'$ is self-dual.

However, if both are non-self-dual then this method doesn't work. In that case, the best zero-free region know is due to Brumley (2006), of the form

$$\sigma > 1 - \frac{c}{Q_{\pi_1}Q_{\pi_2}(|t|+2))^N}.$$

## 2. Statement of Results

We want a "standard" zero-free region, which takes the form

$$\sigma > 1 - \frac{c}{\log(Q_{\pi-1}Q_{\pi-2}(|t|+2))^B}.$$

We will focus on a Rankin-Selberg $L$-functions of the special form $L(s,f\times f)$.

**Theorem 2.1.** *Let $\pi$ be an irreducible cuspidal unramified representation of $\mathrm{GL}(n,\mathbb{A}_{\mathbb{Q}})$ for $n\geq 2$ which is tempered at all finite primes. Then a zero-free region is given by*

$$\sigma > 1 - \frac{c}{\log^5(|t|+2)}.$$

This follows directly from a lower-bound theorem: if $f$ is a Maass cusp form in the space of $\pi$ as above, then

$$|L(1+it,f\times f)| \gg \frac{1}{\log^3(|t|+2)}.$$

2.1. **The Eisenstein Series method.** Our proof is via the Eisenstein series method. In 1976, Jacquet and Shalika introduced this method for $\mathrm{GL}_n$ to prove the non-vanishing of $L(1+it,\pi)$. (This doesn't give a zero-free region, only non-vanishing.)

Shahidi proved non-vanishing of Rankin-Selberg $L$-functions on the 1-line. Then Sarnak showed how to obtain a zero-free region for $\zeta$ using this method, and with Gelbart and Lapid established a zero-free region for all Langlands-Shahidi $L$-functions. In particular, they prove that

$$L(\mathrm{Sym}^9, 1+it) \gg \frac{1}{(|t|+1)^A}.$$

Our approach is based on Sarnak's method.

It goes as follows. Let $\Gamma = \mathrm{SL}_(2,\mathbb{Z})$ and $\Gamma_\infty$ be the stabilizer of $\infty$. Let

$$E_\Gamma(z,s) := \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \mathrm{Im}\,(\gamma z)^s.$$

One considers an integral

$$I = \int_\eta^\infty \eta_0^1 |\zeta(1+2it)|^2 |E_A(z,1/2+it)|^2 d^\times z$$

where $E_A$ is a the "truncated Eisenstein series." Sarnak established upper and lower bounds for this integral, thus deducing the desired inequality.

**Lemma 2.2.**

$$I \ll \frac{1}{\eta}|\zeta(1+2it)|(\log^2 t + 2\log A).$$

*Proof Sketch.* Use the Maass-Selberg relation for GL(2) to obtain an expression for the integral of the $|E_A(z,1/2+it)|^2$. □

**Lemma 2.3.**

$$I \gg \frac{1}{\eta}\frac{1}{\log t}.$$

*Proof Sketch.* This is very tricky. First compute the Fourier expansion of the (truncated) Eisenstein series (the coefficients are basically Bessel function values). Then one uses Parseval's identity to express the desired integral in terms of these Fourier coefficients, and then use sieve theory to show that many terms are non-vanishing. □

2.2. **Outline of proof.** Let $G = \mathrm{GL}(2n,\mathbb{R})$, $\Gamma = \mathrm{SL}(2n,\mathbb{Z})$ and $K = O(2n,\mathbb{R})$. Let $P_{n,n}$ be the maximal parabolic, $N^P$ the unipotent radical, and $M^P$ the standard Levi for $N^P$. We define the cuspidal Eisenstein series using the Iwasawa decomposition:

$$E(z,f;s) := \sum_{\gamma\in P_{n,n}\backslash\Gamma} \left(\frac{|\det m_1|}{|\det m_2|}\right)^{ns} f(m_1)f(m_2)|_\gamma.$$

This has meromorphic continuation and funcitonal equation, and the poles are all at poles of the constant term.

Langlands showed that $E(z,f;s)$ has constant term $C_P$ along the parabolic $P_{n,n}$ and the constant term vanishes for all other parabolics.

We then consider Arthur's "truncated Eisenstein series," which is something like $E(z,f;s)$ minus stuff from the constant terms. Unfortunately, the "sharp truncation" makes it hard to obtain the Fourier series. To deal with this, we use a "mollifier" to create a "smoothed Arthur truncation" of the Eisenstein series. This serves two purpose:

(1) It is easier to compute the Fourier expansion.
(2) It lives in the square-integrable space.

However, we have to go back and compare this to Arthur's original truncated Eisenstein series. We can establish an identity between the truncated and untruncated Eisenstein series. We have to make sure that the growth terms cancel, and in fact the smoothing is "reverse-engineered" to make this true.

The Fourier expansion is obtained by "Laumon's constant term formula" ♠♠♠ TONY: [?] plus induction. Anything, with all this done, we define an integral sort of analogous to the one from before:

$$I = (\dots)\int \left| \int \widehat{E}_A^*(z,f;1+it)g(A/\beta)\frac{dA}{A}\right|^2 \dots$$

Unfolding exhibits this as a standard Rankin-Selberg type integral.

Now we have to mimic Sarnak's strategy as described earlier. For the upper bound, one shifts the line of integration and uses the pole "moved through."

The lower bound is trickier. Because there are many degenerate terms in the Fourier expansion of the truncated Eisenstein series, it is not orthogonal. A key ingredient is an "orthogonalitiy condition" stating that the degenerate part of the Fourier expansion is orthogonal to the non-degenerate part. That kills off "cross-terms."

Also, one needs to prove an analogous result that "many" of the Hecke eigenvalues are not small. We know that they are "not small on average," so that follows as long as no value is exceptionally large. That is where the tempered hypothesis comes in.

# QUANTUM ERGODICITY ON LARGE GRAPHS

NALINI ANANTHARAMAN

## 1. Motivation

The motivation is understand quantum chaos. It is believed that the spectrum of a quantum chaotic system should look like the spectrum of random matrices. There are many conjectures, the most famous of which is the "Quantum Unique Ergodicity" conjecture.

Since the 90s there has been the idea of using *graphs* as a testing ground or toy model for quantum chaos. Here we focus on the case of *large regular (discrete) graphs*. Let $G_n = (V_N, E_N)$ be a $(q+1)$-regular graph of size $N$, and label $V_N = \{1, \ldots, N\}$.

## 2. Statement of results

We study the eigenvalues of the discrete Laplacian: for a function $f : V \to \mathbb{C}$, we set

$$\Delta f(x) = \sum_{y \sim x} f(y) - f(x) = \sum_{y \sim x} f(y) - (q+1) f(x).$$

So we can write $\Delta = A - (q+1)I$, and the interesting part is $A$.

We will examine the behavior in the limit $N \to \infty$, and we assume a technical condition that $G_N$ has "few short loops." More precisely, if the $\lambda_i$ denote the eigenvalues of the adjacency matrix (so $|\lambda_i| \leq q+1$), then

$$\frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i} \xrightarrow{N \to \infty} \rho(s)\, ds \text{ (Plancherel measure)}$$

is supported on on $[-2\sqrt{q}, 2\sqrt{q}]$. ♠♠♠ TONY: [!]

**Theorem 2.1** (Brooks-Lindenstrauss 2011)**.** *Assume that $G_N$ has "few" loops of length $\leq c \log N$. For any $\varepsilon > 0$, there exists $\delta > 0$ such that for every eigenfunction $\phi$, if $B \subset V_N$ is such that*

$$\sum_{x \in B} |\phi(x)|^2 \geq \varepsilon$$

*then $|B| \geq N^{\delta}$.*

**Theorem 2.2** (Anantharaman-Le Masson, 2013)**.** *Assume that $G_N$ has "few" short loops and that it forms an expander family. Let $(\phi_i^{(N)})_{i=1}^N$ be an orthonormal basis of eigenfunctions of the Laplacian on $G_N$. Let $a = a_N : V_N \to \mathbb{C}$ be such that $|a(x)| \le 1$ for all $x \in V_N$ and $\sum_{x \in V_N} a(x) = 0$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left| \sum_{x \in V_N} a(x) |\phi_i^{(N)}(x)|^2 \right|^2 = 0.$$

What we are doing here is viewing each $\phi_i$ as defining a probability distribution on $V_N$ whose mass at $x$ is $\phi_i^{(N)}(x)^2$. The content of the Brooks-Lindenstrauss result is that this is not concentrated on a small set. The content of our theorem is to compare it to the *uniform* distribution on $V_N$. We do this by evaluating the expectation of a mean-zero function. What we obtain above is only the average over the orthonormal basis, but note that the quantity we are averaging is non-negative, so that means most of the terms are small.

The result can be boosted so that $a$ does not have mean 0 in the expected way.

Now we want to consider a slightly different issue.

**Theorem 2.3** (Anantharaman-Le Masson, 2013)**.** *Assume that $G_N$ has "few" short loops and that it forms an expander family. Let $(\phi_i^{(N)})_{i=1}^N$ be an orthonormal basis of eigenfunctions of hte Laplacian on $G_N$. Let $K = K_N : V_N \times V_N \to \mathbb{C}$ be a kernel satisfying some technical conditions. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left| \sum_{x \in V_N} |\langle \phi_N, K\phi_N \rangle \right|^2 = 0.$$

This recovers the previous theorem in the special case where $K$ is an appropriate multiplication operator.

## 3. Idea of Proofs

There are two proofs.

3.1. **Long proof.** The first one is modeled on the proof of quantum ergodicity on a manifold. At each point, the graph looks "locally like a tree." There is a notion of "Fourier-Helgason transform" on the $(q+1)$-regular tree, which tells us how to define "phase space" for a regular tree or graph. This suggests the right definition of "pseudo-differential" operators on the tree and on a finite regular graph. That turns the "PDE question" into a dynamical systems question.

(1) For $f\colon \mathscr{X} \to \mathbb{C}$, the Fourier transform is
$$\widehat{f}(\omega,s) = \sum_{x\in\mathscr{X}} f(x)\overline{e_{s,\omega}(x)}$$
where $s \in \mathbb{T}_q = \mathbb{R}/(2\pi/\log q)$, $\omega \in \partial\mathscr{X}$, and
$$e_{s,\omega}(x) = q^{(1/2+is)h_\omega(x)}$$
is an eigenfunction of the adjacency matrix. They play the role of "plane waves" on the tree.

There is an inversion formula and an analogue of the Plancherel identity. By a result of Cowling-Setti, the Fourier transform is an isomorphism between "rapidly decreasing" functions $f$ and "smooth" functions $\widehat{f}$ satisfying an appropriate symmetry condition with respect to $s \mapsto -s$.

(2) We define the phase space to be
$$\mathscr{X} \times \partial\mathscr{X} \times \mathbb{T}_q = \mathscr{X} \times \partial\mathscr{X} \times \mathbb{R}/(2\pi/\log q).$$

(3) For a function $a(x,\omega,s)$ on $\mathscr{X} \times \partial\mathscr{X} \times \mathbb{T}_q$, we define an operator $\mathrm{Op}(a)$ on $L^2(\mathscr{X})$ by
$$\mathrm{Op}(a)e_{s,\omega}(x) = a(x,\omega,s)e_{s,\omega}(x).$$
According to the "Paley-Wiener" theorem, we get a correspondence between kernels and "smooth" $a(x,\omega,s)$ with appropriate symmetry condition.

There is a problem that the class of smooth $a(x,\omega,s)$ is not closed under multiplication nor "shifts," but Le Masson constructed a suitable subset closed under these operations.

Now we pass from the tree to a finite regular graph. If $G = \Gamma\backslash\mathscr{X}$ and $a(x,\omega,s)$ is invariant under $\Gamma$, then it descends to $G$, and the corresponding kernel is also $\Gamma$-invariant. There's a problem here that operator $\mathrm{Op}(a)$ is *not* bounded (the kernel doesn't decay fast enough) on $L^2$.

Anyway, the variance
$$V(a) = \sum\langle\phi_i, \mathrm{Op}(a)\phi_i\rangle$$
turns out to be "almost invariant" under shift. Iterating, one gets a Birkhoff sum, whose Hilbert-Schmid norm you can control by known results.

3.2. **Sketch of (short) proof.** Now we give a proof that doesn't rely on the Fourier transform. We consider, as before, the Hilbert space of operators $K(x,y)$ on $\mathscr{X} \times \mathscr{X}$ which is $\Gamma$-invariant, with a certain norm. On $\mathscr{H}$ you can define the self-adjoint operator $K \mapsto [\Delta_{\mathscr{X}}, K]$. This is not the same as $K \mapsto [\Delta_G, K]$ on $HS(L^2(G))$, but it is not very different provided certain assumptions on $K$.

Then there are a couple of lemmas. The first characterizing the kernel of this operator as the closure of the span of powers of the Laplacian. The second estimates the "convergence" of operators, and the result follows from some short calculations afterwards.

I feel that this proof *should* be generalizable to not necessarily regular graphs, but I've encountered technical difficulties in doing so. The "easy part" should be proving analogues of the two lemmas, but even those have proven difficult.

# FROM RAMANUJAN GRAPHS TO RAMANUJAN COMPLEXES

ALEX LUBOTZKY

## 1. Ramanujuan Graphs

Let $X$ be a connected $k$-regular graph and $A$ its adjacency matrix, so $A_{v,u}$ is the number of edges between $u$ and $v$. The eigenvalues of $A$ lie in $[-k,k]$.

*Definition* 1.1. $X$ is called a *Ramanujan graph* if for every eigenvalue $\lambda$ of $A$, either $|\lambda| = k$ or $|\lambda| \leq 2\sqrt{k-1}$.

The point is that all the eigenvalues of magnitude less than $k$ come from the (infinite) simply-connected cover, which is a tree. The eigenvalues control the rate of convergence of the random walk to uniform distribution. Ramanujan graphs have the fastest rate of convergence, and are the "best expanders" by a result of Alon-Boppana.

Do Ramanujan graphs actually exist?

## 2. Explicit construction of Ramanujan graphs

Somewhat surprisingly, explicit constructions were found before indirect proofs (by randomness methods) in this case. Let $p \neq q$ be primes congruent to $1 \pmod 4$. Jacobi proved that

$$r_4(n) = \#\{(x_0,x_1,x_2,x_3) \in \mathbb{Z}^4 \mid \sum x_i^2 = n\} = 8\sum_{\substack{d|n \\ 4\nmid d}} d.$$

Thus $r_4(p) = 8(p+1)$.

For our primes (congruent to $1 \pmod 4$), one $x_i$ is odd and the other three are even. Let $S$ be the subset of such tuples where $x_0 > 0$ is odd (essentially a normalization). Then $|S| = p+1$. We can think of $\alpha = (x_0,x_1,x_2,x_3) \in S$ as an integral quaternion $x_0 + x_1 i + x_2 j + x_3 k$. Note that $\alpha \in S \implies \overline{S} \in S$ and $||\alpha|| = \alpha\overline{\alpha} = p$.

If $q \equiv 1 \pmod 4$, then there exists $\varepsilon \in \mathbb{F}_q$ with $\varepsilon^2 = -1$. For $\alpha \in S$, we define

$$\widetilde{\alpha} = \begin{pmatrix} x_0 + \varepsilon x_1 & x_2 + \varepsilon x_3 \\ -x_2 + \varepsilon x_3 & x_0 - \varepsilon x_1 \end{pmatrix} \in \mathrm{PGL}_2(\mathbb{F}_q).$$

This is essentially an explicit way of splitting the quaternion algebra.

**Theorem 2.1** (Lubotsky-Phillips-Sarnak 1986)**.** *Let $H = \langle \widetilde{\alpha} \mid \alpha \in S \rangle$ and $X^{p,q} = Cay(H, \{\widetilde{\alpha}\})$. Then*

*(1) $X^{p,q}$ is a $(p+1)$-regular Ramanujan graph,*

*(2) If $\left(\frac{p}{q}\right) = -1$ then $H = \mathrm{PGL}_2(\mathbb{F}_q)$ and $X^{p,q}$ is bipartite, and otherwise $H = \mathrm{PSL}_2(\mathbb{F}_q)$ and $X^{p,q}$ is not.*

Where do $X^{p,q}$ actually come from? Let $F = \mathbb{Q}_p$ or $\mathbb{F}_p((t))$. Let $\mathscr{O} = \mathbb{Z}_p$ or $\mathbb{F}_p[[t]]$ (i.e. the ring of integers in $F$). Let $G = \mathrm{PGL}_2(F)$ and $K = \mathrm{PGL}_2(\mathscr{O})$ (the maximal compact subgroup of $G$).

In the archimedean world, if you take a Lie group and mod out by a maximal compact then you get a *symmetric space*, e.g. $\mathrm{GL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$ is the usual upper half-plane. But in the non-archimedean world, $G/K$ is a $p+1$-regular tree, the *Bruhat-Tits tree*. If $\Gamma \leq G$ is a discrete cocompact subgroup (i.e. lattice) then

$$\Gamma \backslash G/K = \Gamma \backslash T$$

is a compact, hence *finite*, $(p+1)$-regular graph

**Theorem 2.2.** *$\Gamma \backslash G/K = \Gamma \backslash T$ is Ramanujan if and only if every infinite dimensional irreducible spherical subrepresentation of $L^2(\Gamma \backslash G)$ (as a G-representation) is tempered.*

Spherical means that there is a non-zero $K$-fixed point. Tempered means that the matrix coefficients are in $L^{2+\varepsilon}$, i.e. "weakly contained in $L^2(G)$."

So the combinatorial property of being Ramanujan is equivalent to a representation-theoretic statement. By the Satake isomorphism, the latter can actually be viewed as being number-theoretic.

**Theorem 2.3** (Deligne)**.** *If $\Gamma$ is an arithmetic lattice of $\mathrm{PGL}_2(\mathbb{Q}_p)$ and $\Gamma(I)$ is a congruence subgroup then every irreducible infinite-dimensional spherical subrepresentation of $L^2(\Gamma(I) \backslash G)$ is tempered.*

**Corollary 2.4.** *$\Gamma \backslash G/K$ is a Ramanujan graph.*

The explicit expanders above are obtained from an especially nice $\Gamma$ (namely, Hamiltonian quaternions). There are similar results by Drinfeld in positive characteristic, and similar constructions by Morganstern for all $k = p^{\alpha} + 1$.

## 3. Ramanujan complexes

The generalization of $T = \mathrm{PGL}_2(F)/K$ is the Bruhat-Tits building

$$B_d(F) = G/K = \mathrm{PGL}_d(F)/\mathrm{PGL}_d(\mathscr{O}).$$

This is a $(d-1)$-dimensional contractible simplicial complex.

The vertices of the building $B_D(F)$ come with "colors" $\nu(gK) \in \mathbb{Z}/d\mathbb{Z}$ defined by

$$\nu(gK) = \mathrm{val}_p(\det g) \pmod{d}.$$

There are "colored adjacency operators" (basically Hecke operators)

$$A_i \colon L_2(B_d(F)) \to L_2(B_d(F))$$

by

$$(A_i f)(x) = \sum_{\substack{y\sim x \\ \nu(y)-\nu(x)=i}} f(y).$$

In particular, the adjacency matrix is $\sum_{i=1}^{d-1} A_i$.

It turns out that the $A_i$ are normal commuting operators (but not self-adjoint, as $A_i^* = A_{d-i}$) hence can be diagonalized simultaneously. Then we can view the spectrum of these operators, denoted $\Sigma_d = \mathrm{Spec}\, A_1, \ldots A_{d-1}$, as a subset of $\mathbb{C}^{d-1}$.

*Definition* 3.1. A finite quotient $\Gamma\backslash B_d(F)$, for $\Gamma$ a co-compact discrete subgroup, is a *Ramanujan complex* if every tuple of nontrivial simultaneous eigenvalues $(\lambda) = (\lambda_1, \ldots, \lambda_{d-1})$ of $(A_1, \ldots, A_{d-1})$ acting on $L^2(\Gamma\backslash B_d(F))$ is in $\Sigma_d$.

**Theorem 3.2** (Li)**.** *If a sequence of quotients $X_i = \Gamma_i\backslash B_d(F)$ has injective radius going to $\infty$, then*

$$\Sigma_d \subset \overline{\bigcup \mathrm{Spec}_{X_i}(A_1, \ldots, A_{d-1})}.$$

*Remark* 3.3. For $d = 2$, we saw that there were two "trivial" eigenvalues. In general, there are $d$ "trivial" eigenvalues, but there is the additional subtlety that they need not have the same magnitude.

**Theorem 3.4** (Lubotsky-Samels-Vishne, 2005)**.** *$\Gamma\backslash B_d(F)$ is Ramanujan if and only if every $\infty$-dimensional irreducible spherical subrepresentation of $L^2(\Gamma\backslash \mathrm{PGL}_d(F))$ is tempered.*

**Theorem 3.5** (Lafforgue, 2002)**.** *If $\mathrm{ch}\, F > 0$ and $\Gamma$ is an arithmetic subgroup of $\mathrm{PGL}_d(F)$ and $\Gamma(I)$ is a congruence subgroup, then (under some restrictions) every subrepresentation of $L^2(\Gamma(I)\backslash \mathrm{PGL}_d(F))$ is tempered.*

**Corollary 3.6.** *In the situation above, $\Gamma(I)\backslash B_d(F)$ are Ramanujan complexes.*

## 4. Overlapping properties

**Theorem 4.1** (Boros-Füredi '84)**.** *Given a set $P \subset \mathbb{R}^2$ of size n, there exists a point $z \in \mathbb{R}^2$ covered by $(\frac{2}{9} - o(1))\binom{n}{3}$ of the $\binom{n}{3}$ triangles determined by P.*

This constant $\frac{2}{9}$ is optimal. Barany proved a higher-dimensional generalization: in $\mathbb{R}^d$, there is a constant $c_d$ such that some point is covered by $c_d\binom{n}{d+21}$ of the $d$-simplices.

Incredibly, Gromov proved that if you allow "curvy" triangles, i.e. any curves as sides, then assertion *still* holds with this same constant $2/9$! He also proved a higher-dimensional version. This changed the point of view, that the property has to do with the combinatorics of simplicial complexes rather than the geometry of Euclidean space.

*Definition* 4.2. A simplicial complex $X$ of dimension $d$ has the $\varepsilon$-*geometric* (resp. *topological*) *overlapping property* if for every $f\colon X(0) \to \mathbb{R}^d$ and every affine (resp. continuous) extension $f\colon X \to \mathbb{R}^d$, there exists a point $z \in \mathbb{R}^d$ covered by $\varepsilon|X(d)|$ of the $d$-cells of $X$.

A family of simplicial complexes of dimension $d$ are *geometric* (resp. *topological*) expanders if all members have the overlapping property with the same $\varepsilon$.

The Boros-Füredi theorem can be re-interpreted as saying that the complete simplicial complex on $n$ vertices is geometrically expanding, and Gromov's result can be re-interpreted as saying that they are even topologically expanding.

Gromov asked: can this hold for simplicial complexes of bounded degree (dimension)?

**Theorem 4.3** (Fox-Gromov-Lafforgue-Naor-Pach 2013)**.** *The Ramanujan complexes of dimension d, when $q \gg 0$, are geometric expanders.*

OK then, what about topological expansion?

**Theorem 4.4** (Kaufman-Kazhdan-Lubotsky 2015)**.** *Fix $q \gg 0$, the* 2*-skeletons of the* 3*-dimensional Ramanujan graph are topological expanders.*

There is basically only one proof of topological expansion, and that is Gromov's notion of "$\mathscr{E}$-coboundary expansion." So we have to prove that the complexes in question are coboundary-expansive, which requires some "isoperimetric inequalities." Unfortunately these are false for the Ramanujan graphs.

After some work, we realized that we could prove some results for cochains of "small support."

# ARITHMETIC ASPECTS OF DIAGONALIZABLE ACTIONS

ELON LINDENSTRAUSS

## 1. Setup

Let $G$ be a semisimple linear algebraic group defined over $\mathbb{Q}$. Let $\mathbb{A} = \mathbb{R} \times \prod'_p \mathbb{Q}_p$ be the ring of adeles (a locally compact topological ring). As is well known, $\mathbb{Q} \hookrightarrow \mathbb{A}$ embeds discretely, and we may consider the quotient $X = G(\mathbb{A})/G(\mathbb{Q})$. This has a finite $G(\mathbb{A})$-invariant measure, which we can normalize to have total mass 1.

A less "fancy" way to set this up is as follows. We chose a finite set $S'$ and let $S = S' \cup \{\infty\}$. Then we consider $G_S = G(\mathbb{R}) \times \prod_{p \in S'} G(\mathbb{Q}_p)$. We have a discrete subgroup (lattice) $\Gamma < G_S$ and we consider $G_S/\Gamma$. This is basically the previous "fancy" setup after forgetting some information.

## 2. Homogeneous dynamics

*Definition* 2.1. An $H$-orbit $H \cdot [g]$ for $H < G_S$ is *periodic* if $Stab_H[g]$ has finite covolume in $H$. Equivalently, there exists an $H$-invariant probability measure $m_{H\cdot[g]}$ supported on $H \cdot [g]$.

We ask the following basic question:

> Suppose $H_i \to H$ and $[g_i] \in G_s/\Gamma$ is $H_i$-periodic. Assume $[g_i] \to [g_\infty]$. Does $m_{H_i\cdot[g_i]}$ converge to a "nice" measure (say in the weak* topology)?

*Remark* 2.2. If these measures to converge to some $\mu$, then it will be invariant with respect to $H$.

### 2.1. Invariant measure *I*.

*Definition* 2.3. A probability measure $\mu$ is $H$-*invariant* if $h_*\mu = \mu$ for all $h \in H$. It is *ergodic* if it is an extreme point in the compact convex set of $H$-invariant measures.

A result of Choquet (the "Ergodic Decomposition Theorem") is that any invariant measure can be expressed as an integral of (invariant) ergodic measures over some probability distribution on these.

**Theorem 2.4** (Ratner, Ratner/Margulis-Tomanov)**.** *Suppose $H < G$ is generated by one parameter of unipotents, $\mu$ is $H$-invariant and ergodic. Then there is a group $L \geq H$ such that $\mu$ is an $L$-invariant measure on a single periodic $L$ orbit.*

♠♠♠ TONY: [i.e. $H$-invariant, ergodic orbits come from periodic orbits]

**Theorem 2.5.** *Suppose, in the notation above, that the groups $H_i$ are non-compact with $[g_i] \in G/\Gamma$ being $H_i$-periodic. If $m_{H_i \cdot [g_i]} \to \mu$, $g_i \to g$ then*

*(1) There exists $L$ such that $L.[g]$ is periodic and $\mu = m_{L \cdot [g]}$, and*

*(2) ...* ♠♠♠ TONY: [more stuff, hard to state - morally a positive answer to the earlier question]

2.2. **Invariant measures II - diagonalizable case.** Let $\mathbb{G}$ be a semisimple group defined over $\mathbb{Q}$ and $G = \mathbb{G}(\mathbb{Q}_S)$. Let $\Gamma < G$ be an arithmetic lattice and $T_0 = G$ a maximal split $\mathbb{Q}_S$-torus.

**Theorem 2.6.** *Suppose $A = T_0 \cap G_1$ with $G_1 \triangleleft G$, $\dim A \geq 2$, and $\mu$ is $A$-invariant and ergodic with support not contained in periodic orbits of reductive proper subgroups. Then either $\mu = m_{G/\Gamma}$ or $\mu$ is "small."*

*Remark* 2.7.

(1) Conjecturally, under the above assumptions the only option should be $\mu = m_{G/\Gamma}$.

(2) If $\dim A = 1$, there exists a big zoo of invariant measures.

(3) This implies that if $\mu$ is $A$-invariant and ergodic on $G/\Gamma$ (with no restriction on the support), then $\mu$ is essentially a product of periodic measures, measures invariant under rank 1 groups, and zero-entropy measures.

(4) The first measure classification theorem in this context is by Katok-Spatzier.

Unfortunately, things go horribly wrong for periodic orbits in this context.

# ASYMPTOTICS OF AUTOMORPHIC SPECTRA AND THE TRACE FORMULA

WERNER MÜLLER

## 1. Automorphic forms and $L$-functions

Maass and Selberg introduced spectral theory into the study of automorphic forms. If $\Gamma \subset \mathrm{SL}(2,\mathbb{R})$ is a lattice acting on $\mathbb{H}$, they considered the hyperbolic surface $\Gamma\backslash\mathbb{H}$ and studied eigenfunctions of the Laplace operator. The Laplace operator $\Delta\colon C_c^\infty(X) \to L^2(X)$ turns out to be essentially self-adjoint, so one wants to study the spectral resolution and properties of $L^2$-eigenfuntions. Interesting questions in this theory concern:

- Existence of Maass cusp form, e.g. Weyl law, Phillips-Sarnak conjecture.
- Location of the spectrum, e.g. Selberg's conjecture $\lambda_1(\Gamma(N)) \geq \frac{1}{4}$.

The main difficulty is that $X$ is non-compact, which implies that there is a *continuous* spectrum (conjecturally $[1/4,\infty)$). This is problematic because the intuition from mathematical physics is that embedded eigenvalues are highly unstable. The basic tool available is Selberg's trace formula.

The modern framework is the adelic one. If $G/\mathbb{Q}$ is a (connected) reductive algebraic group and $\omega\colon Z \to \mathbb{C}^\times$ is a central character, then one considers the representation $L^2(Z(\mathbb{A})G(\mathbb{Q})\backslash G(\mathbb{A}),\omega)$ by the right regular representation. There are distinguished subspaces here such as the space of cusp forms, cut out by the vanishing of certain "Fourier-like" coefficients.

Given an automorphic representation $\pi$ of $G$, to each local unramified $\pi_v$ there corresponds a natural conjugacy class $\phi_p(\pi) \in {}^LG$. Given a representation $r$ of ${}^LG$, one obtains a *Langlands L-function* $L_S(s,\pi,r)$.

## 2. Families of automorphic forms

Sarnak introduced the important idea of *families* of automorphic forms. A family $\mathscr{F}$ is a subset of $\mathscr{A}(G)$ cut out by "natural constaints." One is interested in the statistics of these families, e.g.

For a suitable notion of conductor $c$, what are the asymptotics of
$$\mathscr{F}(x) := \{f \in \mathscr{F} : c(f) < x\}?$$

For instance, one might wonder if it is $O(x^a)$. A precise conjecture of Sarnak is that an asymptotic is given in terms of an associated $L$-function.

## 3. Weyl's Law

Let $(X,g)$ be a compact Riemannian manifold of dimension $n$ and let $\Delta = d^*d\colon C^\infty(X) \to C^\infty(X)$ be the Laplace operator. The spectrum of $\Delta$ will be discrete. Weyl's law predicts the asymptotics of eigenvalues: if $N(x)$ denotes the number of magnitude at most $x$, then

$$\boxed{\lim_{x\to\infty} \frac{N(x)}{x^{d/2}} = (2\pi)^{-d}\omega_d \mathrm{Vol}(\Omega).}$$

There are two methods, based on the heat equation and wave equation. These both break down in the non-compact case, as there is a continuous spectrum (described by the Eisenstein series). Each cusp has a "dual" Eisenstein series, and if you take the Eisenstein series for one cusp and expand it around another cusp, the constant term can be interpreted as a "scattering matrix" which contributes the main part of the continuous spectrum.

Lindenstrauss and Venkatesh proved a Weyl law for $G$ a split adjoint semisimple group over $\mathbb{Q}$, using Hecke operators (again, only available for arithmetic groups). Miller and Müller also established results in other cases.

What about estimating the *remainder term*? Lapid and Müller proved a bound for $\mathrm{SL}(n,\mathbb{R})$. ♠♠♠ TONY: [complicated to state...]

# SOME APPLICATIONS OF TRACE FUNCTIONS IN NUMBER THEORY

PHILLIPE MICHEL

## 1. Motivation

We start with the following classical result ($q$ and $p$ are always primes).

**Theorem 1.1** (Equidistribution of Hecke points)**.** *As $q \to \infty$, the integral points of the closed horocycle of height $1/q$*

$$\left\{\frac{n+i}{q} \mid n = 1, 2, \ldots, q-1\right\}$$

*become equidistributed with respect to the hyperbolic measure.*

A very special case of Sarnak's general "Möbius Disjointness conjecture" is that prime Hecke points become equidistributed with respect to the hyperbolic measure

$$\left\{\frac{p+i}{q} \mid p \text{ prime } < q\right\} \subset \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}.$$

Sarnak and Ubis proved that the weak* limit of the probability measure supported at the prime Hecke points is bounded by $\frac{1}{5}d\mu$ and $\frac{9}{5}d\mu$.

Fouvry, Kowalski, and Michel studied this problem and "reduced" it to the problem of obtaining bounds on Kloosterman sums:

$$\sum_{n \ll q} \lambda_f(n)\mathrm{Kl}_2(n;q) \ll q^{1-\delta}, \quad \delta > 0$$

where the $\lambda_f(n)$ are the Hecke eigenvalues of some weight 0 Hecke eigenform. They eventually proved the desired bound, but later found that there was an error in the calculation for the reductions and it was not so useful for the original problem. However, bounds did yield some corollaries:

**Theorem 1.2** (Fouvry-Kowalski-Michel)**.** *For any modular form $f$,*

$$\frac{1}{q-1}\sum_{n=1}^{q-1} f\left(\frac{n+i}{q}\right)e_q(-n^{-1}) \to 0$$

In other words, a certain signed measure supported on the integral points of the height $1/q$ horocycle flow and weighted by $e_q(-n^{-1})$ converges to 0 (in the weak* topology).

Also:

**Theorem 1.3** (Fouvry-Kowalski-Michel)**.** *As $q \to \infty$, the quadratic Hecke points*

$$\left\{\frac{n^2+i}{q}\right\}$$

*become equidistributed with respect to the hyperbolic measure.*

## 2. Trace Functions

The equidistribution statement is true for a class of functions $F\colon \mathbb{F}_q \to C$ called *trace functions*, obtained as $x \in \mathbb{F}_q \mapsto F(x) = \mathrm{tr}(\mathrm{Frob}_q, \mathscr{F}_q)$ where $\mathscr{F}$ is a constructible middle-extension $\ell$-adic sheaf on $\mathbb{A}^1_{\mathbb{F}_q}$ (i.e. a finite-dimensional $\ell$-adic representation which is pure of weight 0 on the lisse locus - hence Deligne applies - and geometrically irreducible).

To such $\mathscr{F}$ we can associate an integer, the "conductor," measuring the "complexity" of the Galois representation underlyinf $\mathscr{F}$. This has constituents the rank, swan conductor, and something else.

*Example* 2.1. Additive characters of $\mathbb{F}_q$ are associated to an Artin-Schreier sheaf.

*Example* 2.2. Multiplicative characters are associated to a Kummer sheaf.

*Example* 2.3. Kloosterman sums are associated to a (Tate-twisted) Kloosterman sheaf,

$$x \mapsto \mathrm{Kl}_k(x) = \frac{1}{p^{k-1/2}} \sum_{x_1+\ldots+x_k=x} e_q(x_1x_2\ldots x_k?)$$

One can construct new trace functions from existing ones by the usual operations: pullback, dual, tensor product. The pushforward is in general something worse (a complex), but there are two important examples of "nice" pushforward for $Y = \mathbb{A}^2_{\mathbb{F}_q}$.

(1) The Fourier transform

$$\widehat{F}(x) = p^{-1/2} \sum_{y\in\mathbb{F}_q} F(y)e_q(xy)$$

(2) Multiplicative convolution

$$F * G(x) = \sum_{y+z=x} F(y)G(z).$$

The main result making it possible to do analytic number theory with trace functions is

**Theorem 2.4** (Deligne, Weil II)**.** *For $\mathscr{F}$ and $\mathscr{G}$ as above,*

$$\sum_{x\in\mathbb{F}_q} F(x)\overline{G}(x) = \alpha_{\mathscr{F},\mathscr{G}}q + O(C(\mathscr{F})C(\mathscr{G})q^{1/2})$$

*with*

$$|\alpha_{\mathscr{F},\mathscr{G}}| = \begin{cases} 1 & \mathscr{F} \cong_{geo} \mathscr{G}, \\ 0 & \text{otherwise.} \end{cases}$$

The group $\mathrm{PGL}_2(\mathbb{F}_q)$ is the automorphism group of $\mathbb{P}^1_{\mathbb{F}_q}$, hence acts on trace functions by pullback. It's interesting to know when $F$ is isomorphic to $\gamma^* F$, so we want to study $\mathrm{Aut}_{\mathscr{F}}(\mathbb{F}_q)$, which is the subgroup of $\mathrm{PGL}_2(\mathbb{F}_q)$ inducing geometric automorphisms of $\mathscr{F}$. This group can be "classified" in some sense: if $\mathscr{F}$ is not of a specific nice form, e.g. extension of Kummer sheaf by Artin-Schreier sheaf, then its automorphism group is bounded by 60.

*Example* 2.5. When using the Cauchy-Schwarz or Hölder inequality, you often get expressions of the form

$$\sum_{x\in\mathbb{F}_q} F(\gamma_1 x)\overline{F}(\gamma_1' x)\dots F(\gamma_p x)\overline{F}(\gamma_p' x)$$

and thus you want to know whether a certain sheaf contains the trivial representation. This requires knowledge of the geometric monodromy groups, which are usually computed by Katz.

*Example* 2.6. If you want to be able to estimate the sum of a trace function over a short interval, there is a result that if $\mathscr{F}$ is not (geometrically) Artin-Schreier (i.e. not an additive character) then

$$\sum_{1\le n\le X} F(n) \ll \sqrt{q}\log q.$$

This is basically obtained by the Polya-Vinogradov method: use Plancherel's theorem, etc.

This is nontrivial if $X \gg q^{1/2}\log q$, and improvements have been made for smaller $X$, with a result of the form

$$\sum_{1\le n\le X} F(n) \ll \sqrt{q}\log(X)/\log q.$$

## 3. Proof of the bound

We want to prove something like

$$\sum_{n \leq X} F(n)\lambda_f(n) \ll q^{1-1/16+o(1)}.$$

This is equivalent to a subconvex bound for $L(f \otimes \chi, 1/2)$. One considers "amplifying" $f$ in the family of modular forms of level $q$. In the end one obtains matrices $\gamma \in \mathrm{GL}_2(\mathbb{Q}) \cap M_2(\mathbb{Z})$ and "correlation sums"

$$C(\widehat{F}, \gamma) = \sum_{x \in \mathbb{F}_q} \widehat{F}(x)\overline{\widehat{f}}(\gamma \cdot x)$$

and the aim is to show that "often" these sums are bounded by $\ll q^{1/2}$, or equivalently that $\gamma \pmod q$ does not fall into $\mathrm{Aut}_{\widehat{\mathscr{F}}}(\mathbb{F}_q)$, which we do by repulsion argument's (basically Linnik's Lemma).

# RAMANUJAN-SELBERG CONJECTURE

FREYDOON SHAHIDI

## 1. The conjectures

Let $f$ be a normalized Maass form for $\Gamma_0(N)$ and $\Delta$ the Laplace operator. Assume that $f$ is an eigenfunction for $\Delta$ and all Hecke operators. Then we may write

$$\Delta f = \frac{1}{4}(1-s^2)f.$$

The Fourier expansion of $f$ involves Whitaker-Bessel functions and some Hecke eigenavlues $a_p := \frac{1}{2}(\alpha_p + \alpha_p^{-1})$.

**Conjecture 1.1** (Ramanujan)**.**

$$|a_p| \leq 2p^{-1/2} \iff \alpha_p = 1.$$

In 2003, Kim-Sarnak showed that $|a_p| \leq p^{-1/2}(p^{7/64} + p^{-7/64})$, or equivalently $p^{-7/64} \leq |\alpha_p| \leq p^{7/64}$. Their method worked only over $\mathbb{Q}$, since one has to work with unit groups. In 2011 Blomer-Brumley extended this to any number field.

Selberg conjectured that every eigenvalue $\lambda \geq \frac{1}{4}$ for $\Gamma$ a congruence subgroup. In that same paper, Kim-Sarnak showed that $\lambda \geq 1/4 - (7/64)^2 = 0.238\ldots$.

Any general progress (Maass forms, arbitrary number fields) has used functoriality and the existence of symmetric power lifts, i.e. functoriality for $\mathrm{Sym}^m\colon\ \mathrm{GL}_2(\mathbb{C}) \to \mathrm{GL}_{m+1}(\mathbb{C})$. The Blomer-Brumley breakthrough follows from $\bigwedge^2(\mathrm{Sym}^3\pi)$ being automorphic, hence $\mathrm{Sym}^4\pi$ is automorphic (it is essentially a four-dimensional constituent).

The automorphic input is this theorem.

**Theorem 1.2.** *Let $\pi$ be a cuspidal automorphic representation of* $\mathrm{GL}_2(\mathbb{A}_F)$*, where $F$ is a number field. Then $L(s, \mathrm{Sym}^4\pi, \mathrm{Sym}^2)$ is absolutely convergent for* $\mathrm{Re}\,(s) > 1$.

These sorts of results are very deep. The strategy goes back to ideas of Jacquet and Shalika. One considers Rankin-Selberg $L$-functions, e.g.

$$L(s, \mathrm{Sym}^4\pi \times \mathrm{Sym}^4\pi) = L(s, \mathrm{Sym}^4\pi, \mathrm{Sym}^2) \cdot L(s, \mathrm{Sym}^4\pi, \bigwedge^2 \pi).$$

And we have the identity

$$L(s, \mathrm{Sym}^4, \overset{2}{\bigwedge}) = L(s, \mathrm{Sym}^3 \pi, \mathrm{Sym}^2 \otimes \omega)$$

where $\omega$ is the central character of $\pi$. There is then a game of deducing things using these "incidental equalities."

Using known results about automorphicity of specific representations, one deduces an equalty of the thing in interest with a Jacquet-Godement $L$-function. By the work of Jacquet-Shalika, we can deduce that this thing is absolutely convergent for $\mathrm{Re}(s) > 1$.

Now let $k$ be a number field, $G$ a quasi-split reductive group over $k$, and $\pi$ a cuspidal representation in $L^2(G(k)\backslash G(\mathbb{A}))$ factoring as $\pi = \otimes_v \pi_v$. There is a hypothetical global Langlands group $L_k$, which we do not know how to define.

Ramanujan conjectured that each $\pi_v$ should be tempered. Unfortunately, this is not true in general, but it is expected to be true for $\mathrm{GL}_n$.

**Conjecture 1.3.** *Let $\pi$ be a cusp form on $G$, quasi-split over $k$. Assume that $\pi$ is (globally) generic, i.e. has a non-zero Whittaker Fourier coefficient. Then $\pi$ is tempered.*

To address the issue of temperedness, Arthur introduced the notion of *Arthur parameters*. These are homomorphisms $\psi\colon L_K \times \mathrm{SL}_2(\mathbb{C}) \to {}^L G$ where $\psi(w,g) = \phi(w)\rho(g)$, where $\rho$ is a complex-analytic morphism. Every automorphic representation is supposed to have an Athur parameter. The tempered representations are parametrzied by $\psi$ for which $\rho = 1$. To each $\psi$ one can attach a Langlands parameter by

$$w \mapsto \psi(w, \begin{pmatrix} |w|^{1/2} & 0 \\ 0 & |w|^{-1/2} \end{pmatrix}).$$

## 2. PROGRESS

The local conjectures are proved in many cases (work of Ban-Liu, Jian-Soudry, Harris-Taylor, Henniart, Scholze, Arthur). They can be reduced to certain questions concerning the Local Langlands Conjectures. There is a quite a bit of evidence that the local conjecture is true, by demonstration for many classical and exceptional groups.

Another application of the Local Langlands Conjectures is the equality of root numbers. This is more difficult. For a Weil-Deligne representation $\rho\colon W_F' \to \mathrm{GL}_n(\mathbb{C})$, we get an $\varepsilon$-factor and an $L$-function. According to the LLC, we also get some other $L$-function and $\varepsilon$-factors, and Cogdell-Shahidi-Tsai show agreement between the stuff for the Langlands-Shahidi

$L$-functions and Artin $L$-functions:

$$\varepsilon(s, \bigwedge^2 \circ \rho, \psi) = \varepsilon(s, \pi(\rho, \bigwedge^2))$$

and also for the $L$-functions. Actually the agreement of $L$-functions was proved by Henniart, but the root numbers are the hard part.

The trick is to induce from a supercuspidal and compare the $\gamma$ functions.

# THE $p$-ADIC LANGLANDS PROGRAM: MOTIVATIONS AND APPLICATIONS

MATTHEW EMERTON

## 1. INTRODUCTION

Algebraic number theorists are interested in an aspect of the Langlands program called *Langlands reciprocity*. This is supposed to be a connection between *automorphic forms* and *Galois representations*.

A famous instance of this is the following: if $f$ is a Hecke eigenform of weight $k$, level $N$, and character $\varepsilon$, then for a prime $\ell \nmid N$ we have a Hecke polynomial $X^2 - a_\ell X + \varepsilon(\ell) l^{k-1}$. There are different ways of thinking about this, but one is that it is the characteristic polynomial of a semisimple conjugacy class in $\mathrm{GL}_2(\mathbb{Q}_\ell)$. We want to think of that conjugacy class as coming from a Frobenius element in the (absolute) Galois group. But which Galois representation does this come from?

If $E$ is an elliptic curve over $\mathbb{Q}$ of conductor $N$, and $\ell \nmid N$, then

$$\#E(\mathbb{F}_\ell) = 1 - a_\ell + \ell$$

which can be thought of the value at $X = 1$ of $X^2 - a_\ell X + \ell^{2-1}$. This is a hint of the connection between the automorphic side and diophantine side.

**Theorem 1.1** (Wiles, Taylor-Wiles, ... )**.** *For each such $E/\mathbb{Q}$, there is a Hecke eigenform $f$ of level $N$ of weight 2 such that $a_\ell(E) = a_\ell(f)$.*

The $p$-adic Langlands program grew out of an attempt to understand more deeply the mathematics involved in this theorem. The proof proceeds via Galois representtions. Deligne showed that for $f$ defined over a number field (i.e. the coefficients $a_\ell$ lie in a number field) equipped with an embedding into $\overline{\mathbb{Q}_p}$, there is a representation

$$\rho_f \colon G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}_p})$$

unramified at $\ell \nmid pN$. The characteristic polynomial of $\rho(\mathrm{Frob}_\ell)$ is then precisely the Hecke polynomial $X^2 - a_\ell X + \varepsilon(\ell) l^{k-1}$.

On the other hand, an elliptic curve $E/\mathbb{Q}$ has a Tate module

$$T_p[E] = \varprojlim_n E[p^n] \cong \mathbb{Z}_p^2.$$

We can extend scalars to $\overline{\mathbb{Q}_p}$ to obtain a two-dimensional representation

$$\rho_E \colon G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}_p}).$$

The modularity of $E$ is proved by showing that $\rho_E \cong \rho_f$.

## 2. Families of representations

Now there is a key point here: the representations $\rho \colon G_{\mathbb{Q}} \to \mathrm{GL}_n(\mathbb{Q}_p)$ "live in families." Think about it like this: suppose that $G_{\mathbb{Q}}$ were finitely presented. Then a representation would be a collection of matrices satisfying certain equations. Now suppose instead that $G_{\mathbb{Q}}$ is the completion of such a group: then the matrices also have to satisfy certain *inequalities*. So the space of representations is a kind of space defined by inequalities, and $\rho_E$ is some point of the space (which turns out to be 3-dimensional). We want to show that $\rho_E$ is one of the countably many points coming from eigenforms.

An idea introduced by Mazur, and developed by Taylor-Wiles, is that it would be nice to enlarge the space of automorphic forms to a 3-dimensional "box" as well. We could then worry about whether or not our particular point comes from an appropriate "classical" point of this box.

> **Goal:** extend the notion of automorphic forms to allow $p$-adic variations.

There are a lot of difficulties. The first one is that automorphic forms are defined as complex functions on a symmetric space. However, it is difficult to connect this to the $p$-adic world. It's somewhat of a miracle that they tend to admit algebraic descriptions; in fact this isn't even known in the generality we suspect, e.g. for certain Maass forms. So why is it true for classical modular forms anyway?

There are many ways to see this, but one of them is from the perspective of Eichler-Shimura: automorphic forms can be described as *cohomology classes*. This is Hodge Theory, or what algebraic number theorists call Eichler-Shimura theory.

## 3. Completed Cohomology

Now the powerful thing about cohomology is that you can take $\mathbb{Z}$-coefficients, so you can immediately see that the Hecke eigenvalues will be algebraic integers. So our first attempt could be to take $\mathbb{Z}_p$ coefficients. This is a crude first step, as it doesn't really "enlarge" the space of automorphic forms meaningfully.

Observe that ramification away from $p$ is pretty rigid, because the wild inertia at $\ell \neq p$ is pro-$\ell$, and can't interact much with $\overline{\mathbb{Q}_p}$ which is nearly pro-$p$. But at $p$ things are less rigid, so we are motivated to consider varying the level by powers of $p$.

Let $G(\mathbb{R})$ be a semisimple algebraic group and $X = G(\mathbb{R})/(\text{maximal compact})$. Let $\Gamma$ be a congruence subgroup, and consider $\Gamma(p^r)$ for all $r \geq 0$. We have towers

$$\ldots \to X/\Gamma(p^r) \to \ldots \to X/\Gamma.$$

We can study the homology of this tower:

$$\widehat{H}_i := \varprojlim_r H_i(X/\Gamma(p^r), \mathbb{Z}_p)$$

which we call the *completed cohomology* of the tower. Note that this can be quite large even though the constituents are countable. It has the structure of a $\mathbb{Z}_p$-module, but even better, it has the action of Hecke operators $T_\ell$ for all $\ell \neq p$. Think of the completed cohomology as being like a Banach space and $\mathbb{T} = \mathbb{Z}_p\langle T_\ell \rangle$ as being like a von Neumann algebra.

$G(\mathbb{Z}/p^r)$ acts as the deck transformations of $X/\Gamma(p^r)$ over $X/\Gamma$. In the limit, $\Lambda = \varprojlim \mathbb{Z}_p[G(\mathbb{Z}/p^r)] =: \mathbb{Z}_p[[G(\mathbb{Z}_p)]]$ acts on the completed cohomology.

*Example* 3.1. If $G$ were abelian, e.g. $\mathbb{G}_a$, then

$$\mathbb{Z}_p[G] = \mathbb{Z}_p[[\mathbb{Z}_p]] \cong \mathbb{Z}_p[[T]].$$

If $G$ were abelian then this would be a power series ring. But $G$ is semisimple, so it's more like a twisted power series ring, but it's like a universal enveloping algebra. (Lazard showed that $\Lambda$ is Noetherian.)

*Remark* 3.2. Actually, even $G(\mathbb{Q}_p)$ acts on $\widehat{H}_i$.

*Example* 3.3. If we take the tower of $S^1$, then $\widehat{H_0} = \mathbb{Z}_p$ (inverse limit of $\mathbb{Z}_p$ by identity) and $\widehat{H_1} = 0$ (intersection of $p^r\mathbb{Z}_p$ over all $r$). One might worry that this limiting process kills off the interesting features, rather than enlarge them. However, it turns out that this doesn't happen, and you can in fact recover the cohomology at any finite level. By some spectral sequence argument,

$$H_i(\Gamma(p^r), \mathbb{Z}_p)\text{``}=\text{''}\Gamma(p^r) - \text{coinvariants in } \widehat{H}_i = \mathbb{Z}_p[G(\mathbb{Z}/p^r)] \otimes_\Lambda \widehat{H}_i.$$

This is only approximately correct; there are correction terms coming from higher Tors. The reason in the above example is that $\mathbb{Z}_p$ is actually a torsion $\Lambda$-module.

If $W$ is a $G$-representation, $W(\mathbb{Z}_p)$ is a $\Gamma$-representation, so it gives local systems on each $X/\Gamma(p^r)$, and in particular on $X/\Gamma$. Roughly,

$$H_i(\Gamma, W)\text{``}=\text{''}W \otimes_\Lambda \widehat{H}_i.$$

Again there are correction terms. The intuition is that if we fix $r$, then $\Gamma(r)$ eventually acts trivially on $W(\mathbb{Z}/p^n)$ for $n \gg 0$, which suggests that one should be able to pull out the $W$. The tensor product then corresponds to "throwing back in" the low level data.

*Example* 3.4. Let $G = \mathrm{SL}_2(\mathbb{Q})$ and $X = \mathbb{H}$. Then $X/\Gamma(p^r)$ is an open modular curve. Then there is an exact sequence

$$0 \to \widehat{H}_i \to \Lambda^n \to \Lambda \to \mathbb{Z}_p \to 0.$$

If there were no $\mathbb{Z}_p$, then this sequence would split, and we would obtain $\widehat{H}_i \cong \Lambda^{n-1}$. Now the $\mathbb{Z}_p$ is not zero, but it's "small" from the perspective of $\Lambda$, as the latter is like a (twisted) power series ring over $\mathbb{Z}_p$ in 3 variables (since we are considering $\mathrm{SL}_2$). So let's pretend that the $\mathbb{Z}_p$ isn't there. Essentially, $\widehat{H}_1 = \Lambda^n$, $H_1(\Gamma(p^r)) = \mathbb{Z}_p[\mathrm{SL}_2(\mathbb{Z}/p^r)]^n$, $H_1(W) = W^n$. One thing grows like the volume, and the other grows like the covolume, which are heuristics expected from the analytic theory. ♠♠♠ TONY: [??]

**Theorem 3.5** (Calegari-Emerton)**.** *If $G(\mathbb{R})$ does not admit discrete series, then $\widehat{H_1}$ is torsion over $\Lambda$.*

**Corollary 3.6.** *In this situation, for some $\varepsilon > 0$*

$$\dim H_i(\Gamma(p^r)) = (covolume)^{1-\varepsilon}$$

The idea is to exploit the algebraic structure. If $\widehat{H_1}$ is free, or at least has a large free part, then the growth is like the covolume. On the other hand, if everything is torsion then you actually get a power savings.

**Theorem 3.7** (Marshall)**.** *If $G = \mathrm{SL}_2(F)$ where $F$ is not totally real, then for some $\varepsilon > 0$*

$$\dim H_1(W) \leq \dim(W)^{1-\varepsilon}.$$

# SOBOLEV TRACE INEEQUALITIES

ALICE CHANG

## 1. Introduction

Let $(M,g)$ be a compact manifold with Riemannian metric $g$. Let $\Delta_g$ denote the Laplace Beltrami operator on $M$ and $0 \le \lambda_0 \le \lambda_1 < \dots$. The *determinant* is

$$\det \Delta_g = \prod \lambda_j.$$

One can define a zeta function by the Mellin transform:

$$\zeta(s) = \frac{1}{\Lambda(s)} \int \operatorname{tr}(e^{t\Delta_g}) t^s \frac{dt}{t}.$$

We define the *height* function to be $\log \Delta_g(0) = \zeta'(0)$.

### 1.1. Polyakov-Alvarez formula.

On compact $(M^n, g)$ we have

$$\operatorname{tr}(e^{-t\Delta_g}) \sim \sum_{k=0}^{\infty} a_k t^{(k-n)/2} \text{ as } t \to 0$$

where $a_0$ is essentially the volume, $a_1$ is essentially the area of the boundary, and $a_2$ is essentially the Euler characteristic.

Polyakov observed that the height function is a "conformal primitive" of the Gaussian curvature.

There's some crazy formula for small variation of the height function in terms of the Gaussian curvature. The Polyakov-Alvarez formula turns out to be another crazy formula for a ratio(?) of height functions in terms of more integrals of Gaussian curvature, which I won't reproduce.

### 1.2. Work of Osgood-Philips-Sarnak.

**Theorem 1.1** (Osgood-Phillips-Sarnak)**.**

1. *An isospectral set of closed* 2*-manifolds is compact in the* $C^\infty$ *topology.*
2. *An isospectral set of planar domains is compact in the* $C^\infty$ *topology.*

Some key ideas:

- on manifolds, each heat coefficient $a_k$ controls the $W^{k,2}$ norm of the heat module. In the positive curvature case, one needs to study the extremal metrics in the formula and establish a sharp inequality.

- On $(S^2, g_0)$, if $\mathrm{vol}(g_w) = \mathrm{vol}(g_0)$ then there's some lower bound on the integral of stuff involving the Gaussian curvature. The point of this is to establish that the canonical metric is extremal.
- On the plane disk, they stablish the classical Milin-Lebedev inequality (essentially a Sobolev trace inequality) in a new way.

How do you generalize this to $n > 2$? It seems nearly impossible. We used $n = 2$ in two crucial ways:

- Any two metrics on $S^2$ are conformal, i.e. $g_1 = e^{2w} g_2$,
- The Laplace-Beltrami operator transforms in a very nice way with respect to conformal transformations: $\Delta_{g_w} = e^{-2w} \Delta_g$

**Theorem 1.2** (Okikiolu)**.** *On* $(S^3, g_0)$*,* $\det(\Delta_{g_0})$ *is a local maximum among all metrics* $g$ *with the same volume as* $g_0$*.*

Recall that an operator $A$ is "conformally covariant" of bidegree $(a, b)$ if

$$A_{g_w}(\phi) = e^{-bw} A_g(e^{aw}\phi)$$

for all $\phi \in C^\infty(M)$. We will study such operators.

*Example* 1.3. On $M^2$, $\Delta_g$ is conformal of bidegree $(0, 2)$. These operators have been extensively studied in higher dimensions too.

♠♠♠ TONY: [overwhelmed...]

# PROBLEMS ON POINTS AND LINES

BEN GREEN

## 1. Introduction

We discuss a classic problem of Sylvester: Take a finite set $P \subset \mathbb{R}^2$, not all on a line. Is there necessarily an *ordinary* lane: a line that passes through exactly two points of $P$?

*Example* 1.1. Equilateral triangle, with midpoints and centroid. There is an obvious example.

**Theorem 1.2** (Sylvester-Gallai Theorem)**.** *The answer is yes.*

*Kelly's Proof.* There is at least one point lying off a line through two other points. We may further assume that amongst all such pairs, the perpendicular distance of this arrangement is minimal. We claim that this line is ordinary, since otherwise we would find a smaller perpendicular distance. □

Here is a refinement of the problem:

| |
|---|
| If $\lvert P \rvert = n$, how many ordinary lines must it have? |

A random arrangement turns out to be bad, with $O(n^2)$ ordinary lines. What if we want to make a set with few ordinary lines? A naïve example is $n-1$ collinear points and another point, so there are obviously $n-1$ ordinary lines.

There is an example due to Böröczky that if $n$ is even, we can make examples with $n/2$ ordinary lines. The more natural setting for these examples is the projective plane. (The questions are evidently equivalent in affine and projective plane.)

*Example* 1.3. We take $n = 12$: 6 points of a regular hexagon inscribed in a circle, and 6 points on the line at infinity corresponding to the chordal directions. The ordinary lines are the tangents to the circle at points on the hexagon.

When $n$ is odd, there are various ways of adding or deleting a point to get $\frac{3n}{4} + o(1)$ ordinary lines.

**Theorem 1.4** (Green-Tao)**.** *Suppose $n \geq 10^{10^{10}}$. Then any set of $n$ non-collinear points has at least*

$$\begin{cases} n/2 & n \text{ even} \\ 3n/4 + o(1) & n \text{ odd} \end{cases}$$

*and the Böröczky examples, up to projective equivalence, are the only equality cases.*

*Example* 1.5. Cubic curves also lead to examples of sets with few ordinary lines. Indeed, if you take a small (torsion) subgroup, then a line through two points will tend to pass through a third, unless there is a degeneracy (tangent line).

What's going on turns out to be that there is more algebraic structure here. Notice that that all the "minimal" examples we have discussed lie on cubic curves (Böröcity's example is a degenerate cubic).

**Theorem 1.6.** *Suppose $k \geq 1$ is fixed. Let $n \geq n_0(k)$. Then any set $P \subset \mathbb{R}^2$ of $n$ points with at most $kn$ ordinary lines lies on the union of $\leq k$ cubic curves ... plus a few degenerate cases.*

To deduce the preceding theorem, you apply this theorem with $k = 1$ to deduce that any such configuration lies on a cubic curve, and then there is some (difficult) accounting.

*Remark* 1.7. The "converse" is definitely not true. A random selection of points on a cubic will have many ordinary lines.

## 2. Idea of the proof

The basic idea is to study a proof of Sylvester's theorem and see what can be leveraged from it. Unfortunately we haven't been able to make any progress with Kelly's proof, but we instead looked at Melchor's proof. This starts by looking at the projective dual $P^*$ of $P$. Euler's formula says

$$V - E + F = 1$$

for the projective dual. Now we digest what these all mean.

(1) $V = \sum_k N_k$, where $N_k$ is the number of $k$-rich lines in $P$ (lines with $k$ points on them; this is just counting each edge once)
(2) $2E = \sum_k 2kN_k$ (this is the graph theoretic assertion that $2E$ is the sum of the degrees, but the degree of $\ell^*$ is twice the number of points on $\ell$)
(3) $F = \sum_s M_s$ ($M_s$ is the number of faces with $s$ edges)
(4) $2E = \sum_s sM_s$.

This gives

$$N_2 - 3 = \sum_{k \geq 3} (k-3) N_k + \sum_{s \geq 3} (s-3) M_s.$$

We immediately see that $N_2 \geq 3$. But we get more information. If $N_2$ is small, then the right hand side is small too, so "most" faces are triangles, and most vertices have degree 6.

The idea is that if $P$ has few ordinary lines, then $P^*$ is "locally a triangular lattice." We then want to show that being a triangular lattice in the dual means that the points lie on a cubic curve.

This relies on two classical facts.

(1) There exists a cubic curve through any 9 points in the plane.
(2) (Chales' Theorem) Any cubic curve passing through 8 of the points of intersection of two triples of lines, passes through the 9th.

The idea is basically that every hexagon on the dual side corresponds to a Cayley-Bacharach configuration on the other side. So you keep picking 9 points and a cubic curve through 8 of them, and you automatically get that the 9th is also on the curve.

Now we'll hint at how additive combinatorics comes into deducing the main theorem. So we have to analyze what happens on cubic curves. Suppose, for instance, that the cubic curve is a union of a line and a parabola, e.g.

$$P = \{(0,b) : b \in B\} \cup \{(a,a^2) : a \in A\}$$

where $|A| = |B| = n/2$. There is a "semblance" of a group law here: you can take two points on the parabolic and get a point on the line. If $P$ has few arbitrary lines, almost all pairs $a_1, a_2 \in A$ have $-a_1 a_2 \in B$. We say that $A$ is an "approximate group." There is a structure theory of approximate groups, which in this case says that they are very close to actual groups. Now $\mathbb{R}^\times$ has no interesting (finite) subgroups, so it turns out that there are no approximate groups either.

## 3. Further Speculations

I'll describe at a few thing that I think are true, but don't know how to prove.

> Suppose $P \subset \mathbb{R}^2$ and $|P| = n$, with no more than 100 points on a line. Suppose there are $\geq \delta n^2$ collinear triples. Is it the case that there are some $\delta n$ points on a cubic?

Now I'll make another speculation concerning where cubic structure arises from combinatorial structure.

Suppose $P \subset \mathbb{F}_p^2$ is a set with no three points on a line. Is it the case that $P$ lies on a cubic apart from $o(p)$ points? (This is false if the cardinality is not prime!)