Evidence for a Spectral Interpretation of the Zeros of Families of Elliptic Curves

Steven Miller

Princeton University

Pisa, June 13, 2002

Introductory Notes

Lot of thesis is book-keeping: don't want to spend an hour keeping track of error terms.

Want to briefly explain the problem, history and set-up – what are the main differences between this and density investigations of other families.

Concentrate on sketching the proof of the main theorem (for $n = 1$): show the new problem that arises (the variation of the conductors), and what we need to fix it.

Give some examples: conditions for the theorem are very easy to check, fun to make high rank examples.

Applications: (1) checking the philosphy of Katz-Sarnak on a much slimmer family; (2): better estimates for excess rank (rank above the family rank); (3): observe potential lower order density terms.

Fundamental Problem: Spacing Between Events

General Formulation: Studying some system, observe values at t_1, t_2, t_3, etc. Question: what rules govern the spacings between events?

Often need to normalize by average spacing.

Example 1: Spacings Between Primes / Prime Pairs.

Example 2: Spacings Between Energy Levels of Nuclei.

Example 3: Spacings Between Eigenvalues of Matrices.

Example 4: Spacings Between Zeros of L-Functions.

Elliptic Curves

Consider $E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, $a_i \in \mathbf{Q}$.

Local data: number of solutions (x, y) mod p. Use to build the L-function.

Modularity Theorem [Wiles]: $L(s, E) = L(s, f)$ for a weight 2 cuspidal newform of level N_E.

$$\Lambda(s, E) = (2\pi)^{-s}N^{s/2}\Gamma(s + \frac{1}{2})L(s, E) = \epsilon_E\Lambda(1 - s, E)$$

By GRH: All zeros on the critical line. Makes sense to talk about spacings between zeros.

Rational solutions form a group, write as $E(\mathbf{Q}) = \mathbf{Z}^r \oplus T$, T are the torsion points, r is the geometric rank.

Birch and Swinnerton-Dyer Conjecture: Geometric rank equals the analytic rank, the order of vanishing of $L(\frac{1}{2}, E)$.

Random Matrix Theory

Consider the group of $N \times N$ matrices from one of the classical compact groups: unitary, symplectic, orthogonal.

One assigns probability measures to matrices from various groups. By explicitly calculating properties associated to an individual matrix and integrating over the group, one can often use the group average to make good predictions about the expected behaviour of statistics from a generic, randomly chosen element.

More generally, can consider other spaces: GUE / GOE: Hermitian / Symmetric matrices with Gaussian probabilities for entries.

Measures of Spacings: n-Level Correlations

Let $\{\alpha_j\}_{j=1}^N$ be an increasing sequence of numbers. For a compact box $B \subset \mathbf{R}^{n-1}$ we define the n-level correlation by

$$\frac{\#\Big\{(\alpha_{j_1}-\alpha_{j_2},\dots,\alpha_{j_{n-1}}-\alpha_{j_n}) \in B, j_i \neq j_k\Big\}}{N} \qquad (0.1)$$

Instead of using a box, can use a test function.

1. $f(x_1,\dots,x_n)$ is symmetric
2. $f(x+t(1,\dots,1)) = f(x)$; ie, f is a function of differences.
3. $f(x) \to 0$ rapidly as $|x| \to \infty$ in the hyperplane $\Sigma_j x_j = 0$.

$$\frac{1}{N}\sum_{\substack{j_1,\dots,j_n\\ distinct}} f(\alpha_{j_1},\dots,\alpha_{j_n})$$

$f(x_1,\dots,x_n) = \chi_B(x_1-x_2,\dots,x_{n-1}-x_n)$ recovers Eq. 0.1; χ_B the characteristic function of the box.

Note the above is over distinct indices. As $N \to \infty$, the contribution from any finite number of indices becomes negligible.

Measures of Spacings: n-Level Correlations (Continued)

In an impressive set of computations, starting with the $10^{20\text{th}}$ zero of $\zeta(s)$, Odlyzko studied the normalized spacings between adjacent zeros and found remarkable agreement with Random Matrix Theory. Specifically, consider the set of $N \times N$ random Hermitian matrices with entries chosen from the Gaussian distribution (the GUE). As $N \to \infty$, the limiting distribution of spacings between adjacent eigenvalues is indistinguishable from what Odlyzko observed!

Montgomery proved that the 2-level correlation for $\zeta(s)$ is the same as that of the GUE, and Rudnick-Sarnak showed the n-level correlations for all automorphic cupsidal L-functions are the same as the GUE.

The universality that Rudnick and Sarnak observed is somewhat surprising, but explainable as follows: the correlations are controlled by the second moments of the a_p's, and while there are many possible limiting distributions for the a_p's, they all have the same second moment.

Katz and Sarnak prove the n-level correlations of all the classical compact groups are the same as $N \to \infty$.

Measures of Spacings: n-Level Density

Let $f(x) = \Pi_i\, f_i(x_i)$ be compactly supported. Assuming $f(x)$ is a product for simplicity. Consider

$$D_{n,E}(f) \;=\; \sum_{\substack{j_1,\dots,j_n \\ distinct}} f_1(L_E\gamma_E^{(j_1)}) \cdots f_n(L_E\gamma_E^{(j_1)}). \quad (0.2)$$

Unlike n-level correlations, possible for a fixed number of zeros to contribute in the limit. For f of compact support, only first few zeros contribute.

In many instances, the behaviour of $L(\frac{1}{2}, f)$ encodes critical information about the function. For example, for L-functions of elliptic curves, the order of vanishing of $L(s, E)$ at $s = \frac{1}{2}$ is conjecturally equal to the geometric rank of the Mordell-Weil group.

If we force the Mordell-Weil group to be large, we expect many zeros at $s = \frac{1}{2}$, and this might influence the behaviour of the neighboring zeros. Hence we are led to study the distribution of the first few, or low lying, zeros.

Similar to choosing an $N \times N$ matrix at random and calculating its eigenvalues, we only get one string of values. If, however, we can find a large number of curves similar to our original one, then we may calculate the zeros of each, and see how they vary from curve to curve.

Families

This leads us to the concept of *family*. Roughly, a family will be a collection of geometric objects and their associated L-functions, where the geometric objects have similar properties.

Iwaniec, Luo and Sarnak considered (among others) all cuspidal newforms of a given level and weight. Rubinstein considered twists by fundamental discriminants $D \in [N, 2N]$ of a fixed modular form.

I studied the family of all elliptic curves and one-parameter families of elliptic curves.

To any geometric family, Katz-Sarnak predict the n-level density depends only on a symmetry group attached to the family. Based on the function field case, for typical elliptic curve families they predict orthogonal symmetries. One can further analyze the distributions depending on the signs of the functional equations. As our elliptic curve families are self-dual, we expect the densities to be controlled by the distribution of signs (all even: SO(even); all odd: SO(odd); equidistributed: O).

Normalization of Zeros

How should we normalize the zeros of the curves in our family? Two choices: (1) locally (using some natural measure from that curve); (2) globally (using some natural measure from the family).

Hope: for f a good even test function with compact support, as $|\mathcal{F}| \to \infty$,

$$\begin{aligned}\frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}} D_{n,E}(f) &= \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\sum_{\substack{j_1,\dots,j_n\\ j_i\neq\pm j_k}}\prod_i f_i\Big(\frac{\log N_E}{2\pi}\gamma_E^{(j_i)}\Big)\\ &\to \int\cdots\int f(x)W_{n,\mathcal{G}(\mathcal{F})}(x)dx\\ &= \int\cdots\int \widehat{f}(u)\widehat{W_{n,\mathcal{G}(\mathcal{F})}}(u)du.\end{aligned}$$

In all results below, we obtain the same results (with significantly easier proofs and no sieving) if instead of rescaling by $\log N_E$ we rescale by the average log-condcutor

$$\log M = \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\log N_E$$

Much of the work is handling the dependence on the conductors.

1-Level Densities

Katz and Sarnak calculate the n-level densities for the classical compact groups. Unlike the correlations, the densities are different for different groups.

The Fourier Transforms for the 1-level densities are

$$\begin{aligned}
\widehat{W_{1,O^+}}(u) &= \delta_0(u) + \frac{1}{2}\eta(u) \\
\widehat{W_{1,O}}(u) &= \delta_0(u) + \frac{1}{2} \\
\widehat{W_{1,O^-}}(u) &= \delta_0(u) - \frac{1}{2}\eta(u) + 1 \\
\widehat{W_{1,Sp}}(u) &= \delta_0(u) - \frac{1}{2}\eta(u) \\
\widehat{W_{1,U}}(u) &= \delta_0(u)
\end{aligned}$$

where $\delta_0(u)$ is the Dirac Delta functional and $\eta(u)$ is 1, $\frac{1}{2}$, and 0 for $|u|$ less than 1, 1, and greater than 1.

Note the three orthogonal densities are indistinguishable for $|u| < 1$. Hence, for test functions supported in $(-1, 1)$, we cannot differentiate the three orthoganal densities, although we can differentiate them from the other densities. Note: the 2-Level Density *does* distinguish the orthogonal groups, even with test functions of arbitrarily small support.

Explicit Formula

Starting Point is the Explicit Formula, which relates sums of test functions over zeros to sums over primes of $a_E(p)$ and $a_E^2(p)$.

$$\begin{aligned}\sum_{\gamma_E^{(j)}} G\Big(\frac{\log N_E}{2\pi}\gamma_E^{(j)}\Big) &= \widehat{G}(0) + G(0) \\ &- 2\sum_p \frac{\log p}{\log N_E}\frac{1}{p}\widehat{G}(\frac{\log p}{\log N_E})a_E(p) \\ &- 2\sum_p \frac{\log p}{\log N_E}\frac{1}{p^2}\widehat{G}(\frac{2\log p}{\log N_E})a_E^2(p) \\ &+ O(\frac{\log\log N_E}{\log N_E}).\end{aligned}$$

Modified Explicit Formula:

$$\begin{aligned}\sum_{\gamma_E^{(j)}} G\Big(\frac{\log X}{2\pi}\gamma_E^{(j)}\Big) &= \frac{\log N_E}{\log X}\widehat{G}(0) + G(0) \\ &- 2\sum_p \frac{\log p}{\log X}\frac{1}{p}\widehat{G}(\frac{\log p}{\log X})a_E(p) \\ &- 2\sum_p \frac{\log p}{\log X}\frac{1}{p^2}\widehat{G}(\frac{2\log p}{\log X})a_E^2(p) \\ &+ O(\frac{\log\log X}{\log X}).\end{aligned}$$

Previous Results

Iwaniec-Luo-Sarnak: Among other examples they prove the following. Let $H_k^\star(N)$ be the space of holomorphic even weight k cuspidal newforms of square-free level N; let $H_k^\pm(N)$ be the subspaces of forms with even (odd) functional equation. For test functions f supported in $(-2,2)$, as $N \to \infty$, the densities of $H_k^\star(N)$, $H_k^+(N)$ and $H_k^-(N)$ agree with O, $SO(\text{even})$ and $SO(\text{odd})$. Explicitly,

$$\lim_{N\to\infty} \frac{1}{|H_k^\star(N)|} \sum_{h \in H_k^\star(N)} D_{1,h}(f) = \int f(x) W_{1,O}(x)dx, \qquad (0.3)$$

and similarly for the other two families.

Main Tools:

(1) Petersson Formula: Let $S_k(N)$ be the space of weight k cusp forms of level N. Let $B_k(N)$ be a basis of forms f with normalized coefficients $\psi_f(n)$. Then for $m, n \geq 1$:

$$\sum_{f \in B_k(N)} \overline{\psi}_f(m)\psi_f(n) \;=\; \delta(m,n) + (2\pi i)^k \sum_{c \equiv 0(N)} \frac{S(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)$$

where $\delta(m,n) = 1$ for $m = n$ and 0 otherwise, J_{k-1} is a Bessel function, and $S(m,n;c)$ is the Kloosterman sum.

(2) Conductors: The analytic conductors of the forms are all equal: $c_h = k^2 N$. $|H_k^\star(N)| \approx 2|H_k^\pm(N)| \approx \frac{k-1}{24}\varphi(N)$, $\varphi(N)$ is Euler's phi-function.

Note: they extend results to $KN \to \infty$, $k \leq K$.

Previous Results (cont)

Rubinstein: $\chi_d(n) = (\frac{d}{p})$. Let D be the set of d with χ_d primitive, $D(X) = \{d \in D : \frac{X}{2} \leq |D| < X\}$. Consider family of twists $L(s, \chi_d)$ of the zeta-function. Let $L = \frac{\log X}{2\pi}$. For test functions f with support $\sum_i \sigma_i < 1$

$$\frac{1}{|D(X)|} \sum_{d \in D(X)} \sum_{\substack{j_1, \dots, j_n \\ distinct}} \prod_i f_i(L\gamma_d^{(j_i)}) \ \rightarrow \ \int_{\mathbf{R}^n} f(x) W_{n,USp}(x) dx.$$

Main Tools:

(1) Orthogonality of characters

(2) Jutila:

$$\sum_{m \leq X^\alpha} \Big| \sum_{d \in D(X)} \chi_d(m) \Big|^2 \ \ll_\epsilon \ X^{1+\alpha} \log^A X. \tag{0.4}$$

Note that the conductors are all approximately the same, and in the Modified Explicit Formula can replace $\log d$ with $\log X$ with cost $O(\frac{1}{\log X})$.

Comments on Previous Results

The proofs have two similarities: there is some nice averaging formula to control the arithmetic quantities, and the conductors are manageable (the log of the conductors is approximately constant).

In Iwaniec-Luo-Sarnak, the conductors were the same for different members of the family.

In Rubinstein, $\log d = \log X + O(1)$.

For families of elliptic curves, let $\Delta(t)$ be the discriminant. Then the conductor $C(t)$ is

$$C(t) \ = \ \prod_{p|\Delta(t)} p^{f_p(t)}$$

For $p > 3$, if the curve is minimal for p then $f_p(t) = 0$ if $p \nmid \Delta(t)$, 1 if $p|\Delta(t)$ and $p \nmid c_4(t)$, and 2 if $p|\Delta(t)$ and $p|c_4(t)$. If $p > 3$ and $p^{12} \nmid \Delta(t)$, then the equation is minimal at p.

Note two t that are close could yield $\Delta(t)$'s with wildly differing factorization, hence the conductors can fluctuate greatly.

1-Level Expansion

$$\begin{aligned}
D_{1,\mathcal{F}}(f) &= \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\sum_j f\Big(\frac{\log N_E}{2\pi}\gamma_E^{(j)}\Big)\\
&= \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\widehat{f}(0)+f_i(0)\\
&\quad - \frac{2}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\sum_p \frac{\log p}{\log N_E}\frac{1}{p}\widehat{f}\Big(\frac{\log p}{\log N_E}\Big)a_E(p)\\
&\quad - \frac{2}{|\mathcal{F}|}\sum_p \frac{\log p}{\log N_E}\frac{1}{p^2}\widehat{f}\Big(2\frac{\log p}{\log N_E}\Big)a_E^2(p)\\
&\quad + O\Big(\frac{\log\log N_E}{\log N_E}\Big)
\end{aligned}$$

Want to move $\frac{1}{|\mathcal{F}|}\Sigma_{E\in\mathcal{F}}$ past the prime sum and the conductors to hit the $a_E^r(p)$ terms.

If instead we scaled by the average log-conductor, we would have $\frac{1}{|\mathcal{F}|}\Sigma_{E\in\mathcal{F}}\, a_E^r(p)$, $r=1,2$. This is just $\frac{|\mathcal{F}|}{p}$ complete sums of $a_E^r(p)$ mod p, which we can often evaluate.

Leads us to study

$$A_{r,\mathcal{F}}(p) \;=\; \sum_{t(p)} a_t^r(p).$$

2-Level Expansion

Need to evaluate terms like

$$\frac{1}{|\mathcal{F}|} \sum_{E \in \mathcal{F}} \prod_{i=1}^{2} \frac{1}{p_i^{r_i}} g_i\Big(\frac{\log p_i}{\log N_E}\Big) a_E^{r_i}(p_i).$$

Goal: want to pass $\frac{1}{|\mathcal{F}|} \sum_{E \in \mathcal{F}}$ past the test function to the $a_E^{r_i}(p_i)$ terms and exploit cancellation of the $a_E(p)$'s.

The following is our best analogue to the Petersson formula. For a one-parameter family (the family of all elliptic curves is handled similarly) define

$$A_{r,\mathcal{F}}(p) \ = \ \sum_{t(p)} a_t^r(p).$$

If $p_1, \dots, p_n$ are distinct primes,

$$\sum_{t(p_1 \cdots p_n)} a_{t_1}^{r_1}(p_1) \cdots a_{t_n}^{r_n}(p_n) = A_{r_1,\mathcal{F}}(p_1) \cdots A_{r_n,\mathcal{F}}(p_n).$$

Thus, to handle the products, it is sufficient to understand the one-dimensional versions. If $y^2 = x^3 + a_2(t)x^2 + a_4(t)x + a_6(t) = f_t(x)$, $a_i(t)$ integer polynomials, then

$$a_t(p) = -\sum_{x(p)} \left(\frac{f_t(x)}{p}\right).$$

Needed Input

For many families

$$\begin{aligned}(1) : A_{1,\mathcal{F}}(p) &= -rp + O(1)\\ (2) : A_{2,\mathcal{F}}(p) &= p^2 - m_{\mathcal{F}}p + O(1),\ m_{\mathcal{F}} > 0.\end{aligned}$$

Consider a one-parameter family $\mathcal{E}$ (of rank r over $\mathbf{Q}(t)$) of elliptic curves E_t over $\mathbf{Q}$. $\mathcal{E}$ is a rational elliptic surface iff $\mathcal{E}$ is birational to $\mathbf{P}^2$.

Silverman and Rosen prove

$$\lim_{X\to\infty} \frac{1}{X} \sum_{p \le X} -A_{\mathcal{E}}(p) \log p = r$$

By partial summation, this is as good as (1).

For surfaces with $j(t)$ non-constant, Michel proves

$$A_{2,\mathcal{F}}(p) = p^2 + O(p^{3/2}).$$

New Results

Rational Surfaces Density Theorem: *Consider a one-parameter family of elliptic curves of rank r over $\boldsymbol{Q}(t)$ that is a rational surface. Assume GRH, $j(t)$ non-constant, and the ABC (or Square-Free Sieve) conjecture if $\Delta(t)$ has an irreducible polynomial factor of degree at least 4. Let $m = \deg C(t)$ and f_i be an even Schwartz function of small but non-zero support σ_i ($\sigma_1 < \min(\frac{1}{2}, \frac{2}{3m})$ for the 1-level density, $\sigma_1 + \sigma_2 < \frac{1}{3m}$ for the 2-level density). Assume the Birch and Swinnerton-Dyer conjecture for interpretation purposes. Possibly after passing to a subsequence,*

$$\begin{aligned}
D_{1,\mathcal{F}}^{(r)}(f_1) &= \widehat{f_1}(0) + \frac{1}{2}f_1(0)\\
D_{2,\mathcal{F}}^{(r)}(f) &= \prod_{i=1}^{2}\left[\widehat{f_i}(0) + \frac{1}{2}f_i(0)\right] + 2\int_{-\infty}^{\infty}|u|\widehat{f_1}(u)\widehat{f_2}(u)du\\
&\quad -2\widehat{f_1 f_2}(0) - f_1(0)f_2(0) + (f_1 f_2)(0)N(\mathcal{F},-1)
\end{aligned}$$

where $N(\mathcal{F},-1)$ is the percent of curves with odd sign.

For small support, the 1-level density of the non-family zeros agrees with SO(even), O, and SO(odd); the 2-level density of the non-family zeros agrees with SO(even), O, and SO(odd) depending on whether the signs are all even, equidistributed in the limit, or all odd.

The densities of the non-family zeros agree with Katz and Sarnak's predictions, at least for small support.

Examples

Constant-Sign Families:

1. $y^2 = x^3 + 2^4(-3)^3(9t+1)^2$, $9t+1$ Square-Free: all even.
2. $y^2 = x^3 \pm 4(4t+2)x$, $4t+2$ Square-Free: $+$ yields all odd, $-$ yields all even.
3. $y^2 = x^3 + tx^2 - (t+3)x + 1$, t^2+3t+9 Square-Free: all odd.

First two rank 0 over $\mathbf{Q}(t)$; third is rank 1. Only assume GRH for first two; add B-SD to interpret third.

Harald Helfgott has shown, if $j(t)$ and $M(t)$ non-constant, that ABC (or Square-Free Sieve) conjecture and Polynomial Moebius imply equidistribution of sign.

Family of Rank 6 over $\mathbf{Q}(t)$:

$$\begin{aligned} y^2 &= x^3 + (2at - B)x^2 + (2bt - C)(t^2 + 2t - A + 1)x \\ &\qquad +(2ct - D)(t^2 + 2t - A + 1)^2 \end{aligned}$$

$$\begin{aligned} A &= 8916100448256000000 \\ B &= -811365140824616222208 \\ C &= 26497490347321493520384 \\ D &= -343107594345448813363200 \\ a &= 16660111104 \\ b &= -1603174809600 \\ c &= 2149908480000 \end{aligned} \tag{0.5}$$

Need GRH. Also need ABC or Square-Free Sieve conjecture to handle sieving, B-SD to interpret result.

Sieving

If conductors were constant and summed over $t \in [N, 2N]$, would have $\frac{N}{p}$ complete sums, each giving $A_{r,\mathcal{F}}(p)$.

Let $D(t)$ be the product of the irreducible factors of $\Delta(t)$. Can often show $C(t)$ is a monotone polynomial of t when $D(t)$ is square-free, and there are $c_{\mathcal{F}}N + o(N)$ such t. (Unconditional if all factors of $D(t)$ of degree ≤ 3; else need ABC or Square-Free Sieve conjecture).

$$\begin{aligned}\sum_{\substack{t=N \\ D(t) \\ sqfree}}^{2N} S(t) &= \sum_{d=1}^{N^{k/2}} \mu(d) \sum_{\substack{D(t)\equiv 0(d^2) \\ t\in[N,2N]}} S(t) \\ &= \sum_{d=1}^{\log^l N} \mu(d) \sum_{\substack{D(t)\equiv 0(d^2) \\ t\in[N,2N]}} S(t) + \sum_{d\geq \log^l N}^{N^{k/2}} \mu(d) \sum_{\substack{D(t)\equiv 0(d^2) \\ t\in[N,2N]}} S(t).\end{aligned}$$

Handle first piece by progressions, handle second piece by Cauchy-Schwartz.

Sieving (cont)

The number of t in the second sum is $o(N)$ (unconditionally if all of $D(t)$'s factors are deg ≤ 3). Denote these t by $\mathcal{T}$. By Cauchy-Schwartz:

$$\begin{aligned}\sum_{t\in\mathcal{T}} S(t) &\ll \Big(\sum_{t\in\mathcal{T}} S^2(t)\Big)^{\frac{1}{2}} \cdot \Big(\sum_{t\in\mathcal{T}} 1\Big)^{\frac{1}{2}} \\ &\ll \Big(\sum_{t\in[N,2N]} S^2(t)\Big)^{\frac{1}{2}} \cdot o\Big(\sqrt{N}\Big).\end{aligned}$$

If we can show $\Sigma_{t=N}^{2N} S^2(t) = O(N)$, then the error term is negligible as $N \to \infty$. Often isn't too bad, as just need order of magnitude, and not exact value.

First piece is handled by progressions: let $\nu(d)$ be the number of incongruent roots of $D(t) \equiv 0 \bmod d^2$; $\nu(d) \ll d^\epsilon$. Let $t_i(d)$ be one of the $\nu(d)$ roots. This gives a sequence of t: $t_i(d), t_i(d)+d^2, \dots, t_i(d)+[\frac{N}{d^2}]d^2$.

If $(d,p)=1$, then (mod p), go through the complete set of residue classes $\frac{N/d^2}{p}$ times. As $d < \log^l N$, $l < 2$, can take all $p > \log^l N$ in the Explicit Formula, incorporating lower p's into the error terms.

Partial Summation

$$\sum_{M}^{N} a_n b_n \;=\; A(N)b_N + \sum_{M}^{N-1} A(u)(b_u - b_{u+1}),\; A(u) = \sum_{n=M}^{u} a_n.$$

Notation: $\tilde{a}_{d,i,p}(t') = a_{t(d,i,t')}(p)$, $G_{d,i,P}(u)$ is related to the test functions.

Applying Partial Summation

$$\begin{aligned} S(d,i,r,p) &= \sum_{t'=0}^{[N/d^2]} \tilde{a}^{r}_{d,i,p}(t')G_{d,i,p}(t') \\ &= \Big(\frac{[N/d^2]}{p} A_{r,\mathcal{F}}(p) + O(p^R)\Big) G_{d,i,p}([N/d^2]) \\ &\quad - \sum_{u=0}^{[N/d^2]-1} \Big(\frac{u}{p} A_{r,\mathcal{F}}(p) + O(p^R)\Big) \\ &\qquad \cdot\Big(G_{d,i,p}(u) - G_{d,i,p}(u+1)\Big) \\ S(r,p) &= \sum_{d=1}^{\log^l N} \mu(d) \sum_{i=1}^{\nu(d)} S(d,i,r,p) \\ &= \sum_{w=1}^{4} \sum_{d=1}^{\log^l N} \mu(d) \sum_{i=1}^{\nu(d)} S_w(d,i,r,p). \end{aligned}$$

$O(p^R)$ is the error from using Hasse to bound the partial sums: $p^R = p^{1+\frac{r}{2}}$.

First, Second and Third Sums

First Sum: Use Taylor Expansion. Gives the main term:

$$\frac{A_{r,\mathcal{F}}(p)G_p(N)}{p}.$$

Second Sum: Summing over primes won't contribute for small support. $G_{d,i,p}$ term is $O(1)$, left with

$$\frac{1}{N}\sum_{p=\log^l N}^{N^\alpha}\frac{1}{p}p^{1+\frac{r}{2}}.$$

Third Sum: Apply Partial Summation again. Taylor Expansion gains a $O(\frac{1}{\log N})$, which is sufficient.

$$\begin{aligned} S_3(d,i,r,p) \ &= \ \Big(G_{d,i,p}(0) - G_{d,i,p}([N/d^2])\Big)\frac{[N/d^2]-1}{p}A_{r,\mathcal{F}}(p) \\ &\quad - \sum_{u=0}^{[N/d^2]-2}\left(G_{d,i,p}(0) - G_{d,i,p}(u+1)\right)\frac{1}{p}A_{r,\mathcal{F}}(p). \end{aligned}$$

Using the Taylor Expansion, we gain a $\frac{1}{\log N}$ in the first term, making it of size $\frac{A_{r,\mathcal{F}}(p)}{p}\frac{[N/d^2]}{\log N} \ll \frac{A_{r,\mathcal{F}}(p)}{p}\frac{|\mathcal{F}|}{d^2\log N}$.

For the second term, we have $< [N/d^2]$ summands, each $\ll \frac{1}{\log N}\frac{S_c(r,p)}{p}$. We again obtain a term of size $\frac{S_c(r,p)}{p}\frac{|\mathcal{F}|}{d^2\log N}$.

Third Sum is a lower order contribution.

Difficult Piece: Fourth Sum I

$$\sum_{u=0}^{[N/d^2]-1} O(p^R)(G_{d,i,p}(u) - G_{d,i,p}(u+1))$$

Using the Taylor Expansion for $G_{d,i,p}(u) - G_{d,i,p}(u+1)$ is not sufficient. This would give $\frac{Np^R}{d^2 \log N}$. Summing over i and d is manageable, and would give us $O(p^R \frac{|\mathcal{F}|}{\log N})$. Dividing by the cardinality of the family gives $O(\frac{p^R}{\log N})$.

The problem is in summing over the primes, as we no longer have $\frac{1}{|\mathcal{F}|}$. We multiply by $\frac{1}{p^r}$.

Consider the case $r = 1$. Then $R = 1 + \frac{r_1}{2} = \frac{3}{2}$, and $\frac{1}{p^r} = \frac{1}{p}$. We have

$$\sum_{p=\log^l N}^{N^{m\sigma}} \frac{1}{p} \frac{p^{\frac{3}{2}}}{\log N} \gg N^{m\sigma}.$$

As $N \to \infty$, this term blows up. We need much better cancellation. Note, by Hasse, $O(p^R) \leq 2^R p^R$.

Fourth Sum: II

$$\begin{aligned} & \sum_{u=0}^{[N/d^2]-1} \Big| G_{d,i,p}(u) - G_{d,i,p}(u+1) \Big| \\ = & \sum_{u=0}^{[N/d^2]-1} \Big| g\Big(\frac{\log p}{\log C(t_i(d)+ud^2)}\Big) - g\Big(\frac{\log p}{\log C(t_i(d)+(u+1)d^2)}\Big)\Big| \end{aligned}$$

If the conductors are monotone, for fixed i, d and p, by the Mean Value Theorem the above is small. It is essential that we look at the above as a problem in bounded variation of g and not $g_{d,i,p}$. Only need g has bounded derivative to bound the u-sum by the support of g (can add boundary terms).

For the 2-Level Density, we would use:

$$\begin{aligned} |a_1a_2 - b_1b_2| &= |a_1a_2 - b_1a_2 + b_1a_2 - b_1b_2| \\ &\leq |a_1a_2 - b_1a_2| + |b_1a_2 - b_1b_2| \\ &= |a_2| \cdot |a_1 - b_1| + |b_1| \cdot |a_2 - b_2| \end{aligned}$$

Note: *if our conductors are not monotone, we cannot apply theorems on bounded variation. The problem is we could transverse* $[0, 1000\sigma]$ *(or a large subset of it) many times.*

Handling the Conductors I

$$C(t) = \prod_{p|\Delta(t)} p^{f_p(t)}, \tag{0.6}$$

where for $p > 3$, if the curve is minimal for p then $f_p(t) = 0$ if $p \nmid \Delta(t)$, 1 if $p|\Delta(t)$ and $p \nmid c_4(t)$, and 2 if $p|\Delta(t)$ and $p|c_4(t)$. If $p > 3$ and $p^{12} \nmid \Delta(t)$, then the equation is minimal at p.

Let $D_1(t)$ be the product of primitive irreducible polynomial factors that $\Delta(t)$ and $c_4(t)$ share. Let $D_2(t)$ be the remaining primitive irreducible polynomial factors of $\Delta(t)$. There are only finitely many primes that can divide $D_1(t)$ and $D_2(t)$ for any t.

We consider only t with $D_1(t)D_2(t)$ square-free; if not possible, square-free save for a fixed number of primes. By passing to a subsequence, can assure always the same power.

Expect the conductors to be like $D_1^2(t)D_2(t)$ except for a finite set of bad primes. This set contains the primes that can divide both, primes where not square-free, primes that divide $\Delta(t)$ for all t, and $p = 2, 3$.

Let P be the product of the bad primes.

Handling the Conductors II

By Tate's Algorithm, can determine $f_p(t)$, which depends on the coefficients $a_i(t)$ mod powers of p. $\Delta(t)$ is not identically zero. Thus, $\exists t_1 > 0$ such that $\forall t \geq t_1$, $\Delta(t) \neq 0$.

Apply Tate's Algorithm to E_{t_1} to determine $f_p(t_1)$ for the bad primes. By choosing m sufficiently large, we can show $f_p(\tau) = f_p(P^m t + t_1) = f_p(t_1)$ for $p|P$, ie, for the bad primes. The number m depends on the curve E_{t_1}.

This is because in Tate's Algorithm, we only need the values modulo a power of p. We have

$$a_i(\tau) \ = \ a_i(P_0^m t + t_1) \ = \ P_0^m t \hat{a}_i(P_0^m t) + a_i(t_1).$$

In applying Tate's Algorithm, we need $a_i(\tau)$ modulo powers of p. If m is sufficiently large, we can ignore $\tilde{a}_i(t)$ in all equivalence checks, as for the powers of p we investigate, $\tilde{a}_i(t) \equiv 0$.

Note, for m enormous, for bad primes, the order of p dividing $D(P^m t + t_1)$ is independent of t. So can find integers st $C(\tau) = c_{bad} \frac{D_1^2(\tau)}{c_1} \frac{D_2(\tau)}{c_2}$.

Application: Bounding Excess Rank

$$D_{1,\mathcal{F}}(f_1) = \widehat{f_1}(0) + \frac{1}{2}f_1(0) + rf_1(0).$$

To estimate the percent with rank at least $r+R$, P_R, we get

$$Rf_1(0)P_R \leq \widehat{f_1}(0) + \frac{1}{2}f_1(0),\ \ R > 1.$$

Note the family rank r has been cancelled from both sides.

By using the 2-level density, however, we get *squares* of the rank on the left hand side. The advantage is we get a cross term rR. The disadvantage is our support is smaller. Once R is large, the 2-level density yields better results.

Potential Lower Order Density Terms

Can often show

$$A_{2,\mathcal{F}}(p) = p^2 - m_{\mathcal{F}} \cdot p + O(1),\ m_{\mathcal{F}} > 0.$$

The p^2 term contributes $-\frac{1}{2}f(0)$; the $m_{\mathcal{F}}$ term contributes something of size $\frac{1}{\log N}$.

The potential lower order density terms, arising from lower order terms in the sums of the second moments of $a_E^2(p)$, could be masked by the errors propagating through our derivations. We have errors of the size $\frac{\log\log N}{\log N}$ arising from the Explicit Formula and the contributions from $a_E^m(p)$, $m \geq 3$.

To truly observe lower order corrections to the densities, a significantly more delicate analysis of these discarded terms are needed. The conductor dependence in the Gamma factors of the Explicit Formula are easily managed. The real difficulty is handling the primes which divide the discriminant and the $m \geq 3$ terms.

We save this for a future project, and content ourselves with observing a potential lower order density term.

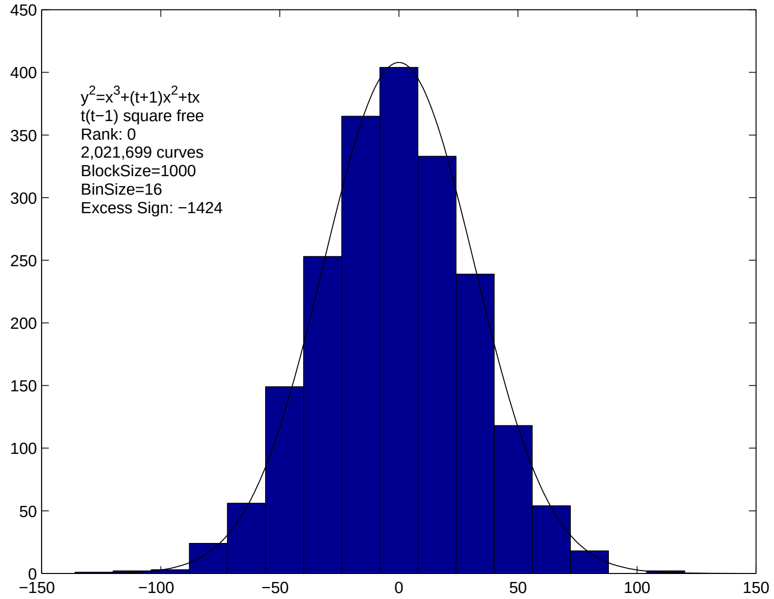
Variation of Sign in One-Parameter Families

For his junior thesis, Atul Pokharel investigated the Restricted Sign Conjecture by studying the distribution of signs in one-parameter families. A representative family is included below. For N curves, the excess of positive to negative signs in intervals of 1000 was computed, for a total of $\frac{N}{1000}$ blocks. If the signs are randomly distributed, one would expect a histogram bin plot to reveal a Gaussian structure, with mean 0 and standard deviation $\sqrt{1000}$. Note this is a far stronger assumption than equidistribution of sign.

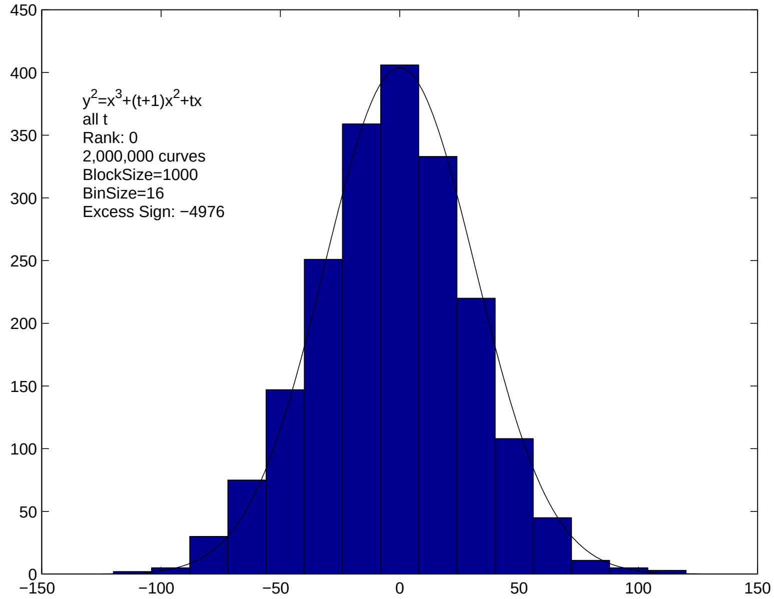
This was tested for many families; further, as in this thesis we sieve to $D(t)$ square-free, the same calculation was performed for the sub-families with $D(t)$ square-free. Restricting t did not seem to change the shape of the observed data. A more sensitive analysis is currently underway.

For the family $y^2 = x^3 + (t+1)x^2 + tx$, we sieve to $D(t) = t(t-1)$ square-free. Approximately 32% of t give $D(t)$ square-free. First, the sign of E_t was computed for $t \in [2, 5 \cdot 10^7]$. Histogram plots were prepared without restricting to $D(t)$. Restricting to $D(t)$ square-free requires evaluating $\mu(t)\mu(t-1)$; as this was lengthy (and this is a preliminary verification), we contented ourselves with checking the first $2 \cdot 10^6$ square-free $D(t)$.

Distribution of signs: $y^2 = x^3 + (t+1)x^2 + tx$

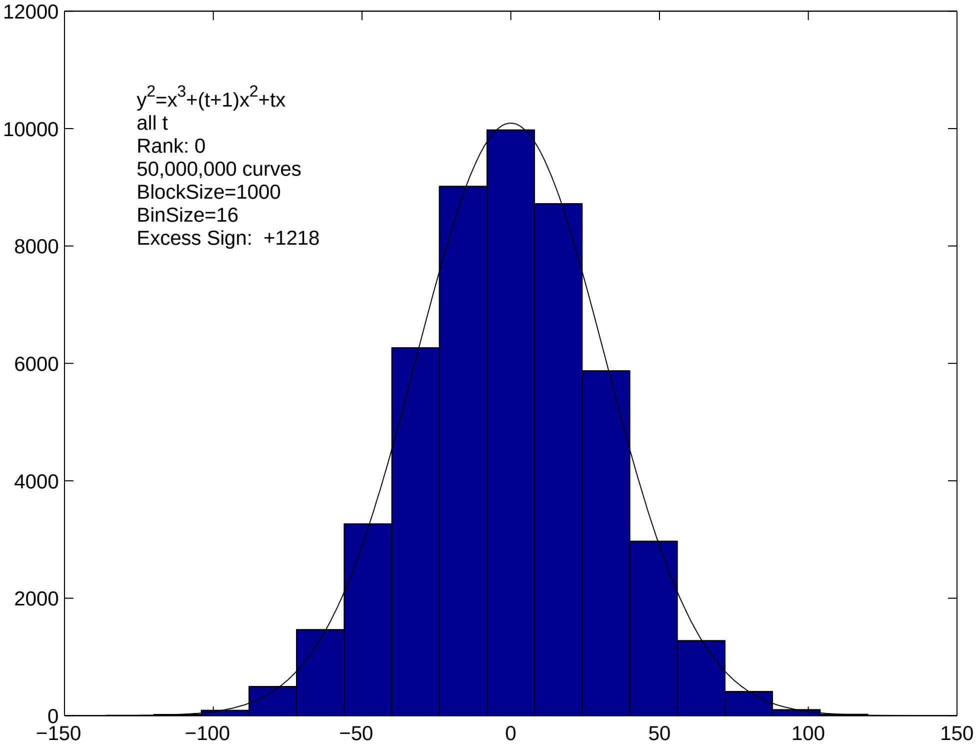


Histogram plot: $D(t)$ square-free, first $2 \cdot 10^6$ such t.



Histogram plot: All $t \in [2, 2 \cdot 10^6]$.

Distribution of signs: $y^2 = x^3 + (t+1)x^2 + tx$



Histogram plot: All $t \in [2, 5 \cdot 10^7]$

The observed behaviour agrees with the predicted behaviour. Note as the number of curves increase (comparing the plot of $5 \cdot 10^7$ points to $2 \cdot 10^6$ points), the fit to the Gaussian improves.