Compositio Math. 104 (2004) 952–992
DOI: 10.1112/S0010437X04000582

One- and two-level densities for rational families of elliptic curves: evidence for the underlying group symmetries

Steven J. Miller

Abstract

Following Katz–Sarnak, Iwaniec–Luo–Sarnak and Rubinstein, we use the one- and two-level densities to study the distribution of low-lying zeros for one-parameter rational families of elliptic curves over $\mathbb{Q}(t)$. Modulo standard conjectures, for small support the densities agree with Katz and Sarnak's predictions. Further, the densities confirm that the curves' L-functions behave in a manner consistent with having r zeros at the critical point, as predicted by the Birch and Swinnerton-Dyer conjecture. By studying the two-level densities of some constant sign families, we find the first examples of families of elliptic curves where we can distinguish $SO(\text{even})$ from $SO(\text{odd})$ symmetry.

1. Introduction

1.1 n-level correlations and densities

Assuming the Generalized Riemann Hypothesis (GRH), the zeros of any L-function lie on the critical line, and therefore it is possible to investigate statistics of the normalized zeros. The general philosophy, born out in many examples (see [CFKRS02]), is that the behavior of random matrices/ensembles of random matrices behaves in a similar manner to that of L-functions/families of L-functions. By a family $\mathcal{F}$ we mean a collection of geometric objects and their associated L-functions, where the geometric objects have similar properties.

We expect that there is a symmetry group $\mathcal{G}(\mathcal{F})$ (one of the classical compact groups $U(N)$, $SU(N)$, $USp(2N)$, $SO(\text{even})$ and $SO(\text{odd})$) which can be associated to a family of L-functions, and that the behavior of the eigenvalues of matrices in $\mathcal{G}(\mathcal{F})$ should (after appropriate normalizations) equal the behavior of the zeros of L-functions.

Iwaniec *et al.* [ILS00] considered (among other examples) all cuspidal newforms of a given level and weight. Rubinstein [Rub98] considered twists by the fundamental discriminants D of a fixed modular form.

We study the family of all elliptic curves and various one-parameter families of elliptic curves. Thus, in our case the notion of a family is the standard notion from geometry: we have a collection of curves over a base and the geometry is much clearer in our examples than in [ILS00] and [Rub98].

Let $\{\alpha_j\}$ be an increasing sequence of numbers tending to infinity, such as eigenvalues or zeros normalized to have mean spacing 1. For a compact box $B \subset \mathbb{R}^{n-1}$, define the *$n$-level correlation* by

$$\lim_{N\to\infty} \frac{\#\{(\alpha_{j_1} - \alpha_{j_2}, \ldots, \alpha_{j_{n-1}} - \alpha_{j_n}) \in B, j_i \in \{1, \ldots, N\}, j_i \neq j_k\}}{N}. \tag{1.1}$$

Received 14 October 2002, accepted in final form 2 June 2003.
2000 Mathematics Subject Classification 11M26 (primary), 11G05, 11G40, 11M26 (secondary).
Keywords: n-level density, low lying zeros, elliptic curve L-functions, Birch and Swinnerton-Dyer conjecture.

Note that the n-level correlations are unaffected by removing finitely many zeros. Instead of using a box, one can study a smoothed version with a test function on $\mathbb{R}^n$ (see [RS96]).

For test functions whose Fourier transform has small support, Montgomery [Mon73] proved that the two-level (and Hejhal [Hej94] proved that the three-level) correlation for the zeros of $\zeta(s)$ is the same as that of the eigenvalues of complex Hermitian matrices with entries independently chosen from Gaussian distributions (the GUE ensemble of matrices). Rudnick and Sarnak [RS96] proved that the n-level correlations for the zeros of any automorphic cuspidal L-function are the same as that of the zeros of the GUE. The universality is due to the fact that the correlations are controlled by the second moment of a_p, and while there are many possible limiting distributions, all have the same second moment.

Katz and Sarnak [KS99a] proved that the classical compact groups have the same n-level correlations. In particular, we cannot use the n-level correlations to distinguish GUE behavior, $U(N)$, from the other classical compact groups. We are led to investigate another statistic which will depend on the underlying group.

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 (Birch and Swinnerton-Dyer conjecture). If we force the Mordell–Weil group to be large, we expect many zeros exactly at $s = \frac{1}{2}$, and this might influence the behavior of the neighboring zeros. Hence, we are led to study the distribution of the first few, or low-lying, zeros, and the fascinating possibility that there could be a difference in statistics for zeros near $\frac{1}{2}$ and those zeros higher up.

Let $f(x)$ be an even Schwartz function whose Fourier transform is supported in a neighborhood of the origin. We assume that f is of the form $\prod_{i=1}^{n} f_i(x_i)$. The *n-level density* for the family $\mathcal{F}$ with test function f is

$$D_{n,\mathcal{F}}(f) = \frac{1}{|\mathcal{F}|} \sum_{E\in\mathcal{F}} \sum_{\substack{j_1,\dots,j_n \\ j_i\neq\pm j_k}} f_1\left(\frac{\log N_E}{2\pi}\gamma_E^{(j_1)}\right)\cdots f_n\left(\frac{\log N_E}{2\pi}\gamma_E^{(j_n)}\right), \tag{1.2}$$

where $\gamma_E^{(j_i)}$ runs through the non-trivial zeros of the curve E, and N_E is its conductor. We rescale the zeros by $\log N_E$ as this is the order of the number of zeros with imaginary part less than a large absolute constant (see [ILS00]). As f_i is Schwartz, most of the contribution is due to the zeros near the critical point. We use the explicit formula (Equation (2.3)) to relate sums of test functions over zeros to sums over primes of $a_E(p)$ and $a_E^2(p)$.

Katz and Sarnak [KS99a] determined the $N \to \infty$ limits for the n-level densities of eigenvalues near 1 for the classical compact groups (see § 3); their calculations can be modified to determine the densities of classical compact groups with a forced number of eigenvalues at 1. Forcing eigenvalues at 1 corresponds to L-functions with zeros forced at the critical point.

1.2 Results

To any geometric family in the function field case, the results of Katz and Sarnak [KS99a, KS99b] state that the n-level density of zeros near $\frac{1}{2}$ depends only on a symmetry group attached to the family. In particular, for generic families of elliptic curves the relevant symmetry is orthogonal. One can further analyze the distributions depending on the signs of the functional equations. As the families of elliptic curves 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).

For an elliptic curve E_t, let $D(t)$ be the product of the irreducible polynomial factors of the discriminant $\Delta(t)$, and let $C(t)$ be the conductor. Let B be the largest square dividing $D(t)$ for all t. Pass to a subsequence $ct + t_0$, and call $t \in [N, 2N]$ *good* if $D(ct + t_0)$ is square-free, except for primes $p|B$ where the power of such $p|D(t)$ is independent of t.

The main result is Theorem 5.8, which can be summarized as follows.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Rational Surfaces Density Theorem. *Consider a one-parameter family of elliptic curves of rank r over $\mathbb{Q}(t)$ which constitutes a rational surface. Assume the GRH, $j(E_t)$ non-constant, and if $\Delta(t)$ has an irreducible polynomial factor of degree at least 4, assume the ABC conjecture.*

After passing to a subsequence, for t good, $C(t)$ is a polynomial. Let f_i be an even Schwartz function of small but non-zero support σ_i ($\sigma_1 < \min(\frac{1}{2}, 2/3m)$ for the one-level density, $\sigma_1 + \sigma_2 < 1/3m$ for the two-level density).

The one-level density agrees with the orthogonal densities plus a term which equals the contributions from r zeros at the critical point. The two-level density agrees with SO(even), O and SO(odd), depending on whether the signs are all even, equidistributed in the limit or all odd, plus a term which equals the contribution from r zeros at the critical point. Thus, for small support, the densities of the zeros agree with Katz and Sarnak's predictions. Further, the densities confirm that the curves' L-functions behave in a manner consistent with having r zeros at the critical point, as predicted by the Birch and Swinnerton-Dyer conjecture.

The ABC conjecture is used to handle large prime divisors of polynomials of degree 4 or more (see [Gra98]). In place of ABC, one could assume the square-free sieve conjecture.

For the one-level densities, the three orthogonal densities agree for test functions with support less than 1, split (i.e. are distinguishable) for support greater than 1, but are all distinguishable from U and Sp for any support. Hence, unlike the n-level correlations, the one-level density is already sufficient to observe non-GUE and non-symplectic behavior.

The polynomial growth of the conductor in families of elliptic curves makes it difficult to evaluate the sums over primes for test functions with moderate support. Converting to our language, for small support the one-level densities for many families have been shown to be equal to the Katz–Sarnak predictions: all elliptic curves (Brumer and Heath-Brown [Bru92, BH], support less than $\frac{2}{3}$); twists of a given curve (support less than 1); one-parameter families (Silverman [Sil98], small support).

None of these are sufficient to distinguish the three orthogonal candidates. Further, previous investigations have rescaled each curve's zeros by the average of the logarithms of the conductors. This greatly simplifies the calculations; however, the normalization is no longer natural for each curve, as each curve can sit in infinitely many families, each with a different average spacing. By using local normalizations for each curve's zeros, the n-level density for a family becomes the average of the n-level densities for each curve.

The utility of the two-level density is that, even for test functions with arbitrarily small support, the three candidate orthogonal symmetries *are* distinguishable, and in a very satisfying way. The three candidates differ by a factor which encodes the distribution of sign in the family, and all differ from the GUE's two-level density.

We study several families of constant sign, and we will see that the densities are as expected. Thus, for these constant sign families, the two-level density reflects the predicted symmetry, which is invisible through the one-level density because of support considerations.

Similar to the universality Rudnick and Sarnak [RS96] found in studying n-level correlations, our universality follows from the sums of $a_t^2(p)$ in our families (the second moments). For non-constant $j(E_t)$, this follows from a Sato–Tate law proved by Michel [Mic95, Theorem 2.3].

1.3 Structure of the paper

We first calculate sums of the Fourier coefficients of elliptic curves. We quote the predicted densities, and then calculate useful expansions for the one- and two-level densities for families of elliptic curves over $\mathbb{Q}(t)$. We derive the density results, conditional on the evaluation of many elliptic curve sums. We calculate these sums for one-parameter rational families of elliptic curves. We conclude with several examples (four constant sign families, a rank 1 and a rank 6 rational family).

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

We need excellent control over the conductors to evaluate the above sums; the estimation is so delicate that if the log of conductors are of size $m \log N$, fluctuations of size $O(1)$ yield error terms greater than the expected main terms.

The key observation is that the error terms can be controlled if the conductors are monotone. By straightforward sieving and applications of Tate's algorithm (to calculate the conductors), given a one-parameter rational family of elliptic curves, we may pass to a positive percentage sub-family where the conductors are monotone. The proofs of these results are given in the appendices.

In this paper, we concentrate on rational elliptic surfaces, because here Tate's conjecture is known. Rosen and Silverman [RS98] show that Tate's conjecture implies that certain sums over primes are related to the rank of the family over $\mathbb{Q}(t)$. This allows us to interpret some of our density terms as the contributions from r critical point zeros.

The modifications needed to handle the family of all elliptic curves, parametrized by

$$y^2 = x^3 + ax + b, \quad a \in [-N^2, N^2],\ b \in [-N^3, N^3], \tag{1.3}$$

are straightforward and can be found in [Mil02].

Finally, if instead we normalize by the average of the logarithms of the conductors, we obtain the same results but with significantly less work. This is done in [Mil02] for one-parameter families and the family of all elliptic curves.

2. Elliptic curve preliminaries

2.1 Definitions

Consider a one-parameter family $\mathcal{E}$ of elliptic curves E_t over $\mathbb{Q}(t)$:

$$\mathcal{E} : y^2 + a_1(t)xy + a_3(t)y = x^3 + a_2(t)x^2 + a_4(t)x + a_6(t), \quad a_i(t) \in \mathbb{Z}[t]. \tag{2.1}$$

For each curve E_t, let $\Delta(t)$ be its discriminant and $C(t)$ its conductor. Let $D(t)$ denote the product of the irreducible polynomial factors dividing $\Delta(t)$. We take $t \in [N, 2N]$ such that $D(t)$ is square-free.

Let $a_t(p) = a_{E_t}(p) = p+1-N_{t,p}$, where $N_{t,p}$ is the number of solutions of $E_t \bmod p$ (including ∞). If $y^2 = x^3 + A(t)x + B(t)$, then

$$a_t(p) = -\sum_{t(p)} \left(\frac{x^3 + A(t)x + B(t)}{p} \right). \tag{2.2}$$

2.2 Assumptions

We assume the following at various points.

Generalized Riemann Hypothesis (For elliptic curves). *Let $L(s, E)$ be the (normalized) L-function of an elliptic curve E. The non-trivial zeros ρ of $L(s, E)$ have $\mathrm{Re}(\rho) = \frac{1}{2}$.*

Occasionally we assume the Riemann hypothesis for the Riemann zeta-function and Dirichlet L-functions.

Birch and Swinnerton-Dyer Conjecture [BS63, BS65]. *Let E be an elliptic curve of geometric rank r over $\mathbb{Q}$ (the Mordell–Weil group is $\mathbb{Z}^r \oplus T$). Then the analytic rank (the order of vanishing of the L-function at the critical point) is also r.*

We only assume the above for interpretation purposes.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Tate's Conjecture for Elliptic Surfaces [Tat65]. *Let $\mathcal{E}/\mathbb{Q}$ be an elliptic surface and $L_2(\mathcal{E},s)$ be the L-series attached to $H^2_{\text{ét}}(\mathcal{E}/\overline{\mathbb{Q}},\mathbb{Q}_l)$. The series $L_2(\mathcal{E},s)$ has a meromorphic continuation to $\mathbb{C}$ and $-\text{ord}_{s=1}L_2(\mathcal{E},s) = \text{rank}\, NS(\mathcal{E}/\mathbb{Q})$, where $NS(\mathcal{E}/\mathbb{Q})$ is the $\mathbb{Q}$-rational part of the Néron–Severi group of $\mathcal{E}$. Further, $L_2(\mathcal{E},s)$ does not vanish on the line* $\text{Re}(s) = 1$.

Most of the one-parameter families that we investigate are rational surfaces, in which case Tate's conjecture is known (see [RS98]).

ABC Conjecture. *Fix $\epsilon > 0$. For co-prime positive integers a, b and c with $c = a + b$ and $N(a,b,c) = \prod_{p|abc} p$, $c \ll_\epsilon N(a,b,c)^{1+\epsilon}$.*

The full strength of ABC is never needed; rather, we need a consequence of ABC, the square-free sieve (see [Gra98]).

Square-Free Sieve Conjecture. *Fix an irreducible polynomial $f(t)$ of degree at least 4. As $N \to \infty$, the number of $t \in [N,2N]$ with $f(t)$ divisible by p^2 for some $p > \log N$ is $o(N)$.*

For irreducible polynomials of degree at most 3, the above is known, complete with a better error than $o(N)$ [Hoo76, ch. 4].

We use the square-free sieve to handle the variations in the conductors. If our evaluation of the log of the conductors is off by as little as a small constant, the prime sums become untractable. This is why many works normalize by the average log-conductor.

Restricted Sign Conjecture (For the family $\mathcal{F}$). *Consider a one-parameter family $\mathcal{F}$ of elliptic curves. As $N \to \infty$, the signs of the curves E_t are equidistributed for $t \in [N,2N]$.*

The restricted sign conjecture often fails. First, there are families with constant $j(E_t)$ where all curves have the same sign.

Helfgott [Hel] has recently related the restricted sign conjecture to the square-free sieve conjecture and standard conjectures on sums of Moebius.

Polynomial Moebius. *Let $f(t)$ be a non-constant polynomial such that no fixed square divides $f(t)$ for all t. Then $\sum_{t=N}^{2N} \mu(f(t)) = o(N)$.*

The polynomial Moebius conjecture is known for linear $f(t)$.

Helfgott showed that the square-free sieve and polynomial Moebius imply the restricted sign conjecture for many families. More precisely, let $M(t)$ be the product of the irreducible polynomials dividing $\Delta(t)$ and not $c_4(t)$.

Equidistribution of Sign in a Family Theorem [Hel]. *Let $\mathcal{F}$ be a one-parameter family with $a_i(t) \in \mathbb{Z}[t]$. If $j(E_t)$ and $M(t)$ are non-constant, then the signs of E_t, $t \in [N,2N]$, are equidistributed as $N \to \infty$. Further, if we restrict to good t, $t \in [N,2N]$ such that $D(t)$ is good (usually square-free), the signs are still equidistributed in the limit.*

The above is only used to calculate $N(\mathcal{F},-1)$, the percentage of odd curves. Without this, we can still calculate the one-level densities for small support, and all but one term in the two-level densities, $N(\mathcal{F},-1)f_1(0)f_2(0)$.

2.3 Explicit formula

The starting point for working with zeros of the L-functions of elliptic curves is the explicit formula (see [Mes86]), which relates sums over zeros to sums over primes.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

For an elliptic curve E with conductor N_E,

$$\sum_{\gamma_E^{(j)}} G\left(\gamma_E^{(j)}\frac{\log N_E}{2\pi}\right) = \widehat{G}(0) + G(0) - 2\sum_p \frac{\log p}{\log N_E}\frac{1}{p}\widehat{G}\left(\frac{\log p}{\log N_E}\right)a_E(p) - 2\sum_p \frac{\log p}{\log N_E}\frac{1}{p^2}\widehat{G}\left(\frac{2\log p}{\log N_E}\right)a_E^2(p) + O\left(\frac{\log\log N_E}{\log N_E}\right). \tag{2.3}$$

2.4 Sums of $a_t(p)$

Using the explicit formula, we find that we need to handle sums like

$$\sum_{t=N}^{2N} a_t^{r_1}(p_1)\cdots a_t^{r_n}(p_n). \tag{2.4}$$

We record these results for later use. Define

$$A_{r,\mathcal{F}}(p) = \sum_{t(p)} a_t^r(p). \tag{2.5}$$

Lemma 2.1. *Let $p_1,\dots,p_n$ be distinct primes and $r_i \geqslant 1$. Then*

$$\sum_{t(p_1\cdots p_n)}\prod_{i=1}^n a_t^{r_i}(p_i) = \prod_{i=1}^n A_{r_i,\mathcal{F}}(p_i). \tag{2.6}$$

The proof is a straightforward induction, using the fact that $a_{t+mp}(p) = a_t(p)$.

Lemma 2.1 is our best analogue to the Petersson formula, which is used in [ILS00] to obtain large support for the density functions.

The value $A_{1,\mathcal{F}}(p)/p$ is bounded independent of p [Del80]. Rosen and Silverman [RS98] proved the following conjecture of Nagao [Nag97].

Theorem 2.2 [RS98]. *For a one-parameter family $\mathcal{E}$ of elliptic curves over $\mathbb{Q}(t)$, if Tate's conjecture is true, then*

$$\lim_{X\to\infty}\frac{1}{X}\sum_{p\leqslant X} -\frac{A_{1,\mathcal{F}}(p)}{p}\log p = \operatorname{rank}\mathcal{E}(\mathbb{Q}(t)). \tag{2.7}$$

Tate's conjecture is known for rational surfaces (see [RS98]). An elliptic surface $y^2 = x^3 + A(t)x + B(t)$ is rational if and only if one of the following is true:

1) $0 < \max\{3\deg A, 2\deg B\} < 12$;
2) $3\deg A = 2\deg B = 12$ and $\operatorname{ord}_{t=0} t^{12}\Delta(t^{-1}) = 0$.

Theorem 2.3 [Mic95]. *Consider a one-parameter family over $\mathbb{Q}(t)$ with non-constant $j(E_t)$. Then*

$$A_{2,\mathcal{F}}(p) = p^2 + O(p^{3/2}). \tag{2.8}$$

2.5 Sieving and conductors

To evaluate the sums of $\prod_i a_t^{r_i}(p_i)$, it is necessary to restrict t to arithmetic progressions; in order to bound some of the error terms, we will see that the conductors $C(t)$ must be monotone.

Let

$$\begin{aligned}\mathcal{T}_{\text{sqfree}} &= \{t\in[N,2N] : D(t) \text{ is square-free}\}\\ \mathcal{T}_N &= \{t\in[N,2N] :\ d^2\nmid D(t) \text{ for } 2\leqslant d\leqslant \log^l N\}.\end{aligned} \tag{2.9}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Clearly $\mathcal{T}_{\text{sqfree}} \subset \mathcal{T}_N$. We show that $\mathcal{T}_N$ is a union of arithmetic progressions, and $|\mathcal{T}_N - \mathcal{T}_{\text{sqfree}}| = o(N)$.

Thus, except for $o(N)$ values of t, we can write t good (where the conductors are monotone) as a union of arithmetic progressions. For proofs, see Theorems A.5 and B.2.

3. One- and two-level density kernels for the classical compact groups

By [KS99a], the n-level densities for the classical compact groups are

$$\begin{aligned} W_{n,O^+}(x) &= \det(K_1(x_i,x_j))_{i,j\leqslant n} \\ W_{n,O^-}(x) &= \det(K_{-1}(x_i,x_j))_{i,j\leqslant n} + \sum_{k=1}^{n}\delta(x_k)\det(K_{-1}(x_i,x_j))_{i,j\neq k} \\ &= (W_{n,O^-})_1(x) + (W_{n,O^-})_2(x) \\ W_{n,O}(x) &= \tfrac{1}{2}W_{n,O^+}(x) + \tfrac{1}{2}W_{n,O^-}(x) \\ W_{n,U}(x) &= \det(K_0(x_i,x_j))_{i,j\leqslant n} \\ W_{n,Sp}(x) &= \det(K_{-1}(x_i,x_j))_{i,j\leqslant n} \end{aligned} \tag{3.1}$$

where $K(y) = \sin\pi y/\pi y$, $K_\epsilon(x,y) = K(x-y) + \epsilon K(x+y)$ for $\epsilon = 0, \pm 1$, O^+ denotes the group $SO(\text{even})$ and O^- the group $SO(\text{odd})$.

3.1 One-level densities

Let $I(u)$ be the characteristic function of $[-1,1]$.

Theorem 3.1 (One-level densities). *We have*

$$\begin{aligned} \widehat{W_{1,O^+}}(u) &= \delta(u) + \tfrac{1}{2}I(u) \\ \widehat{W_{1,O}}(u) &= \delta(u) + \tfrac{1}{2} \\ \widehat{W_{1,O^-}}(u) &= \delta(u) - \tfrac{1}{2}sI(u) + 1 \\ \widehat{W_{1,Sp}}(u) &= \delta(u) - \tfrac{1}{2}I(u) \\ \widehat{W_{1,U}}(u) &= \delta(u). \end{aligned} \tag{3.2}$$

For functions whose Fourier transforms are supported in $[-1,1]$, the three orthogonal densities are indistinguishable, although they are distinguishable from U and Sp. To detect differences between the orthogonal groups using the one-level density, one needs to work with functions whose Fourier transforms are supported beyond $[-1,1]$.

3.2 Two-level densities

Theorem 3.2 ($\mathcal{G} = SO(\text{even})$, O or $SO(\text{odd})$). *Let $c(\mathcal{G}) = 0, \frac{1}{2}, 1$ for $\mathcal{G} = SO(\text{even}), O, SO(\text{odd})$. For even functions supported in $|u_1| + |u_2| < 1$*

$$\begin{aligned} \iint \widehat{f}_1(u_1)\widehat{f}_2(u_2)\widehat{W_{2,\mathcal{G}}}(u)\,du_1\,du_2 &= [\widehat{f}_1(0) + \tfrac{1}{2}f_1(0)][\widehat{f}_2(0) + \tfrac{1}{2}f_2(0)] + 2\int |u|\widehat{f}_1(u)\widehat{f}_2(u)\,du \\ &\quad - 2\widehat{f_1 f_2}(0) - f_1(0)f_2(0) + c(\mathcal{G})f_1(0)f_2(0). \end{aligned} \tag{3.3}$$

For arbitrarily small support, the three two-level densities differ. One increases by a factor of $\frac{1}{2}f_1(0)f_2(0)$ moving from $\widehat{W_{2,O^+}}$ to $\widehat{W_{2,O}}$ to $\widehat{W_{2,O^-}}$.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

THEOREM 3.3 ($\mathcal{G} = Sp$). *We have*

$$\begin{aligned}\iint \widehat{f}_1(u_1)\widehat{f}_2(u_2)\widehat{W_{2,Sp}}(u)\,du_1\,du_2 &= [\widehat{f}_1(0)+\tfrac{1}{2}f_1(0)][\widehat{f}_2(0)+\tfrac{1}{2}f_2(0)] + 2\int |u|\widehat{f}_1(u)\widehat{f}_2(u)\,du \\ &\quad - 2\widehat{f_1f_2}(0) - f_1(0)f_2(0) - f_1(0)\widehat{f}_2(0) - \widehat{f}_1(0)f_2(0) + 2f_1(0)f_2(0). \end{aligned} \tag{3.4}$$

THEOREM 3.4 ($\mathcal{G} = U$). *We have*

$$\iint \widehat{f}_1(u_1)\widehat{f}_2(u_2)\widehat{W_{2,U}}\,du_1\,du_2 = \widehat{f}_1(0)\widehat{f}_2(0) + \int |u|\widehat{f}_1(u)\widehat{f}_2(u)\,du - \widehat{f_1f_2}(0). \tag{3.5}$$

For test functions with arbitrarily small support, the two-level densities for the classical compact groups are mutually distinguishable.

4. Expansions for the one- and two-level densities for elliptic curve families

For $i = 1$ and 2, let f_i be an even Schwartz function whose Fourier transform is supported in $(-\sigma_i, \sigma_i)$ and $f(x_1, x_2) = f_1(x_1)f_2(x_2)$, $\widehat{f}(u_1, u_2) = \widehat{f}_1(u_1)\widehat{f}_2(u_2)$.

4.1 One-level density: $D_{1,\mathcal{F}}(f)$

We have

$$\begin{aligned} D_{1,\mathcal{F}}(f) &= \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\sum_{\gamma_E^{(j)}} f_1\left(\gamma_E^{(j)}\frac{\log N_E}{2\pi}\right) \\ &= \widehat{f}_1(0) + f_1(0) - 2\sum_p \frac{1}{p}\frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\frac{\log p}{\log N_E}\widehat{f}_1\left(\frac{\log p}{\log N_E}\right)a_E(p) \\ &\quad - 2\sum_p \frac{1}{p^2}\frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\frac{\log p}{\log N_E}\widehat{f}_1\left(\frac{2\log p}{\log N_E}\right)a_E^2(p) + O\left(\frac{\log\log N_E}{\log N_E}\right). \end{aligned} \tag{4.1}$$

As the one-level density sums are sub-calculations which arise in the two-level investigations, we postpone their determination for now.

4.2 Two-level density: $D_{2,\mathcal{F}}(f)$ and $D^*_{2,\mathcal{F}}(f)$

Recall that the two-level density $D_{2,\mathcal{F}}(f)$ is the sum over all indices j_1, j_2 with $j_1 \neq \pm j_2$.

DEFINITION 4.1. The density $D^*_{2,\mathcal{F}}(f)$ differs from the two-level density $D_{2,\mathcal{F}}(f)$ in that j_1 may equal $\pm j_2$.

We first calculate $D^*_{2,\mathcal{F}}(f)$, and then subtract off the contribution from $j_1 = \pm j_2$. Assuming the GRH, we may write the zeros as $1 + i\gamma^{(j)}$, with $\gamma^{(j)} = -\gamma^{(-j)}$. We have

$$\begin{aligned} D^*_{2,\mathcal{F}}(f) &= \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\sum_{j_1}\sum_{j_2} f_1(L\gamma_E^{(j_1)})f_2(L\gamma_E^{(j_2)}) \\ &= \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\prod_{i=1}^{2}\left[\widehat{f}_i(0) + f_i(0) - 2\sum_{p_i}\frac{\log p_i}{\log N_E}\frac{1}{p_i}\widehat{f}_i\left(\frac{\log p_i}{\log N_E}\right)a_E(p_i)\right. \\ &\quad \left. - 2\sum_{p_i}\frac{\log p_i}{\log N_E}\frac{1}{p_i^2}\widehat{f}_i\left(2\frac{\log p_i}{\log N_E}\right)a_E^2(p_i) + O\left(\frac{\log\log N_E}{\log N_E}\right)\right] \\ &= \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\prod_{i=1}^{2}[\widehat{f}_i(0) + f_i(0) + S_{i,1} + S_{i,2}]. \end{aligned} \tag{4.2}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

We use Theorem D.1 to drop the error terms, as they do not contribute in the limit as $|\mathcal{F}| \to \infty$. The astute reader will notice that Theorem D.1 requires us to know the one-level density, and we have postponed that calculation; however, in the process of calculating the two-level density we will determine the required sums for the one-level density (without using Theorem D.1 to evaluate them). Thus, there is no harm in removing the error terms.

There are five types of sums that we need to investigate: $(1/|\mathcal{F}|)\sum_{E\in\mathcal{F}} S_{i,1}$, $(1/|\mathcal{F}|)\sum_{E\in\mathcal{F}} S_{i,2}$, $(1/|\mathcal{F}|)\sum_{E\in\mathcal{F}} S_{1,1}S_{2,1}$, $(1/|\mathcal{F}|)\sum_{E\in\mathcal{F}} S_{1,2}S_{2,2}$ and $(1/|\mathcal{F}|)\sum_{E\in\mathcal{F}} S_{1,1}S_{2,2}$ $(i \neq j)$. In $S_{i,j}$, i refers to which prime (p_1 or p_2) and j refers to the power of $a_E(p_\alpha)$ (1 or 2). The first and the second sums are what we need to calculate the one-level densities.

4.2.1 $j_1 = \pm j_2$. Let $\rho = 1 + i\gamma_E^{(j)}$ be a zero. For a curve with even functional equation, we may label the zeros by

$$\cdots \leqslant \gamma_E^{(-2)} \leqslant \gamma_E^{(-1)} \leqslant 0 \leqslant \gamma_E^{(1)} \leqslant \gamma_E^{(2)} \leqslant \cdots, \quad \gamma_E^{(-k)} = -\gamma_E^{(k)}, \tag{4.3}$$

while for a curve with odd functional equation we label the zeros by

$$\cdots \leqslant \gamma_E^{(-1)} \leqslant 0 \leqslant \gamma_E^{(0)} = 0 \leqslant \gamma_E^{(1)} \leqslant \cdots, \quad \gamma_E^{(-k)} = -\gamma_E^{(k)}. \tag{4.4}$$

We exclude the contribution from $j_1 = \pm j_2$. If an elliptic curve has even functional equation, j_i ranges over all non-zero integers, and $\gamma_E^{(-j)} = -\gamma_E^{(j)}, j \neq -j$. Since the test functions are even, the sum over all pairs (j_1, j_2) with $j_1 = \pm j_2$ is twice the sum over all pairs (j, j), which is $D_{1,E}(f_1 f_2)$, i.e. the one-level density for a curve E with test function $f_1(x)f_2(x)$.

If an elliptic curve has odd functional equation, j_i ranges over all integers. The curve vanishes to odd order at the critical point $s = 1$. Except for one zero (labelled $\gamma_E^{(0)}$), for every non-zero j, $\gamma_E^{(-j)} = -\gamma_E^{(j)}$, and $j \neq -j$. Twice the sum over pairs (j, j) minus the contribution from the pair $(0, 0)$ equals the sum over all pairs (j_1, j_2) with $j_1 = \pm j_2$. Thus, the curves with odd sign contribute $D_{1,E}(f_1 f_2) - f_1(0)f_2(0)$.

Let $\epsilon_E = \pm 1$ be the sign of the functional equation for E, and define

Definition 4.2. $N(\mathcal{F}, -1) = (1/|\mathcal{F}|)\sum_{E\in\mathcal{F}}((1 - \epsilon_E)/2)$, i.e. the percentage of curves with odd sign.

Summing over $E \in \mathcal{F}$ yields $D_{1,\mathcal{F}}(f_1 f_2) - N(\mathcal{F}, -1)f_1(0)f_2(0)$ for $j_1 = \pm j_2$.

4.2.2 *Two-level density expansion.*

Lemma 4.3 (Two-level density expansion). *We have*

$$D_{2,\mathcal{F}}(f) = \frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\prod_{i=1}^{2}[\widehat{f}_i(0) + f_i(0) + S_{i,1} + S_{i,2}]$$
$$- 2D_{1,\mathcal{F}}(f_1 f_2) + (f_1 f_2)(0)N(\mathcal{F}, -1) + O\left(\frac{\log\log N}{\log N}\right). \tag{4.5}$$

To evaluate the above, we only need to know the percentage of curves with odd sign, not which curves are even or odd. For the three- and higher-level densities, we have to execute sums over the subset of curves with odd sign.

4.3 Useful expansion for the one- and two-level densities for one parameter families

Let $\mathcal{E}$ denote a one-parameter family of elliptic curves E_t over $\mathbb{Q}(t)$, $t \in [N, 2N]$, and let $\mathcal{F}$ denote a sub-family of $\mathcal{E}$. In the applications, $\mathcal{F}$ will be obtained by sieving to $D(t)$ good, where $D(t)$ is the product of the irreducible polynomial factors of $\Delta(t)$.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

4.3.1 *Required prime sums.*

LEMMA 4.4 (Prime sums). *Let $C(N)$ be a power of N. By Lemmas C.2, C.3 and C.4:*

i)
$$\sum_p \frac{\log p}{\log C(N)} \frac{1}{p} \widehat{f}_1\left(\frac{\log p}{\log C(N)}\right) = \frac{1}{2} f_1(0) + O\left(\frac{1}{\log N}\right);$$

ii)
$$\sum_p \frac{\log p}{\log C(N)} \frac{1}{p} \widehat{f}_1\left(2\frac{\log p}{\log C(N)}\right) = \frac{1}{4} f_1(0) + O\left(\frac{1}{\log N}\right);$$

iii)
$$\sum_p \frac{\log^2 p}{\log^2 C(N)} \frac{1}{p} \widehat{f}_1 \widehat{f}_2\left(\frac{\log p}{\log C(N)}\right) = \frac{1}{2} \int_{-\infty}^{\infty} |u| \widehat{f}_1(u) \widehat{f}_2(u)\, du + O\left(\frac{1}{\log N}\right).$$

If instead we are summing over primes congruent to a mod m, we use Lemmas C.1 and C.5, and the right-hand sides are modified by $1/\varphi(m)$.

4.3.2 *Expansions of sums.* We use the expansion from Lemma 4.3. Recall that

$$S_{i,j} = -2 \sum_{p_i} \frac{\log p_i}{\log C(t)} \frac{1}{p_i^j} \widehat{f}_i\left(2^{j-1} \frac{\log p_i}{\log C(t)}\right) a_t^j(p_i). \tag{4.6}$$

In $S_{i,j}$, i refers to the prime (p_1, p_2) and j refers to the power of $a_t(p)$ $(a_t(p), a_t^2(p))$.

To determine the one- and two-level densities, there are eight sums over $t \in \mathcal{F}$ to evaluate: $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{1,1}$ and $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{2,1}$; $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{1,2}$ and $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{2,2}$; $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{1,1}S_{2,2}$ and $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{2,1}S_{1,2}$; $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{1,1}S_{2,1}$ and $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{1,2}S_{2,2}$.

We have written the sums in pairs where the two sums are handled similarly. Substituting the definitions leads to five types of sums:

i)
$$-2\sum_p \frac{1}{p} \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \frac{\log p}{\log C(t)} \widehat{f}_1\left(\frac{\log p}{\log C(t)}\right) a_t(p);$$

ii)
$$-2\sum_p \frac{1}{p^2} \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \frac{\log p}{\log C(t)} \widehat{f}_1\left(2\frac{\log p}{\log C(t)}\right) a_t^2(p);$$

iii)
$$4\sum_{p_1}\sum_{p_2} \frac{1}{p_1 p_2^2} \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \frac{\log p_1}{\log C(t)} \frac{\log p_2}{\log C(t)} \widehat{f}_1\left(\frac{\log p}{\log C(t)}\right) \widehat{f}_2\left(2\frac{\log p}{\log C(t)}\right) a_t(p_1) a_t^2(p_2);$$

iv)
$$4\sum_{p_1}\sum_{p_2} \frac{1}{p_1 p_2} \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \frac{\log p_1}{\log C(t)} \frac{\log p_2}{\log C(t)} \widehat{f}_1\left(\frac{\log p}{\log C(t)}\right) \widehat{f}_2\left(\frac{\log p}{\log C(t)}\right) a_t(p_1) a_t(p_2);$$

v)
$$4\sum_{p_1}\sum_{p_2} \frac{1}{p_1^2 p_2^2} \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \frac{\log p_1}{\log C(t)} \frac{\log p_2}{\log C(t)} \widehat{f}_1\left(2\frac{\log p}{\log C(t)}\right) \widehat{f}_2\left(2\frac{\log p}{\log C(t)}\right) a_t^2(p_1) a_t^2(p_2).$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

In the above sums, we use Lemma C.7 to restrict to primes greater than $\log^l N$, $l < 2$. Label the five sums $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S(t;p)$ by $T_k(p)$ and $T_k(p_1,p_2)$. Trivially, by Hasse's bound some of the above do not contribute.

In the third sum, if $p_1 = p_2 = p$, we get a result $\ll (1/\log N)\sum_p (p^{3/2}\log p/p^3) = O(1/\log N)$. In the fifth sum, if $p_1 = p_2 = p$, we get a result $\ll (1/\log N)\sum_p (p^2\log p/p^4) = O(1/\log N)$.

Thus, we only study the third and fifth sums when $p_1 \neq p_2$. The fourth sum has the potential to contribute when $p_1 = p_2$. Hence, we split it into two cases: $p_1 \neq p_2$ and $p_1 = p_2$.

4.3.3 *Conditions on the family to evaluate the sums.*

Conditions 4.5 (On the family $\mathcal{F}$). Let $T_k(p)$ and $T_k(p_1,p_2)$ $(= (1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S(t;p))$ equal:

i)
$$\frac{\log p}{\log C(N)}\widehat{f}_1\left(\frac{\log p}{\log C(N)}\right)\left[-r + O\left(p^{-\alpha} + \frac{p^\beta}{|\mathcal{F}|} + \frac{1}{\log^\gamma N}\right)\right];$$

ii)
$$\frac{\log p}{\log C(N)}\widehat{f}_1\left(2\frac{\log p}{\log C(N)}\right)\left[p + O\left(p^{1-\alpha} + \frac{p^\beta}{|\mathcal{F}|} + \frac{p}{\log^\gamma N}\right)\right];$$

iii)
$$\frac{\log p_1}{\log C(N)}\frac{\log p_2}{\log C(N)}\widehat{f}_1\left(\frac{\log p_1}{\log C(N)}\right)\widehat{f}_2\left(2\frac{\log p_2}{\log C(N)}\right) \times\left[-rp_2 + O\left(p_1^{-\alpha_1}p_2^{1-\alpha_2} + \frac{p_1^{\beta_1}p_2^{\beta_2}}{|\mathcal{F}|} + \frac{p_2}{\log^\gamma N}\right)\right];$$

iv) (a)
$$\frac{\log p_1}{\log C(N)}\frac{\log p_2}{\log C(N)}\widehat{f}_1\left(\frac{\log p_1}{\log C(N)}\right)\widehat{f}_2\left(\frac{\log p_2}{\log C(N)}\right) \times\left[r^2 + O\left(p_1^{1-\alpha_1}p_2^{1-\alpha_2} + \frac{p_1^{\beta_1}p_2^{\beta_2}}{|\mathcal{F}|} + \frac{1}{\log^\gamma N}\right)\right] \quad \text{if } p_1 \neq p_2;$$

(b)
$$\frac{\log^2 p}{\log^2 C(N)}\widehat{f}_1\widehat{f}_2\left(\frac{\log p}{\log C(N)}\right)\left[p + O\left(p^{1-\alpha} + \frac{p^\beta}{|\mathcal{F}|} + \frac{p}{\log^\gamma N}\right)\right] \quad \text{if } p_1 = p_2 = p;$$

v)
$$\frac{\log p_1}{\log C(N)}\frac{\log p_2}{\log C(N)}\widehat{f}_1\left(2\frac{\log p_1}{\log C(N)}\right)\widehat{f}_1\left(2\frac{\log p_2}{\log C(N)}\right) \times\left[p_1p_2 + O\left(p_1^{1-\alpha_1}p_2^{1-\alpha_2} + \frac{p_1^{\beta_1}p_2^{\beta_2}}{|\mathcal{F}|} + \frac{p_1p_2}{\log^\gamma N}\right)\right]$$

where $\alpha, \beta, \gamma > 0$, $\alpha_i, \beta_i \geqslant 0$ and whenever two α_i or β_i occur, at least one is positive.

By Lemma 4.4 we can evaluate the eight $S_{i,j}$ sums for a family satisfying Conditions 4.5 as follows.

Lemma 4.6 ($S_{i,j}$ sums). *If the family satisfies Conditions 4.5, then (up to lower order terms which do not contribute for small support):*

i) $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{i,1} = rf_i(0)$;

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

ii) $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{i,2} = -\frac{1}{2}f_i(0)$;

iii) $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{1,1}S_{2,2} + S_{2,1}S_{1,2} = -\frac{1}{2}rf_1(0)f_2(0) + -\frac{1}{2}rf_1(0)f_2(0)$;

iv) $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{1,1}S_{2,1} = r^2 f_1(0)f_2(0) + 2\int_{-\infty}^{\infty} |u|\widehat{f_1}(u)\widehat{f_2}(u)\,du$;

v) $(1/|\mathcal{F}|)\sum_{t\in\mathcal{F}} S_{1,2}S_{2,2} = \frac{1}{4}f_1(0)f_2(0)$.

4.3.4 *One- and two-level densities, assuming certain conditions on the family.* Substituting Lemma 4.6 into the one- and two-level density expansions we obtain the following.

LEMMA 4.7 (One- and two-level densities). *Assume $|\mathcal{F}|$ is a positive multiple of N and $\mathcal{F}$ satisfies Conditions 4.5. Up to lower order correction terms (which vanish as $|\mathcal{F}| \to \infty$), for even Schwartz functions with small support,*

$$D_{1,\mathcal{F}}(f) = \widehat{f_1}(0) + \tfrac{1}{2}f_1(0) + rf_1(0) \tag{4.7}$$

and

$$\begin{aligned} D_{2,\mathcal{F}}(f) = \prod_{i=1}^{2}[\widehat{f_i}(0) + \tfrac{1}{2}f_i(0)] + 2\int_{-\infty}^{\infty} |u|\widehat{f_1}(u)\widehat{f_2}(u)\,du - 2\widehat{f_1 f_2}(0) - f_1(0)f_2(0) \\ + (f_1 f_2)(0)N(\mathcal{F},-1) + (r^2 - r)f_1(0)f_2(0) + r\widehat{f_1}(0)f_2(0) + rf_1(0)\widehat{f_2}(0). \end{aligned} \tag{4.8}$$

Let $D_{1,\mathcal{F}}^{(r)}(f_1)$ and $D_{2,\mathcal{F}}^{(r)}(f_1)$ be the one- and two-level densities from which the contributions of r family zeros at the critical point have been subtracted. Then

$$D_{1,\mathcal{F}}^{(r)}(f_1) = \widehat{f_1}(0) + \tfrac{1}{2}f_1(0) \tag{4.9}$$

and

$$D_{2,\mathcal{F}}^{(r)}(f_1) = \prod_{i=1}^{2}[\widehat{f_i}(0) + \tfrac{1}{2}f_i(0)] + 2\int_{-\infty}^{\infty} |u|\widehat{f_1}(u)\widehat{f_2}(u)\,du - 2\widehat{f_1 f_2}(0) - f_1(0)f_2(0) + (f_1 f_2)(0)N(\mathcal{F},-1). \tag{4.10}$$

Thus, removing the contribution from r family zeros, for test functions of small support the two-level density of the remaining zeros agrees with SO(even) if all curves are even, O if half are even and half odd, and SO(odd) if all are odd.

Proof. The one-level density is immediate from substitution. Substituting for the eight $S_{i,j}$ sums for $D_{2,\mathcal{F}}(f)$ yields (up to lower order terms which do not contribute for small support)

$$\begin{aligned} D_{2,\mathcal{F}}(f) &= \prod_{i=1}^{2}[\widehat{f_i}(0) + f_i(0)] + [\widehat{f_1}(0) + f_1(0)]rf_2(0) + [\widehat{f_2}(0) + f_2(0)]rf_1(0) + r^2 f_1(0)f_2(0) \\ &\quad + 2\int_{-\infty}^{\infty} |u|\widehat{f_1}(u)\widehat{f_2}(u)\,du + [\widehat{f_1}(0) + f_1(0)]\left(-\frac{1}{2}f_2(0)\right) + [\widehat{f_2}(0) + f_2(0)]\left(-\frac{1}{2}f_1(0)\right) \\ &\quad - \frac{1}{2}rf_1(0)f_2(0) - \frac{1}{2}rf_1(0)f_2(0) + \frac{1}{4}f_1(0)f_2(0) \\ &\quad - 2D_{1,\mathcal{F}}(f_1 f_2) + (f_1 f_2)(0)N(\mathcal{F},-1) + O\left(\frac{\log\log N}{\log N}\right) \\ &= \prod_{i=1}^{2}\left[\widehat{f_i}(0) + \frac{1}{2}f_i(0)\right] + 2\int_{-\infty}^{\infty} |u|\widehat{f_1}(u)\widehat{f_2}(u)\,du \\ &\quad + 2rf_1(0)f_2(0) + r\widehat{f_1}(0)f_2(0) + rf_1(0)\widehat{f_2}(0) - rf_1(0)f_2(0) + r^2 f_1(0)f_2(0) \\ &\quad - 2D_{1,\mathcal{F}}(f_1 f_2) + (f_1 f_2)(0)N(\mathcal{F},-1). \end{aligned} \tag{4.11}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Substituting

$$D_{1,\mathcal{F}}(f_1 f_2) = \widehat{f_1 f_2}(0) + \tfrac{1}{2} f_1(0) f_2(0) + r f_1(0) f_2(0) \tag{4.12}$$

yields

$$\begin{aligned} D_{2,\mathcal{F}}(f) &= \prod_{i=1}^{2} [\widehat{f}_i(0) + \tfrac{1}{2} f_i(0)] + 2\int_{-\infty}^{\infty} |u| \widehat{f}_1(u) \widehat{f}_2(u)\, du + r f_1(0) f_2(0) + r\widehat{f}_1(0) f_2(0) + r f_1(0)\widehat{f}_2(0) \\ &\quad + r^2 f_1(0) f_2(0) - 2\widehat{f_1 f_2}(0) - f_1(0) f_2(0) - 2r f_1(0) f_2(0) + (f_1 f_2)(0) N(\mathcal{F}, -1) \\ &= \prod_{i=1}^{2} [\widehat{f}_i(0) + \tfrac{1}{2} f_i(0)] + 2\int_{-\infty}^{\infty} |u| \widehat{f}_1(u) \widehat{f}_2(u)\, du - 2\widehat{f_1 f_2}(0) - f_1(0) f_2(0) \\ &\quad + (f_1 f_2)(0) N(\mathcal{F}, -1) + (r^2 - r) f_1(0) f_2(0) + r\widehat{f}_1(0) f_2(0) + r f_1(0)\widehat{f}_2(0). \end{aligned} \tag{4.13}$$

If the family has rank r over $\mathbb{Q}(t)$, there is a natural interpretation of these terms. By the Birch and Swinnerton-Dyer conjecture (only used for interpretation purposes) and Silverman's specialization theorem, for all t sufficiently large, each curve's L-function has at least r zeros at the critical point. We isolate the contributions from r family zeros.

Assume that there are r family zeros at the critical point. Let $L_t = \log C(t)/2\pi$. Recall that the one-level density is $D_{1,\mathcal{F}}(f) = \widehat{f}(0) + \frac{1}{2} f(0) + r f(0)$. Let j_i range over all zeros of a curve, and j_i' range over all but the r family zeros. We have

$$\begin{aligned} &D_{2,\mathcal{F}}(f) \\ &= \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \sum_{j_1} \sum_{j_2} f_1(L_t \gamma_{E_t}^{(j_1)}) f_2(L_t \gamma_{E_t}^{(j_2)}) - 2D_{1,\mathcal{F}}(f_1 f_2) + (f_1 f_2)(0) N(\mathcal{F}, -1) \\ &= \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \left(r f_1(0) + \sum_{j_1'} f_1(L_t \gamma_{E_t}^{(j_1')}) \right) \left(r f_2(0) + \sum_{j_2'} f_2(L_t \gamma_{E_t}^{(j_2')}) \right) \\ &\quad - 2D_{1,\mathcal{F}}(f_1 f_2) + (f_1 f_2)(0) N(\mathcal{F}, -1) \\ &= \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \sum_{j_1'} \sum_{j_2'} f_1(L_t \gamma_{E_t}^{(j_1')}) f_2(L_t \gamma_{E_t}^{(j_2')}) + r f_1(0) D_{1,\mathcal{F}}(f_2) + D_{1,\mathcal{F}}(f_1) r f_2(0) - r^2 f_1(0) f_2(0) \\ &\quad - 2D_{1,\mathcal{F}}(f_1 f_2) + (f_1 f_2)(0) N(\mathcal{F}, -1) \\ &= \frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \sum_{j_1'} \sum_{j_2'} f_1(L_t \gamma_{E_t}^{(j_1')}) f_2(L_t \gamma_{E_t}^{(j_2')}) + (f_1 f_2)(0) N(\mathcal{F}, -1) + r f_1(0) \left(\widehat{f}_2(0) + \left(r + \frac{1}{2} \right) f_2(0) \right) \\ &\quad + \left(\widehat{f}_1(0) + \left(r + \frac{1}{2} \right) f_1(0) \right) r f_2(0) - r^2 f_1(0) f_2(0) - 2 \left(\widehat{f_1 f_2}(0) + \frac{1}{2} f_1(0) f_2(0) + r f_1(0) f_2(0) \right) \\ &= \left[\frac{1}{|\mathcal{F}|} \sum_{t\in\mathcal{F}} \sum_{j_1'} \sum_{j_2'} f_1(L_t \gamma_{E_t}^{(j_1')}) f_2(L_t \gamma_{E_t}^{(j_2')}) - 2 \left(\widehat{f_1 f_2}(0) + \frac{1}{2} f_1(0) f_2(0) \right) + (f_1 f_2)(0) N(\mathcal{F}, -1) \right] \\ &\quad + r f_1(0)\widehat{f}_2(0) + r\widehat{f}_1(0) f_2(0) + (r^2 - r) f_1(0) f_2(0) \\ &= D_{2,\mathcal{F}}^{(r)}(f_1) + r f_1(0)\widehat{f}_2(0) + r\widehat{f}_1(0) f_2(0) + (r^2 - r) f_1(0) f_2(0). \end{aligned} \tag{4.14}$$

□

We isolate the following.

Lemma 4.8. *The contribution from r critical point zeros is*

$$r f_1(0)\widehat{f}_2(0) + r\widehat{f}_1(0) f_2(0) + (r^2 - r) f_1(0) f_2(0). \tag{4.15}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

5. Calculation of the one- and two-level densities for elliptic curve families

Let $\mathcal{E}$ be a one-parameter family of elliptic curves E_t with discriminants $\Delta(t)$ and conductors $C(t)$. For many families, we can evaluate the conductors exactly if we sieve to a subfamily $\mathcal{F}$ defined as the $t \in [N, 2N]$ with $D(t)$ good, where $D(t) = a_k t^k + \cdots + a_0$ ($a_k \geqslant 1$) is the product of the irreducible polynomial factors of $\Delta(t)$. Good usually means square-free, although occasionally it means square-free except for a fixed set of primes, and for these special primes the power of $p|D(t)$ is independent of t.

Let our family $\mathcal{F}$ be the set of good $t \in [N, 2N]$ where the conductors are given by a monotone polynomial in t. We use this polynomial for the conductors at non-good t; this is permissible as these curves are not in our family and do not originally appear in our sums.

For each d, let

$$T(d) = \{t \in [N, 2N] : d^2 | D(t)\}. \tag{5.1}$$

Let $S(t)$ be some quantity associated to the elliptic curve E_t. We study

$$\sum_{\substack{t=N \\ D(t) \text{ good}}}^{2N} S(t) = \sum_{d=1}^{(2a_k N)^{k/2}} \mu(d) \sum_{t \in T(d)} S(t). \tag{5.2}$$

In particular, setting $S(t) = 1$ yields the cardinality of the family. In all the families we investigate, $|\mathcal{F}| = c_{\mathcal{F}} N + o(N)$, $c_{\mathcal{F}} > 0$.

Let $t_1(d), \ldots, t_{\nu(d)}(d)$ be the incongruent roots of $D(t) \equiv 0 \bmod d^2$. The presence of $\mu(d)$ allows us to restrict to d square-free. For small d, we may take the $t_i(d) \in [N, N + d^2)$. For such d,

$$\sum_{t \in T(d)} S(t) = \sum_{i=1}^{\nu(d)} \sum_{t'=0}^{[N/d^2]} S(t_i(d) + t' d^2) + O(\nu(d) \|S\|_\infty). \tag{5.3}$$

The error piece is from boundary effects for the last value of t'. $T(d)$ restricts us to $t \in [N, 2N]$; as each $t_i(d) \geqslant N$, and at most one is exactly N, it is possible in summing to $t' = [N/d^2]$ that we have added an extra term.

5.1 Assumptions for sieving

We evaluate the sums under the following assumptions:

i) for square-free $D(t)$, the conductors $C(t)$ are given by a monotone polynomial in t;

ii) a positive percentage of $t \in [N, 2N]$ have $D(t)$ square-free, i.e. $|\mathcal{F}| = c_{\mathcal{F}} N + o(N)$.

We constantly use Lemma A.2 ($\nu(d) \ll d^\epsilon$ for square-free d) and

$$\sum_{\substack{t=N \\ D(t) \text{ good}}}^{2N} 1 = \sum_{d=1}^{\log^l N} \mu(d) \sum_{\substack{t=N \\ D(t) \equiv 0 (d^2)}}^{2N} 1 + o(N) = c_{\mathcal{F}} N + o(N), \quad c_{\mathcal{F}} > 0. \tag{5.4}$$

We show that the family satisfies Conditions 4.5. We evaluate the sums over $t \in \mathcal{F}$ below and then execute the summation over the prime(s). $\widehat{f}_i$ is supported in $(-\sigma_i, \sigma_i)$. There are no contributions (for σ_i sufficiently small) in the prime sum(s) for sufficiently small error terms.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

5.2 Definition of terms for sieving

Recall $A_{r,\mathcal{F}}(p) = \sum_{t(p)} a_t^r(p)$. For distinct primes, by Lemma 2.1

$$\sum_{t(p_1\cdots p_n)} \prod_{j=1}^{n} a_t^{r_i}(p_j) = \prod_{j=1}^{n} A_{r_i,\mathcal{F}}(p_i). \tag{5.5}$$

By Lemma C.7, we may assume that all of our primes (in the expansion from the explicit formula in the n-level densities) are at least $\log^l N$, $l \in [1,2)$. We can incorporate these errors into our existing error terms; the result will still be a lower order term which will not contribute for small support.

$S(t)$ will equal $\widetilde{a}_P(t)G_P(t)$, where for distinct primes p_1 and p_2

$$\begin{aligned}
\widetilde{a}_P(t) &= a_t^{r_1}(p_1)a_t^{r_2}(p_2)\\
G_P(t) &= \prod_{\substack{j=1\\ r_j\neq 0}}^{2} \frac{\log p_j}{\log C(t)} f_j\left(2^{r_j-1}\frac{\log p_j}{\log C(t)}\right)\\
(r_1,r_2) &\in \{(1,0),(0,1),(2,0),(1,1),(0,2),(1,2),(2,1),(2,2)\}.
\end{aligned} \tag{5.6}$$

Thus $\widetilde{a}_P(t)G_P(t)$ is merely a convenient way of encoding the eight sums we need to examine for the one- and two-level densities.

Actually, this is slightly off. We have to study

$$\prod_{\substack{j=1\\ r_j\neq 0}}^{2} \frac{1}{p_j^{r_j}} \frac{\log p_j}{\log C(t)} g_j\left(2^{r_j-1}\frac{\log p_j}{\log C(t)}\right) a_t^{r_j}(p_j). \tag{5.7}$$

If both r_j are non-zero and the two primes are equal, we obtain

$$\frac{1}{p^{r_1+r_2}}\left(\frac{\log p}{\log C(t)}\right)^2 \times \cdots \times a_t^{r_1+r_2}(p). \tag{5.8}$$

For example, if $r_1 = r_2 = 1$ we would get $(\log p/\log C(t))^2 \times \cdots \times a_t^2(p)$. Thus, the definition of G_P needs to be slightly modified. We want to deal with distinct primes p_1 and p_2. There will be no contribution for equal primes if $r_1 + r_2 \geqslant 3$; simply bound each $a_t(p)$ by Hasse. There is a contribution if $r_1 = r_2 = 1$. By modifying the definition of G_P we may regard it as a case where $r = (2,0)$; however, we have $(\log p/\log C(t))^2$ instead of $(\log p/\log C(t))$, and instead of $f_1(\cdots)$ we have $f_1 f_2(\cdots)$. Note that we evaluate the test functions at $\log p/\log C(t)$ and not $2(\log p/\log C(t))$. We have

$$G_P(t) = \prod_{\substack{j=1\\ r_j\neq 0}}^{2} \left(\frac{\log p_j}{\log C(t)}\right)^{\kappa(r)} g_j\left(2^{r_j-\kappa(r)}\frac{\log p_j}{\log C(t)}\right), \tag{5.9}$$

where $\kappa(r)$ is 2 if $r = (2,0)$, and this arises from $p_1 = p_2 = p$, and $\kappa(r) = 1$ otherwise; $g_j = f_j$ unless $r = (2,0)$ arising from $p_1 = p_2 = p$, in which case $g_1 = f_1 f_2$.

We may now assume that the primes are distinct. Define

$$\begin{aligned}
P &= \prod_{\substack{j=1\\ r_j\neq 0}}^{2} p_j, \quad r = (r_1,r_2),\ r_j \in \{0,1,2\}\\
S_c(r,P) &= \sum_{t(P)} \widetilde{a}_P(t) = \sum_{t(P)} a_t^{r_1}(p_1)a_t^{r_2}(p_2) = A_{r_1,\mathcal{F}}(p_1)A_{r_2,\mathcal{F}}(p_2),
\end{aligned} \tag{5.10}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

where for convenience we set $A_0(p) = 1$. We often have incomplete sums of $\widetilde{a}_P(t)$ mod P. Let $S_I(r, P)$ denote a generic incomplete sum. By Hasse,

$$S_I(r,P) \leqslant P \cdot 2^{r_1}\sqrt{p_1^{r_1}} \cdot 2^{r_2}\sqrt{p_2^{r_2}} = 2^{r_1+r_2}p_1^{1+r_1/2} \cdot p_2^{1+r_2/2} = 2^r P^{1+r/2}, \tag{5.11}$$

where the last expression is a convenient abuse of notation:

$$2^r = 2^{r_1+r_2}, \quad P^r = p_1^{r_1} \cdot p_2^{r_2}. \tag{5.12}$$

For fixed i and d, we evaluate the arguments at $t = t_i(d) + t'd^2$. Let

$$\widetilde{a}_{d,i,P}(t') = \widetilde{a}_P(t_i(d) + t'd^2), \quad G_{d,i,P}(t') = G_P(t_i(d) + t'd^2). \tag{5.13}$$

5.3 Ranges and contributions of sums over primes

Each prime sum is to (approximately) $C(N)^{\sigma_j/2^{r_j-\kappa(r)}} \approx N^{m\sigma_j/2^{r_j-\kappa(r)}}$, as $C(t)$ is a degree m polynomial. We assume $\sigma_j < \frac{1}{2}$ as we do not worry about $p^2 > N$. This is harmless, as handling the error terms forces the support to be significantly less than $\frac{1}{2}$.

Lemma 5.1 (Contributions from sums over primes). *For $r_j = 1$, summing $p^{1/2}/|\mathcal{F}|$ does not contribute for $\sigma_j < 2/3m$. For $r_j = 2$, summing $1/|\mathcal{F}|$ does not contribute for $\sigma_j < 2/m$ for $\kappa(r) = 1$ and $1/m$ for $k(r) = 2$. As we often have two sums, dividing the above supports by two ensures that all errors are manageable: write $1/|\mathcal{F}|$ as $(1/\sqrt{|\mathcal{F}|})(1/\sqrt{|\mathcal{F}|})$.*

5.3.1 *Expected result.* To simplify the proof, we assume

$$\begin{aligned} A_{1,\mathcal{F}}(p) &= -rp + O(1) \\ A_{2,\mathcal{F}}(p) &= p^2 + O(p^{3/2}). \end{aligned} \tag{5.14}$$

For a general rational surface, $A_{1,\mathcal{F}}(p) \neq -rp + O(1)$; however, an analysis of the arguments below shows that we only need to be able to handle sums such as

$$\sum_p \frac{\log p}{\log X} f\left(\frac{\log p}{\log X}\right) \frac{A_{1,\mathcal{F}}(p)}{p^2}. \tag{5.15}$$

For surfaces where Tate's conjecture is known, we may replace $A_{1,\mathcal{F}}(p)$ in the above sum with the rank of the family over $\mathbb{Q}(t)$ (see Lemma C.6 and [RS98]). For notational simplicity, in the proof below we assume that $A_{1,\mathcal{F}}(p) = -rp + O(1)$, and content ourselves with noting that a similar proof works in general.

We have $A_{r_j}(p_j) = c_j \cdot p_j^{r_j}$ plus lower order terms not contributing for any support. (This is not quite true. For families where the curves have complex multiplication, $a_t(p)$ often vanishes for half the primes, and has double the expected contribution for the other primes. This case is handled similarly, using Lemmas C.1 and C.5.)

Hence, $S_c(r, P) = c_1c_2p_1^{r_1}p_2^{r_2} = c_1c_2P^r$ plus lower terms. For each pair (d, i) we expect (if we can manage the conductors) to have approximately $(N/d^2)/P$ complete sums of $S_c(r, P) = c_1c_2P^r$. We hit this with $(1/N)(\log p_j/\log C(t))(1/p_j^{r_j})$ for each non-zero r_j. We have approximately $(\log p_j/\log C(t))(1/P^r)$.

A sum like $\sum_{p_j}(\log p_j/\log C(t))(1/p_j)g(\log p_j/\log C(t))$ contributes; if we had an additional $1/\log N$ there would be no net contribution. Thus, we expect terms of the size P^r to contribute and $P^r/\log N$ not to contribute.

We rewrite Conditions 4.5 in a more tractable form, using $A_{1,\mathcal{F}}(p)$, $A_{2,\mathcal{F}}(p)$ and $S_c(r, P)$. Assume that the family satisfies Equation (5.14) (or the related equation if $a_t(p)$ vanishes for half the primes).

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Then

i) $P = p$, $\widetilde{a}_P(t) = a_t(p)$:
$$\frac{S_c(r,P)}{P} = \frac{-rp + O(1)}{p} = -r + O\left(\frac{1}{p}\right);$$

ii) $P = p$, $\widetilde{a}_P(t) = a_t^2(p)$:
$$\frac{S_c(r,P)}{P} = \frac{p^2 + O(p^{3/2})}{p} = p + O(\sqrt{p});$$

iii) $P = p_1p_2$, $\widetilde{a}_P(t) = a_t(p_1)a_t^2(p_2)$:
$$\frac{S_c(r,P)}{P} = \frac{-rp_1p_2^2 + O(p_1p_2^{3/2})}{p_1p_2} = -rp_2 + O(\sqrt{p_2});$$

iv) $P = p_1p_2$, $\widetilde{a}_P(t) = a_t(p_1)a_t(p_2)$:

(a)
$$\frac{S_c(r,P)}{P} = \frac{r^2p_1p_2 + O(p_1+p_2)}{p_1p_2} = r^2 + O(\sqrt{p_1}+\sqrt{p_2}) \quad \text{if } p_1 \neq p_2;$$

(b)
$$\frac{S_c(r,P)}{P} = \frac{p^2 + O(p^{3/2})}{p} = p + O(\sqrt{p}) \quad \text{if } p_1 = p_2 = p;$$

v) $P = p_1p_2$, $\widetilde{a}_P(t) = a_t^2(p_1)a_t^2(p_2)$:
$$\frac{S_c(r,P)}{P} = \frac{p_1^2p_2^2 + O(p_1^{3/2}p_2^{3/2})}{p_1p_2} = p_1p_2 + O(\sqrt{p_1p_2}).$$

We have proved the following.

Lemma 5.2 (Conditions to evaluate the five types of sums). *Assume that the family satisfies Equation (5.14). If, up to lower order terms, the five sums (Conditions 4.5) are $G_P(N)(S_c(r,P)/P)$, then the family satisfies Conditions 4.5.*

5.4 Taylor expansion of $G_{d,i,P}(t')$

Fix i and d. We calculate the first-order Taylor expansion of $G_{d,i,P}(t') = G_P(t_i(d)+t'd^2)$. $G_{d,i,P}$ only involves t' through expressions like $\log p_j/\log C(t)$, where $t = t_i(d)+t'd^2$. Let $C(t) = h_mt^m+\cdots+h_0$.

The derivative of $G_{d,i,P}$ in t' will involve nice functions times factors like
$$\begin{aligned}\frac{d}{dt'}\frac{\log p_j}{\log C(t)} &= -\frac{\log p_j}{\log^2 C(t)}\frac{d}{dt'}\log C(t_i(d)+t'd^2)\\ &= -\frac{\log p_j}{\log^2 C(t)}\frac{mh_mt^{m-1}d^2+\cdots}{h_mt^{m-1}\cdot(t_i(d)+t'd^2)+\cdots}\\ &\leqslant \left(\frac{10m}{|h_m|}\max_{0\leqslant k\leqslant m-1}|m-k|\cdot|h_{m-k}|\right)\frac{\log p_j}{\log^2 C(t)}\frac{d^2}{t_i(d)+t'd^2},\end{aligned} \tag{5.16}$$
provided that N is sufficiently large.

As $p_j \leqslant C(t)^\sigma$, where σ is related to the support of G, $\log p_j/\log C(t) \leqslant \sigma$. As $C(t)$ is of size a power of t, we have the following.

Lemma 5.3 (Taylor expansion of $G_{d,i,P}$). *We have*
$$G_{d,i,P}(t') = G_{d,i,P}(0) + O\left(\frac{1}{\log N}\right). \tag{5.17}$$
The constant above does not depend on p_j, d or i.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

By the mean value theorem there exist $\xi \in [0, t']$, corresponding to $t_\xi = t_i(d) + \xi d^2 \in [N, 2N + d^2] \subset [N, 2.1N]$, such that

$$G_{d,i,P}(t') = G_{d,i,P}(0) + \frac{d}{dt'} G_{d,i,P}\Big|_{t'=\xi} (t' - 0). \tag{5.18}$$

First, we have derivatives of $\log p_j / \log C(t)$ which can be universally bounded from the support of G. Second, we evaluate G and its derivative at $2^{r_j - \kappa(r)}(\log p_j / \log C(t_\xi))$. We see it is sufficient to universally bound functions like $(d/dt')g(\log p / \log C(t))$.

We have $\log C(t_\xi) \approx \log C(N)$. Evaluating the derivative at ξ, by Equation (5.16) we have something bounded by $(1/\log C(t_\xi))(d^2/t_i(d) + \xi d^2)$. We then multiply by $t' - 0$. Thus we are bounded by $(1/\log C(N))(t' d^2/t_i(d) + \xi d^2)$. As $t_i(d) \geqslant N$ and $t' d^2 \leqslant N$, the bound is at most $1/\log C(N)$.

LEMMA 5.4 (Further Taylor expansion of $G_{d,i,P}$). *We have*

$$G_{d,i,P}(t') = G_P(N) + O\left(\frac{1}{\log N}\right). \tag{5.19}$$

The constant above does not depend on p_j, d or i.

The proof is similar to the previous lemma. $G_{d,i,P}(0) = G_P(t_i(d))$, $t_i(d) \in [N, N + d^2]$. Thus, to replace $G_{d,i,P}(0)$ with $G_P(N)$ involves Taylor expanding $G_P(t)$ around $t = N$.

This allows us to replace all the conductors of curves with $D(t)$ good with the value from $t = N$ with small error. This is very convenient, as $G_P(N)$ has no t', i or d dependence. Consequently, we are able to move it past all summations except over primes, which allows us to take advantage of cancellations in t-sums of the values $a_t(p)$.

5.5 Removing the $\nu(d)\|S\|_\infty$ term for $d < \log^l N$

We have

$$\sum_{t \in T(d)} S(t) = \sum_{i=1}^{\nu(d)} \sum_{t'=0}^{[N/d^2]} S(t_i(d) + t' d^2) + O(\nu(d)\|S\|_\infty). \tag{5.20}$$

We show that the $O(\nu(d)||S||_\infty)$ piece does not contribute for $d < \log^l N$. Using Hasse to trivially bound $\|S\|_\infty$ gives $2^r P^r$. We hit this with $1/P^r$ and sum over the primes, which will be at most $O(N^\sigma)$. We now sum over $d < \log^l N$, obtaining a value

$$\ll N^\sigma \sum_{d=1}^{\log^l N} \nu(d) \ll N^\sigma \sum_{d=1}^{\log^l N} d^\epsilon \ll N^\sigma \log^{l(1+\epsilon)} N. \tag{5.21}$$

We then divide by the cardinality of the family, which is assumed to be a multiple of N. There is no contribution for $\sigma_1 + \sigma_2 < 1$.

5.6 Sieving

Let B be the largest square which divides $D(t)$ for all t. Recall by t good we mean that $D(t)$ is square-free except for primes dividing B, and for $p|B$ the power of $p|D(t)$ is independent of t. By Theorem A.5, possibly after passing to a subsequence, we can approximate t good by

$$\sum_{t \in [N,2N] \ t \text{ good}} S(t) = \sum_{\substack{d=1 \\ (d,B)=1}}^{\log^l N} \mu(d) \sum_{\substack{t \in [N,2N] \\ D(t) \equiv 0 (d^2)}} S(t) + O\Big(\sum_{t \in \mathcal{T}} S(t)\Big), \tag{5.22}$$

where the set of t good is $c_{\mathcal{F}} N + o(N)$, $c_{\mathcal{F}} > 0$, $\mathcal{T}$ is the set of $t \in [N, 2N]$ such that $D(t)$ is divisible by the square of a prime $p > \log^l N$ and $|\mathcal{T}| = o(N)$.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

5.7 Contributions from $d < \log^l N$

We would like to use Lemma 5.4 to replace $G_{d,i,P}(t')$ with $G_P(N)$ plus a manageable error. This works for pairs such as $r = (2,0)$ or $r = (2,2)$ but fails for pairs such as $r = (1,0)$, so there we need to evaluate $(1/|\mathcal{F}|)\sum_{E\in\mathcal{F}}(1/p)S(r,p)$. Replacing $\widetilde{a}_p(t)$ with $|a_t(p)| \leqslant 2\sqrt{p}$ gives a value

$$\ll \frac{1}{|\mathcal{F}|}\frac{N}{p}\sqrt{p}, \tag{5.23}$$

which is disastrous when we sum over p. The reason we must trivially bound $\widetilde{a}_P(t)$ is the Taylor expansion. We evaluate the derivative at $\xi(t') = \xi(p_j, i, d; t')$. The dependence of the other parameters prevents us from obtaining complete sums (mod P) and using that cancellation for control. We need to keep the cancellation from summing $\widetilde{a}_P(t)$.

We use partial summation twice. Note that we may always replace a $G_{d,i,P}(t')$ with a $G_P(N)$ at a cost of $1/\log N$.

Let $\widetilde{A}_P(u) = \sum_{t'=0}^{u}\widetilde{a}_P(t')$. As $(p_i, d) = 1$ (this is why we are assuming $d \leqslant \log^l N$ and $p_i \geqslant \log^l N$), every time t' increases by P we have a complete sum of the $\widetilde{a}_P$. Thus,

$$\begin{aligned}\widetilde{A}_P(u) &= \left[\frac{u}{P}\right] S_c(r,P) + O(P^{1+r/2}) = \frac{u}{P}S_c(r,P) + O(P^R)\\ R &= 1 + \frac{r}{2}, \quad P^R = \prod_{\substack{j=1\\ r_j\neq 0}}^{2} p_j^{1+r_j/2}.\end{aligned} \tag{5.24}$$

In the above, the first error term is from our bound for the incomplete sum of at most P terms, each term bounded by $\sqrt{p_1^{r_1}p_2^{r_2}} = P^{r/2}$. Dropping the greatest integer brackets costs at most $S_c(r,P) = O(P^r)$. We have $P^r = p_1^{r_1}p_2^{r_2}$ and $P^{1+r/2} = p_1^{1+r_1/2}p_2^{1+r_2/2}$. As $r_j \in \{0,1,2\}$, $r_j \leqslant 1 + r_j/2$. Thus, we may incorporate the error from removing the greatest integer brackets into the $O(P^R)$ term.

$$\begin{aligned}S(d,i,r,P) &= \sum_{t'=0}^{[N/d^2]} \widetilde{a}_{d,i,P}(t')G_{d,i,P}(t')\\ &= \left(\frac{[N/d^2]}{P}S_c(r,P) + O(P^R)\right) G_{d,i,P}\left(\left[\frac{N}{d^2}\right]\right)\\ &\quad - \sum_{u=0}^{[N/d^2]-1}\left(\frac{u}{P}S_c(r,P) + O(P^R)\right)(G_{d,i,P}(u) - G_{d,i,P}(u+1))\end{aligned}$$

$$S(r,P) = \sum_{d=1}^{\log^l N}\mu(d)\sum_{i=1}^{\nu(d)} S(d,i,r,P) = \sum_{w=1}^{4}\sum_{d=1}^{\log^l N}\mu(d)\sum_{i=1}^{\nu(d)} S_w(d,i,r,P). \tag{5.25}$$

5.7.1 *First sum:* $([N/d^2]/P)S_c(r,P)G_{d,i,P}([N/d^2])$. Summing over i and d yields

$$\begin{aligned}S_1(r,P) &= \sum_{d=1}^{\log^l N}\mu(d)\sum_{i=1}^{\nu(d)}\frac{[N/d^2]}{P}S_c(r,P)G_{d,i,P}\left(\left[\frac{N}{d^2}\right]\right)\\ &= \frac{S_c(r,P)}{P}\sum_{d=1}^{\log^l N}\mu(d)\sum_{i=1}^{\nu(d)}\left[\frac{N}{d^2}\right]\left(G_P(N) + O\left(\frac{1}{\log N}\right)\right)\end{aligned}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

$$= \frac{S_c(r,P)G_P(N)}{P}\sum_{d=1}^{\log^l N}\mu(d)\sum_{i=1}^{\nu(d)}\sum_{t'=0}^{[N/d^2]}\left(1+O\left(\frac{1}{\log N}\right)\right)$$

$$= \frac{S_c(r,P)G_P(N)}{P}\sum_{d=1}^{\log^l N}\mu(d)\left(O(\nu(d))+\sum_{\substack{t=N\\ D(t)\equiv 0(d^2)}}^{2N}1\right)\left(1+O\left(\frac{1}{\log N}\right)\right)$$

$$= \frac{S_c(r,P)G_P(N)}{P}|\mathcal{F}| + \frac{S_c(r,P)}{P}\cdot o(N). \tag{5.26}$$

In the last line, the error term follows from Equation (5.4) (which gives the d, t-sums are $|\mathcal{F}|+o(N)$) and Lemma A.2 (which gives $\nu(d)\ll d^{\epsilon}$). Dividing by $|\mathcal{F}| = c_{\mathcal{F}}N + o(N)$, the error term will not contribute when we sum over primes, leaving us with $S_c(r,P)G_P(N)/P$.

5.7.2 *Second sum:* $O(P^R)G_{d,i,P}([N/d^2])$. Summing over i and d yields

$$S_2(r,P) \ll \sum_{d=1}^{\log^l N}|\mu(d)|\sum_{i=1}^{\nu(d)}P^R\left|G_{d,i,P}\left(\left[\frac{N}{d^2}\right]\right)\right|$$

$$\ll P^R\sum_{d=1}^{\log^l N}|\mu(d)|\sum_{i=1}^{\nu(d)}\|G\|_{\infty}$$

$$\ll P^R\sum_{d=1}^{\log^l N}|\mu(d)|\sum_{i=1}^{\nu(d)}1. \tag{5.27}$$

As $\nu(d)\ll d^{\epsilon}$, we obtain

$$S_2(r,P) \ll P^R\log^{l(1+\epsilon)}N \leqslant P^R\log^{2l}N = P^{1+r/2}\log^{2l}N. \tag{5.28}$$

We divide by $|\mathcal{F}| = c_{\mathcal{F}}N + o(N)$, hit it with $1/P^r$ and then sum over the primes. By Lemma 5.1, for small support ($\sigma = \sigma_1+\sigma_2 < 2/3m$) there is no contribution.

5.7.3 *Third sum:* $\sum_{u=0}^{[N/d^2]-1}(u/P)S_c(r,P)(G_{d,i,P}(u)-G_{d,i,P}(u+1))$. We apply partial summation, where $a_u = G_{d,i,P}(u)-G_{d,i,P}(u+1)$ and $b_u = (u/P)S_c(r,P)$. Thus

$$S_3(d,i,r,P) = \left(G_{d,i,P}(0)-G_{d,i,P}\left(\left[\frac{N}{d^2}\right]\right)\right)\frac{[N/d^2]-1}{P}S_c(r,P)$$
$$-\sum_{u=0}^{[N/d^2]-2}(G_{d,i,P}(0)-G_{d,i,P}(u+1))\frac{1}{P}S_c(r,P). \tag{5.29}$$

Using the Taylor expansion, we gain a $1/\log N$ in the first term, making it of size

$$\frac{S_c(r,P)}{P}\frac{[N/d^2]}{\log N} \ll \frac{S_c(r,P)}{P}\frac{|\mathcal{F}|}{d^2\log N}.$$

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

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

We sum over i and d and obtain

$$\begin{aligned} S_3(r,P) &\ll \sum_{d=1}^{\log^l N} |\mu(d)| \sum_{i=1}^{\nu(d)} \frac{S_c(r,P)}{P} \frac{|\mathcal{F}|}{d^2 \log N} \\ &\ll \frac{S_c(r,P)}{P} \frac{|\mathcal{F}|}{\log N} \sum_{d=1}^{\log^l N} \sum_{i=1}^{\nu(d)} \frac{1}{d^2} \\ &\ll \frac{S_c(r,P)}{P} \frac{|\mathcal{F}|}{\log N} \sum_{d=1}^{\log^l N} \frac{\nu(d)}{d^2}. \end{aligned} \tag{5.30}$$

As $\nu(d) \ll d^\epsilon$, $S_3(r,P) \ll (S_c(r,P)/P)(|\mathcal{F}|/\log N)$.

5.7.4 *Fourth sum:* $\sum_{u=0}^{[N/d^2]-1} O(P^R)(G_{d,i,P}(u) - G_{d,i,P}(u+1))$. Using the Taylor expansion for $G_{d,i,P}(u) - G_{d,i,P}(u+1)$ is insufficient. That gives $NP^R/d^2 \log N$. Summing over i and d is manageable, giving $O(P^R|\mathcal{F}|/\log N)$. Dividing by the cardinality of the family yields $O(P^R/\log N)$.

The problem is in summing over the primes, as we no longer have $1/|\mathcal{F}|$. We multiply by $1/P^r$. We recall the definitions of r and R and unwind the above.

Consider the case $r = (1,0)$. Then $P = p_1 = p$, $R = 1 + r_1/2 = 3/2$, and $1/P^r = 1/p$. We have

$$\sum_{p=\log^l N}^{N^{m\sigma}} \frac{1}{p} \frac{p^{3/2}}{\log N} \gg N^{m\sigma}. \tag{5.31}$$

As $N \to \infty$, this term diverges. We need significantly better cancellation in

$$S_4(r,P) = \sum_{d=1}^{\log^l N} \mu(d) \sum_{i=1}^{\nu(d)} \sum_{u=0}^{[N/d^2]-1} O(P^R)(G_{d,i,P}(u) - G_{d,i,P}(u+1)). \tag{5.32}$$

Taking absolute values and using the maximum of the $O(P^R)$ terms gives

$$S_4(r,P) \ll P^R \sum_{d=1}^{\log^l N} \sum_{i=1}^{\nu(d)} \sum_{u=0}^{[N/d^2]-1} |G_{d,i,P}(u) - G_{d,i,P}(u+1)|. \tag{5.33}$$

The constant is independent of P. Taking the maximum of the P^R term involves the maximum of either the incomplete sum or one complete sum. Using Hasse, the constant is at most $2^{r_1+r_2}$. Thus, the constant in Equation (5.33) does not depend on P.

If exactly one of the r_j is non-zero, then

$$G_{d,i,P}(u) - G_{d,i,P}(u+1) = g\left(\frac{\log p}{\log C(t_i(d) + ud^2)}\right) - g\left(\frac{\log p}{\log C(t_i(d) + (u+1)d^2)}\right) \tag{5.34}$$

for some Schwartz function g of compact support.

If both of the r_j are non-zero, we may write $G_{d,i,P}(u)$ as the product of two functions, say g_1 and g_2. Thus

$$G_{d,i,P}(u) = \prod_{j=1}^{2} g_j\left(\frac{\log p_j}{\log C(t_i(d) + ud^2)}\right). \tag{5.35}$$

Recall that

$$\begin{aligned} |a_1a_2 - b_1b_2| &= |a_1a_2 - b_1a_2 + b_1a_2 - b_1b_2| \\ &\leqslant |a_1a_2 - b_1a_2| + |b_1a_2 - b_1b_2| = |a_2| \cdot |a_1 - b_1| + |b_1| \cdot |a_2 - b_2|. \end{aligned} \tag{5.36}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

We apply the above to our function $G_{d,i,P}(u) = g_1(d,i,p_1;u)g_2(d,i,p_2;u)$. Each $g_j(d,i,p_j;u)$ can be bounded independently of d, i, p_j and u, as each g_j is a Schwartz function defined in terms of the n-level density test functions. Let $B = \max_j \|g_j\|_\infty$. Then

$$\begin{aligned} S_4(d,i,r,P)(u) &= G_{d,i,P}(u) - G_{d,i,P}(u+1) \\ &= \prod_{\substack{j=1\\ r_j\neq 0}}^{2} g_j\left(\frac{\log p_i}{\log C(t_i(d)+ud^2)}\right) - \prod_{\substack{j=1\\ r_j\neq 0}}^{2} g_j\left(\frac{\log p_j}{\log C(t_i(d)+(u+1)d^2)}\right) \\ &\leqslant \sum_{\substack{j=1\\ r_j\neq 0}}^{2} B\cdot\left|g_j\left(\frac{\log p_j}{\log C(t_i(d)+ud^2)}\right) - g_j\left(\frac{\log p_j}{\log C(t_i(d)+(u+1)d^2)}\right)\right|. \end{aligned} \tag{5.37}$$

We sum the above over u, i and d. Let $t_{i,d}(u) = t_i(d) + ud^2$. We have

$$\begin{aligned} S_4(r,P) &\leqslant 2^r P^R \sum_{d=1}^{\log^l N} |\mu(d)| \sum_{i=1}^{\nu(d)} \sum_{u=0}^{[N/d^2]-1} S_4(d,i,r,P)(u) \\ &\leqslant 2^r P^R \sum_{d=1}^{\log^l N} \sum_{i=1}^{\nu(d)} \sum_{\substack{j=1\\ r_j\neq 0}}^{2} B \sum_{u=0}^{[N/d^2]-1} \left|g_j\left(\frac{\log p_j}{\log C(t_{i,d}(u))}\right) - g_j\left(\frac{\log p_j}{\log C(t_{i,d}(u+1))}\right)\right|. \end{aligned} \tag{5.38}$$

We show the u-sums are bounded independent of p_j, i, d, and N. We may add

$$\left|g_j(0) - g_j\left(\frac{\log p_j}{\log C(t_i(d))}\right)\right| + \left|g_j\left(\frac{\log p_j}{\log C(t_i(d)+[N/d^2]d^2)}\right) - g_j(1000\sigma)\right|. \tag{5.39}$$

As each g_j is a Schwartz function, they are of bounded variation. Let $x_u(d,i,p_j) = p_j/\log N_{t_i(d)+ud^2}$. As the conductors are monotone increasing, $x_u(d,i,p_j) > x_{u+1}(d,i,p_j)$. Thus, we have a partition of $[0,1000\sigma]$, and we may now apply theorems on bounded variation to bound the u-sum independent of p_j, i, d and N, obtaining $\ll 1000\sigma$.

The above is an exercise in the bounded variation of $g(x)$ on $[0,\sigma]$. If we were to regard this as a problem in the bounded variation of $g_{j;p_j,d,i}$ we would have u ranging over at least $[0,[N/d^2]]$. Even though we would gain a $1/\log N$ from the derivatives, the bounded variation bound depends on the size of the interval, which here is of length $[N/d^2]$. We could also argue that each g_j has continuous, bounded first derivative on $[0,1000\sigma]$. By the mean value theorem, the u-sum is $\ll \|g_j'\|_\infty \cdot |1000\sigma - 0|$.

Thus, the u- and j-sums are universally bounded. We are left with $\ll P^R$. Summing over i and d gives $\ll P^R \log^{l(1+\epsilon)} N$. We multiply by $1/P^r$ and sum over the primes. The prime sums give $N^{h(\sigma)}$; dividing by the cardinality of the family (a multiple of N), we find there is no contribution for small support.

Note that if our conductors are not monotone, we cannot apply theorems on bounded variation. The problem is that we could transverse $[0,1000\sigma]$ (or a large subset of it) up to N/d^2 times. This is why S_4 is the most difficult of the error pieces, and why we needed to obtain polynomial expressions for the conductors for t good.

5.7.5 *Summary of contributions for $d < \log^l N$.*

Lemma 5.5 (Contributions for $d < \log^l N$). *Based on our sieving assumptions for the family (for $D(t)$ good the conductors are given by a monotone polynomial in t, a positive percentage of $t \in [N,2N]$ give $D(t)$ good), the main term contribution from $d < \log^l N$ is $(S_c(r,P)/P)G_P(N)|\mathcal{F}|$. The error terms are either of size $(S_c(r,P)/P)o(|\mathcal{F}|)$, which will not contribute when we sum over primes, or are such that their sum over primes will not contribute.*

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

5.8 Contributions from $t \in \mathcal{T}$

5.8.1 *Preliminaries.* We are left with estimating the contributions from the troublesome set

$$\mathcal{T} = \{t \in [N, 2N] : \exists d > \log^l N \text{ with } d^2 | D(t)\}. \tag{5.40}$$

In Theorem A.5 we show that $|\mathcal{T}| = o(N)$. By Cauchy–Schwartz

$$\left|\sum_{t\in\mathcal{T}} S(t)\right| \leqslant \left(\sum_{t\in\mathcal{T}} S^2(t)\right)^{1/2}\left(\sum_{t\in\mathcal{T}} 1\right)^{1/2} \leqslant \left(\sum_{t=N}^{2N} S^2(t)\right)^{1/2} o(\sqrt{N}). \tag{5.41}$$

We then sum over the primes, and need to show that the sum over t is $O(N)$. As it stands, however, this is not sufficient to control the error. As a quick sketch; assume that $S(t) = a_t(p)g(\log p/\log C(t))$. Ignoring the t-dependence in the conductors, we have

$$\begin{aligned}\sum_{t=N}^{2N} S(t) &\approx g^2\left(\frac{\log p}{\log C(N)}\right)\frac{N}{p}\sum_{t(p)} a_t^2(p)\\ &\approx g^2\left(\frac{\log p}{\log C(N)}\right)\frac{N}{p}p^2 = O(Np).\end{aligned} \tag{5.42}$$

Taking the square-root, we hit it with $1/p$ and sum over $p \leqslant N^\sigma$, which is not $O(\sqrt{N})$.

We have that $S(t)$ is the product of at most two terms involving factors such as $a_t^{r_j}(p_j)$. We hit this with factors $p_j^{-r_j}$ and sum over p. Thus, instead of $S(t)$ consider $S_1(t)S_2(t)$, where $S_j(t)$ incorporates the sum over primes to the jth power and all relevant factors. We have

$$\begin{aligned}S &= \sum_{t=N}^{2N}\left[\prod_{\substack{j=1\\ r_j\neq 0}}^{2}\sum_{p_j\geqslant \log^l N} p_j^{-r_j} g_j\left(\frac{\log p_j}{\log C(t)}\right)a_t^{r_j}(p_j)\right]^2\\ &= \sum_{t=N}^{2N}\prod_{w=1}^{2}\prod_{\substack{j=1\\ r_j\neq 0}}^{2}\sum_{p_{j_w}\geqslant \log^l N} p_{j_w}^{-r_{j_w}} g_{j_w}\left(\frac{\log p_{j_w}}{\log C(t)}\right)a_t^{r_{j_w}}(p_{j_w}).\end{aligned} \tag{5.43}$$

We proceed in a similar manner as in the $d \leqslant \log^l N$ case, except now there are no d and i, and we potentially have four factors instead of one or two. On expanding, we combine terms where we have the same prime occurring multiple times. There are several types of sums: four distinct primes (four factors), three distinct primes (three factors), ..., all primes the same (one factor). We proceed with the worst case, when there are four factors; the other cases are handled similarly.

5.8.2 *A specific case: four distinct primes.* Assume that we have four distinct primes. Relabelling, we have $p^{-r_i}a_t^{r_i}(p_i)$ for $i = 1$ to 4. Let $P = \prod_{i=1}^4 p_i$. Interchange the t-summation with the p_i-summations. As before, we apply partial summation to $\sum_{t=N}^{2N}\prod_{i=1}^4 a_t^{r_i}(p_i)\cdot g_i(p_i,t)p^{-r_i} = \sum_{t=N}^{2N} a(P,t)\cdot b(P,t)$, the only change being the addition of the factors $\prod_i p^{-r_i}$. Now

$$A(u) = \sum_{t=N}^{u} a(P,t) = \frac{u-N}{P}S_c(r,P) + O\left(\prod_{i=1}^4 p_i^{1+r_i/2}\right), \quad S_c(r,P) = \prod_{i=1}^4 A_{r_i,\mathcal{F}}(p_i)$$

by Lemma 2.1. Let $P^R = \prod_{i=1}^4 p_i^{1+r_i/2}$; the error in the partial summation is $O(P^R)$.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

As in Equation (5.25) we have

$$\begin{aligned} S &= \prod_{i=1}^{4} \sum_{p_i} \sum_{t=N}^{2N} a_t^{r_i}(p_i) \cdot p^{-r_i} G(P,t) \\ &= \prod_{i=1}^{4} \sum_{p_i} \left(\frac{N}{P} S_c(r,P) + O(P^R) \right) p_i^{-r_i} G(P,2N) \\ &\quad - \prod_{i=1}^{4} \sum_{p_i} \sum_{u=N}^{2N-1} \left(\frac{u-N}{P} S_c(r,P) + O(P^R) \right) p_i^{-r_i} (G(P,u) - G(P,u+1)). \end{aligned} \tag{5.44}$$

For $r \geqslant 2$, by Hasse $A_{r,\mathcal{F}}(p) \leqslant 2^r p^{1+r/2}$. For $r = 1$, $A_{1,\mathcal{F}}(p) \ll p$ by [Del80]. Hence, for all r, $A_{r,\mathcal{F}}(p) \ll p^r$. We have

$$\prod_{i=1}^{4} \frac{S_c(P)}{p_i} p_i^{-r_i} \ll \prod_{i=1}^{4} \frac{A_{r_i,\mathcal{F}}(p_i)}{p_i^{1+r_i}} \ll \prod_{i=1}^{4} \frac{p_i^{r_i}}{p_i^{1+r_i}} = \prod_{i=1}^{4} \frac{1}{p_i}. \tag{5.45}$$

We can immediately handle the first sum. Inserting absolute values yields something like

$$\prod_{i=1}^{4} \sum_{p_i} \frac{\log p_i}{\log C(2N)} \left| g_i\left(\frac{\log p_i}{\log C(2N)} \right) \right| \frac{1}{p_i} \ll \prod_{i=1}^{4} O(1) \tag{5.46}$$

where the last result (the sums over the primes) follows from Corollary C.2.

Pulling out the prime factors and using partial summation again, the third sum is handled similarly.

The second and fourth pieces are more difficult, and result in significantly decreased support. We analyze this loss later. For now, we need only note that the second sum is $\prod_i \sum_{p_i} p_i^{r_i/2}$. For test functions of small support, this sum is $o(N)$.

There is a slight obstruction in applying the same argument to the fourth sum; namely, that $G(P,u)$ could be the product of four factors. Similar to the identity $|a_1a_2 - b_1b_2| \leqslant |a_1| \cdot |a_1 - b_1| + |b_1| \cdot |a_2 - b_2|$, we have

$$\begin{aligned} &|a_1a_2a_3a_4 - b_1b_2b_3b_4| \\ &\quad \leqslant |a_2a_3a_4| \cdot |a_1 - b_1| + |b_1a_3a_4| \cdot |a_2 - b_2| + |b_1b_2a_4| \cdot |a_3 - b_3| + |b_1b_2b_3| \cdot |a_4 - b_4| \\ &\quad \leqslant \prod_{j=1}^{4} (|a_j| + |b_j| + 1) \sum_{i=1}^{4} |a_i - b_i|. \end{aligned} \tag{5.47}$$

The rest of the proof in this case is identical to the fourth sum in the $d \leqslant \log^l N$ case.

Note that as we have always inserted absolute values before summing over primes, it is permissible to extend from the primes that are distinct to all possible 4-tuples.

5.8.3 *Handling the other cases.* The other cases (especially cases where some primes are equal) are handled similarly. The only real change is if we have less than four factors, and this only affects the fourth sum. For example, if we have three factors instead of four, set $a_4 = b_4 = 1$ in Equation (5.47).

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

5.9 Determining the admissible supports of the test functions

The largest errors arise from $r_i = 1$ terms, using Hasse to trivially bound partial sums of $a_t(p)$ by $p^{3/2}$ (at most p terms, each term at most $2\sqrt{p}$). Let $C(t)$ be a polynomial of degree m for t good. We assume that all supports are at most $\frac{1}{2}$ (as otherwise p^2 could exceed N, changing some of our arguments above). In the one-level densities, we encounter errors like

$$\sum_{p=\log^l N}^{N^{\sigma m}} \frac{1}{p}\frac{\log p}{\log N^m} g\left(\frac{\log p}{\log N^m}\right) p^{3/2} \ll \sum_{p=\log^l N}^{N^{\sigma m}} p^{1/2} \ll N^{3\sigma m/2}. \tag{5.48}$$

We divide by $|\mathcal{F}|$, a multiple of N. The errors are negligible for $\sigma < \min(2/3m, 1/2)$.

In the two-level density, the worst case (not including the Cauchy–Schwartz arguments used to handle the over-counting of almost square-free numbers) was when we had two $r_i = 1$ terms. We have two functions of support σ_1 and σ_2, and we obtain

$$\prod_{i=1}^{2} \sum_{p_i=\log^l N}^{N^{\sigma_i m}} \frac{1}{p}\frac{\log p_i}{\log N^m} g\left(\frac{\log p_i}{\log N^m}\right) p_i^{3/2} \ll \prod_{i=1}^{2} \sum_{p_i=\log^l N}^{N^{\sigma m}} p_i^{1/2} \ll N^{3(\sigma_1+\sigma_2)m/2}. \tag{5.49}$$

We divide by a multiple of N and see the errors are negligible for $\sigma_1 + \sigma_2 < \min(2/3m, \frac{1}{2})$. Thus, for $\sigma_1 = \sigma_2$, the support of each test function is half that from the one-level density.

In applying Cauchy–Schwartz, we further decrease the allowable support. The worst case is where we have four distinct primes with $r_i = 1$. We sum as before and obtain $N^{3(\sigma_1+\sigma_2)m}$ (there is no factor of 2 as two of the primes are associated with test functions with support σ_1 and two with σ_2). We take the square-root and this must be $O(\sqrt{N})$. Thus, we now find $\sigma_1 + \sigma_2 < \frac{1}{2}(2/3m)$. Setting $\sigma_1 = \sigma_2$ yields that the support is one-quarter that of the one-level density.

5.10 One- and two-level densities

Assume that the original family has rank r over $\mathbb{Q}(t)$. The Birch and Swinnerton-Dyer conjecture and Silverman's specialization theorem imply, for all t sufficiently large, that each curve's L-function has r family zeros at the critical point.

The Birch and Swinnerton-Dyer conjecture is only used for interpretation purposes. The results below are derived independently of this conjecture, however, assuming that this allows us to interpret some of the n-level density terms as contributions from expected family zeros.

Definition 5.6 (Non-family density). Let $D^{(r)}_{n,\mathcal{F}}(f)$ be the n-level density from the non-family zeros (i.e., the trivial contributions from r family zeros have been removed).

Theorem 5.7 ($D_{n,\mathcal{F}}(f)$ and $D^{(r)}_{n,\mathcal{F}}(f)$, $n = 1$ or 2). *For any one-parameter family of rank r over $\mathbb{Q}(t)$ satisfying:*

i) *for t good (relative to $D(t)$), the conductors $C(t)$ are a monotone polynomial in t;*

ii) *up to $o(N)$, the good $t \in [N, 2N]$ are obtainable by sieving up to $d = \log^l N$; further, the number of such t is $|\mathcal{F}| = c_{\mathcal{F}} N + o(N)$, $c_{\mathcal{F}} > 0$;*

iii) *$A_{1,\mathcal{F}}(p) = -rp + O(1)$, $A_{2,\mathcal{F}}(p) = p^2 + O(p^{3/2})$;*

then, for f_i even Schwartz functions of small but non-zero support σ_i,

$$\begin{aligned} D_{1,\mathcal{F}}(f) &= \widehat{f}_1(0) + \tfrac{1}{2} f_1(0) + r f_1(0), \\ D^{(r)}_{1,\mathcal{F}}(f_1) &= \widehat{f}_1(0) + \tfrac{1}{2} f_1(0), \end{aligned} \tag{5.50}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

and

$$\begin{aligned} D_{2,\mathcal{F}}(f) &= \prod_{i=1}^{2}[\widehat{f_i}(0)+\tfrac{1}{2}f_i(0)]+2\int_{-\infty}^{\infty}|u|\widehat{f_1}(u)\widehat{f_2}(u)\,du-2\widehat{f_1f_2}(0)-f_1(0)f_2(0)\\ &\quad+(f_1f_2)(0)N(\mathcal{F},-1)+(r^2-r)f_1(0)f_2(0)+r\widehat{f_1}(0)f_2(0)+rf_1(0)\widehat{f_2}(0),\\ D^{(r)}_{2,\mathcal{F}}(f_1) &= \prod_{i=1}^{2}[\widehat{f_i}(0)+\tfrac{1}{2}f_i(0)]+2\int_{-\infty}^{\infty}|u|\widehat{f_1}(u)\widehat{f_2}(u)\,du-2\widehat{f_1f_2}(0)-f_1(0)f_2(0)\\ &\quad+(f_1f_2)(0)N(\mathcal{F},-1). \end{aligned} \tag{5.51}$$

Removing the contribution from r family zeros, for small support the two-level density of the remaining zeros agrees with SO(even), O or SO(odd) if the signs are all even, equidistributed or all odd. If Tate's conjecture is true for the surface, we may interpret r as the rank of $\mathcal{E}$ over $\mathbb{Q}(t)$.

Let $m = \deg C(t)$. For the one-level density, $\sigma < \min(\frac{1}{2}, 2/3m)$. For the two-level density, $\sigma_1+\sigma_2 < 1/3m$. For families where $\Delta(t)$ has no irreducible factors of degree 4 or more, the sieving is unconditional, otherwise the results are conditional on ABC or the square-free sieve conjecture.

Proof. When we sieve we obtain $S_c(r,P)G_P(N)/P$ plus lower order terms. By Theorem 5.2, the family satisfies Conditions 4.5. Thus Lemma 4.7 is applicable. □

As remarked, we do not need to assume $A_{1,\mathcal{F}}(p) = -rp+O(1)$. A more cumbersome proof (using Lemma C.6) handles $A_{1,\mathcal{F}}(p)$ for surfaces where Tate's conjecture is known.

To apply Theorem 5.7, we need:

i) that the conductors are monotone polynomials for $D(t)$ good;

ii) that a positive percentage of $D(t)$ are good, and all but $o(N)$ of the good t may be taken in the required arithmetic progressions;

iii) knowledge of $A_{1,\mathcal{F}}(p)$ and $A_{2,\mathcal{F}}(p)$.

For rational surfaces, by passing to a subsequence the above conditions are satisfied. By changing $t \to ct+t_0$, Tate's algorithm yields that $C(t)$ is a monotone polynomial for $D(t)$ good (see Theorem B.2). By Theorem A.5, $|\mathcal{F}| = c_{\mathcal{F}}N+o(N)$, $c_{\mathcal{F}} > 0$ (i.e. a positive percentage of $D(t)$ are good). If Tate's conjecture is true, Theorem 2.2 gives $A_{1,\mathcal{F}}(p)$; if $j(E_t)$ is non-constant, Michel's theorem (Theorem 2.3) gives $A_{2,\mathcal{F}}(p)$. We have proved the following.

Theorem 5.8 (Rational surfaces density theorem). *Consider a one-parameter family of elliptic curves of rank r over $\mathbb{Q}(t)$ that is a rational surface. Assume the GRH, $j(E_t)$ is non-constant, and the ABC or square-free sieve conjecture if $\Delta(t)$ has an irreducible polynomial factor of degree at least 4. Let f_i be an even Schwartz function of small but non-zero support σ_i and $m = \deg C(t)$. For the one-level density, $\sigma < \min(\frac{1}{2}, 2/3m)$. For the two-level density, $\sigma_1+\sigma_2 < 1/3m$. Assume the Birch and Swinnerton-Dyer conjecture for interpretation purposes.*

Let $M(t)$ be the product of the irreducible polynomials dividing $\Delta(t)$ and not $c_4(t)$. If $M(t)$ is non-constant, then the signs of E_t, t good, are equidistributed as $N \to \infty$ (see [Hel]). In this case, $N(\mathcal{F},-1) = \frac{1}{2}$.

After passing to a subsequence,

$$\begin{aligned} D_{1,\mathcal{F}}(f_1) &= \widehat{f_1}(0)+\tfrac{1}{2}f_1(0)+rf_1(0),\\ D^{(r)}_{1,\mathcal{F}}(f_1) &= \widehat{f_1}(0)+\tfrac{1}{2}f_1(0), \end{aligned} \tag{5.52}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

and

$$D_{2,\mathcal{F}}(f) = \prod_{i=1}^{2}[\widehat{f_i}(0) + \tfrac{1}{2}f_i(0)] + 2\int_{-\infty}^{\infty}|u|\widehat{f_1}(u)\widehat{f_2}(u)\,du - 2\widehat{f_1 f_2}(0) - f_1(0)f_2(0)$$
$$+ (f_1 f_2)(0)N(\mathcal{F},-1) + (r^2 - r)f_1(0)f_2(0) + r\widehat{f_1}(0)f_2(0) + rf_1(0)\widehat{f_2}(0),$$
$$D^{(r)}_{2,\mathcal{F}}(f_1) = \prod_{i=1}^{2}[\widehat{f_i}(0) + \tfrac{1}{2}f_i(0)] + 2\int_{-\infty}^{\infty}|u|\widehat{f_1}(u)\widehat{f_2}(u)\,du - 2\widehat{f_1 f_2}(0) - f_1(0)f_2(0)$$
$$+ (f_1 f_2)(0)N(\mathcal{F},-1). \tag{5.53}$$

The two-level non-family density is SO(even) *(SO*(odd), *O) if all curves are even (odd, the signs are equidistributed).*

Thus, for small support, the one- and two-level non-family density agrees with the predictions of Katz and Sarnak; further, the densities confirm that the curves' L-functions behave in a manner consistent with having r zeros at the critical point, as predicted by the Birch and Swinnerton-Dyer conjecture.

6. Examples

6.1 Constant sign families

We consider several families where the sign of the functional equation is always positive or negative. We verify the Katz–Sarnak predictions [KS99a, KS99b], assuming only the GRH.

6.1.1 $\mathcal{F} : y^2 = x^3 + 2^4(-3)^3(9t+1)^2$, $9t+1$ *square-free.* Let $\mathcal{F} : y^2 = x^3 + 2^4(-3)^3(9t+1)^2$, $t \in [N, 2N]$, $9t+1$ square-free. Note that $y^2 = x^3 + 2^4(-3)^3D^2$ is equivalent to $y^3 = x^3 + Dz^3$. Birch and Stephens [BS66] calculate the sign of the functional equation for $y^3 = x^3 + Dz^3$, D cube-free. It is

$$\epsilon_{E_D} = -w_3 \prod_{p\neq 3} w_p, \tag{6.1}$$

where $w_3 = -1$ if $D \equiv \pm 1, \pm 3(9)$ and 1 otherwise, $w_p = -1$ if $p|D$, $p \equiv 2(3)$ and 1 otherwise, and D is cube-free.

Consider $D = D(t) = 9t+1$. This result modulo 9 is 1, so $-w_3$ is 1. Assume that a prime congruent to 2 mod 3 divides $9t+1$. If there were only one such prime, the remaining primes would be congruent to 1 mod 3, and the product over all primes dividing $9t+1$ would be congruent to 2 mod 3, a contradiction. Hence, the number of primes congruent to 2 mod 3 dividing $9t+1$ is even. For $9t+1$ square-free, this proves the functional equation is even.

Applying Tate's algorithm (see [Mil02]), we find the conductors $C(t)$ are $3^3(9t+1)^2$ for $9t+1$ square-free. We have $\delta_D = 1$, $k = 1$, $a_k = 9$ so $\mathcal{P} = \{2, 3\}$. As $\nu(2) = 1$ and $\nu(3) = 0$, by Theorem A.5 $c_{\mathcal{F}} > 0$.

For $p \equiv 2(3)$, $x \to x^3$ is an automorphism and $a_t(p) = 0$. Therefore, in the following we assume that all primes are congruent to 1 mod 3, for any sum involving a prime congruent to 2 mod 3 is zero.

For $p > 3$ and $p \equiv 1$ mod 3, direct calculation gives

$$\begin{aligned} A_{1,\mathcal{F}}(p) &= 0 \\ A_{2,\mathcal{F}}(p) &= 2p^2 - 2p = 2p^2 + O(p). \end{aligned} \tag{6.2}$$

From Michel's theorem (Theorem 2.3), we expect $A_{2,\mathcal{F}}(p) = p^2 + O(p^{3/2})$; however, his theorem is only applicable for non-constant $j(E_t)$. As $j(E_t)$ is constant, we must directly compute $A_{2,\mathcal{F}}(p)$. Further, as $a_t(p)$ trivially vanishes for half of the primes, we expect and observe twice the predicted

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

contribution at the other primes. Finally, we will see later that the correction term to $A_{2,\mathcal{F}}(p)$ contributes a potential lower order term to the density functions.

By Dirichlet's theorem for primes in arithmetic progressions (using Lemma C.1 instead of Corollaries C.2 and C.3), we see that the factors of 2 compensate for the restriction to primes congruent to 1 mod 3, and this will be harmless in the applications.

Thus, the family satisfies the conditions of Theorem 5.8 with $r = 0$. We verify (for small support) the Katz–Sarnak predictions. As all the signs are even, conditional only on the GRH, we observe SO(even) symmetry, which is distinguishable from SO(odd) and O symmetry.

6.1.2 $\mathcal{F} : y^2 = x^3 \pm 4(4t+2)x$, $4t+2$ *square-free.* Let $\mathcal{F} : y^2 = x^3 + 4(4t+2)x$, $4t+2$ square-free. We need to study sums of $((x^3 \pm 4(4t+2)x)/p)$. For $p > 2$, changing variables by $t \to t - 2^{-1}$, $t \to \pm 16^{-1}t$, we are led to study sums of $((x^3 + tx)/p)$. If $p \equiv 3 \bmod 4$, then $(-1/p) = -1$. Changing variables $x \to -x$ shows that $a_t(p) = -\sum_{x(p)}(f_t(x)/p)$ vanishes; therefore, in the following we only consider $p \equiv 1 \bmod 4$.

Birch and Stephens [BS66] calculate the sign of the functional equation for this family. For general D, D not divisible by 4 or any fourth power, the sign of the functional equation for the curve $y^2 = x^3 + 4Dx$ is

$$w_\infty w_2 \prod_{p^2 \| D} w_p, \tag{6.3}$$

where $w_\infty = \operatorname{sgn}(-D)$, $w_2 = -1$ if $D \equiv 1, 3, 11, 13 \bmod 16$ and 1 otherwise, $w_p = -1$ for $p \equiv 3(4)$ and $w_p = 1$ for other $p \geqslant 3$.

By restricting to positive, even, square-free D, we force the sign of the functional equation to be odd. Hence, $\epsilon_D = -1$ if $D = 4t + 2$, D square-free. If we had taken $D = -(4t + 2)$, $4t + 2$ square-free, we would have found $\epsilon_D = +1$.

From Tate's algorithm, for $D(t) = \pm(4t + 2)$ square-free, $C(t) = 2^6(4t + 2)^2$. We have $\delta_D = 1$, $k = 1$, $a_k = 4$, so $\mathcal{P} = \{2\}$. As $\nu(2) = 0$, by Theorem A.5 $c_{\mathcal{F}} > 0$.

For $p > 2$ and $p \equiv 1 \bmod 4$, direct calculation gives

$$\begin{aligned} A_{1,\mathcal{F}}(p) &= 0 \\ A_{2,\mathcal{F}}(p) &= 2p^2 - 2p = 2p^2 + O(p). \end{aligned} \tag{6.4}$$

For the family $\mathcal{F}_\pm : y^2 = x^3 \pm 4(4t + 2)x$, $4t + 2$ square-free, all curves in $\mathcal{F}_-$ have even sign and all curves in $\mathcal{F}_+$ have odd sign. The families satisfy the conditions of Theorem 5.8 with $r = 0$. We verify (for small support) the Katz–Sarnak predictions. As all the signs are even (odd), conditional only on the GRH, we observe SO(even) (SO(odd)) symmetry.

6.1.3 $\mathcal{F} : y^2 = x^3 + tx^2 - (t + 3)x + 1$. For this family (due to Washington [Was87])

$$\begin{aligned} c_4(t) &= 2^4(t^2 + 3t + 9) \\ \Delta(t) &= 2^4(t^2 + 3t + 9)^2 \\ j(E_t) &= 2^8(t^2 + 3t + 9). \end{aligned} \tag{6.5}$$

Washington ([Was87]) proved that the rank is odd for $t^2 + 3t + 9$ square-free, assuming the finiteness of the Tate–Shafarevich group. Rizzo [Riz] proved that the rank is odd for all t. While $j(E_t)$ is non-constant, $M(t) = 1$ ($M(t)$ is the product of all irreducible polynomials dividing $\Delta(t)$ but not $c_4(t)$). Thus, Helfgott's results on the equidistribution of sign are not applicable.

For sieving convenience, we replace t with $12t + 1$. Let $D(t) = 144t^2 + 60t + 13$. Tate's algorithm yields $C(t) = 2^3(144t^2 + 60t + 13)$ for $D(t)$ square-free.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

We have $\delta_D = -2^4 3^5$, $k = 2$, $a_k = 2^4 3^2$, so $\mathcal{P} = \{2, 3\}$. The value $D(t)$ is a primitive integral polynomial. For $p \nmid 6$ the number of incongruent solutions of $D(t) \equiv 0 \bmod p^2$ equals the number of incongruent solutions of $D(t) \equiv 0 \bmod p$ (see [Nag81]). As $\nu(2) = \nu(3) = 0$, by Theorem A.5, $c_{\mathcal{F}} > 0$.

Direct calculation gives

$$A_{1,\mathcal{F}}(p) = -p\left[1 + \left(\frac{-1}{p}\right)\right]. \tag{6.6}$$

Hence, $A_{1,\mathcal{F}}(p)$ is $-2p$ for $p \equiv 1(4)$ and 0 for $p \equiv 3(4)$. By Theorem 2.2, the rank over $\mathbb{Q}(t)$ is 1.

As $j(E_t)$ is non-constant, by Michel's theorem $A_{2,\mathcal{F}}(p) = p^2 + O(p^{3/2})$.

The conditions of Theorem 5.8 are satisfied with $r = 1$. We again verify the Katz–Sarnak predictions: there are two pieces to our densities. The first equals the contribution from one zero at the critical point; the second agrees with SO(odd) for small support.

6.2 Rational families

We give two examples of rational families of elliptic curves over $\mathbb{Q}(t)$. See [Mil02] for the proofs, as well as a new method to generate rational families of moderate rank.

6.2.1 *Rank* 1 *example.* Consider the rational family $y^2 = x^3 + 1 + tx^2$:

$$\begin{aligned} c_4(t) &= 16t^2 \\ \Delta(t) &= -16(4t^3 + 27) \\ j(E_t) &= -256\frac{t^6}{4t^3 + 27} \\ M(t) &= 4t^3 + 27. \end{aligned} \tag{6.7}$$

If we replace t with $6t + 1$, we can easily calculate the conductors for $D(t) = 4(6t + 1)^3 + 27$ square-free. In [Mil02], $C(t) = 2^2(4(6t + 1)^3 + 27)$ was shown for $D(t)$ square-free. By Hooley [Hoo76, Theorem 3, p. 69], as $D(t)$ is an irreducible polynomial of degree 3, $c_{\mathcal{F}} > 0$.

Direct calculations [Mil02] gives $A_{1,\mathcal{F}}(p) = -p$, and a more involved calculation gives

$$A_{2,\mathcal{F}}(p) = p^2 - 3ph_{3,p}(2) - 1 + p\sum_{x(p)}\left(\frac{4x^3 + 1}{p}\right) = p^2 + O(p^{3/2}),$$

where $h_{3,p}(2)$ is one if 2 is a cube mod p and zero otherwise. Note that this shows that Michel's bound for $A_{2,\mathcal{F}}(p)$ is sharp.

As $j(E_t)$ and $M(t)$ are non-constant, we expect the signs to be equidistributed.

The rational surfaces density theorem is applicable and we obtain orthogonal symmetry for the density of the non-family zeros.

6.2.2 *Rank* 6 *example.* We now give a more exotic example. See [Mil02] for the details. Let

$$\begin{aligned} A &= 8\,916\,100\,448\,256\,000\,000 \\ B &= -811\,365\,140\,824\,616\,222\,208 \\ C &= 26\,497\,490\,347\,321\,493\,520\,384 \\ D &= -343\,107\,594\,345\,448\,813\,363\,200 \\ a &= 16\,660\,111\,104 \\ b &= -1\,603\,174\,809\,600 \\ c &= 2\,149\,908\,480\,000. \end{aligned} \tag{6.8}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

The rational family $y^2 = x^3t^2 + 2g(x)t - h(x)$, $g(x) = x^3 + ax^2 + bx + c$ and $h(x) = (A-1)x^3 + Bx^2 + Cx + D$, has $A_{1,\mathcal{F}}(p) = -6p + O(1)$ for p large. Therefore, the family has rank 6 over $\mathbb{Q}(t)$. Writing in Weierstrass normal form yields

$$\begin{aligned} y^2 &= x^3 + (2at - B)x^2 + (2bt - C)(t^2 + 2t - A + 1)x + (2ct - D)(t^2 + 2t - A + 1)^2 \\ c_4(t) &= 2^{19}3^7 7^1 13^1(1475t^3 + \cdots - 7\,735\,999\,878\,503\,076\,170\,786\,750\,620\,939) \\ c_6(t) &= -2^{25}3^{11}(625t^5 + \cdots) \\ j(E_t) &= \frac{50\,141\,357\,421\,875t^9 + \cdots}{-1\,171\,875t^{10} + \cdots} \\ \Delta(t) &= -2^{44}3^{18}5^6(75t^{10} + \cdots). \end{aligned} \tag{6.9}$$

This is a rational surface, $j(E_t)$ and $M(t)$ are non-constant. Thus, by the rational surfaces density theorem, we verify the Katz–Sarnak predictions for a family of rank 6 over $\mathbb{Q}(t)$.

7. Summary and future work

Our main result is that, modulo standard conjectures, the fluctuations of the non-family low-lying zeros in one-parameter families of elliptic curves agree with the Katz–Sarnak conjectures. Further, a family of rank r over $\mathbb{Q}(t)$ has a density correction which equals the contribution of r zeros at the critical point, providing further evidence for the Birch and Swinnerton-Dyer conjecture.

We have found four families where the observed density agrees with the density of one (and only one) symmetry group. As expected, the first piece equals the contribution from r zeros at the critical point (where r is the geometric rank of the family), and the second equals SO(even) if all curves have even sign and SO(odd) if all curves have odd sign.

For these four families, we assumed only the GRH. We are able to unconditionally handle the dependence of the conductors on t, the signs of the functional equations and the error terms.

In general, the greatest difficulty is handling the variation in the conductors. Unlike the other families investigated [ILS00, Rub98], the conductors of elliptic curves vary wildly in a given family. If the discriminant $\Delta(t)$ has an irreducible factor of degree 4 or greater, either ABC or the square-free sieve conjecture must be assumed to perform the necessary sieving; if all irreducible factors are of degree at most 3, the sieving is unconditional.

The crucial observation is that, if we sieve to a positive percentage subset where the conductors are monotone, then we can bound the error terms. Note the extreme delicacy of our arguments: for conductors of size $\log N$, we cannot bound the error terms if the conductors range from $\log N - \log c$ to $\log N + \log c$ for some constant c.

It was observed in [Mil02] that in every family where $A_{2,\mathcal{F}}(p)$ can be directly calculated,

$$A_{2,\mathcal{F}}(p) = p^2 + h(p) - m_{\mathcal{F}}p + O(1), \tag{7.1}$$

where $h(p)$ is of size $p^{3/2}$ and averages to zero, and $m_{\mathcal{F}}$ is a positive constant, often different for different families.

We have shown all rational families (with the same distribution of signs) have equal one- and two-level densities. We can, however, try to expand the densities in powers of $1/\log N$. The different $m_{\mathcal{F}}p$ terms will lead to potential corrections to the densities of size $1/\log N$, giving the exciting possibility of distinguishing different families by lower order corrections to the common densities.

Unfortunately, the size of the errors in the one- and two-level densities are $O(\log\log N/\log N)$; thus, a significantly more delicate analysis is needed before we can expand the densities.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Appendix A. Sieving families of elliptic curves

Given a one-parameter family of elliptic curves E_t, we need to control the conductors $C(t)$ to determine the one- and two-level densities. Let the curves have discriminants $\Delta(t)$, and let $D(t)$ be the product of the irreducible polynomial factors of $\Delta(t)$.

$D(t)$ may always be divisible by a fixed square; let B be the largest square dividing $D(t)$ for all t. We prove in Theorem B.2 that for a rational elliptic surface, by passing to a subsequence $\tau = c_1 t + c_0$ for $D(\tau)/B$ square-free, $C(t)$ is given by a polynomial in τ. Call such t (or $D(t)$ or τ) good.

In order to evaluate the sums of $\prod_i a_t^{r_i}(p_i)$, it is necessary to restrict t to arithmetic progressions; however, restricting to t good ($D(\tau)/B$ square-free) does not yield t in arithmetic progressions.

We overcome this difficulty by doing a partial sieve with good bounds on over-counting. For notational convenience, we consider the case where $B = 1$ below, and indicate how to modify for general B.

Let $S(t)$ be some quantity associated to our family which we desire to sum over $\mathcal{T}_{\text{sqfree}}$, where

$$\begin{aligned} \mathcal{T}_{\text{sqfree}} &= \{t \in [N, 2N] : D(t) \text{ is square-free}\} \\ \mathcal{T}_N &= \{t \in [N, 2N] : d^2 \nmid D(t) \text{ for } 2 \leqslant d \leqslant \log^l N\}. \end{aligned} \tag{A.1}$$

Clearly $\mathcal{T}_{\text{sqfree}} \subset \mathcal{T}_N$. We show that $\mathcal{T}_N$ is a union of arithmetic progressions, and $|\mathcal{T}_N - \mathcal{T}_{\text{sqfree}}| = o(N)$.

The main obstruction is estimating the number of $t \in [N, 2N]$ with $D(t)$ divisible by the square of a prime $p \geqslant \log^l N$. If $k = \deg D(t)$,

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

For $k > 3$, the second piece is too difficult to estimate: there are too many d terms (d runs to $N^{k/2}$). If all the irreducible factors of $D(t)$ are of degree at most 3, the second piece is small. For factors of degree at most 2, this follows immediately, while for factors of degree 3 it follows from Hooley [Hoo76]. For larger degrees, we need the ABC conjecture (or one of its consequences, the square-free sieve conjecture).

A.1 Incongruent solutions of polynomials

Recall the following basic facts (see, for example, [Nag81]) for an integral polynomial $D(t)$ of degree k and discriminant δ.

i) Let p be a prime not dividing the coefficient of x^k. Then $D(t) \equiv 0 \bmod p$ has at most k incongruent solutions.

ii) Let $D(t) \equiv 0 \bmod p_i^{\alpha_i}$ have ν_i incongruent solutions. If the primes are distinct, there are $\prod_{i=1}^r \nu_i$ incongruent solutions of $D(t) \equiv 0 \bmod \prod_{i=1}^r p_i^{\alpha_i}$.

iii) Suppose that $p \nmid \delta$. Then the number of incongruent solutions of $D(t) \equiv 0 \bmod p$ equals the number of incongruent solutions of $D(t) \equiv 0 \bmod p^\alpha$.

Definition A.1. Let $\nu(d)$ be the number of incongruent solutions of $D(t) \equiv 0 \bmod d^2$.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

LEMMA A.2. *For d square-free, $\nu(d) \ll d^{\epsilon}$.*

The proof combines the above facts with the standard bound of the divisor function, $\tau(d) \ll d^{\epsilon}$.

A.2 Common prime divisors of polynomials

LEMMA A.3. *Let $f(t)$ and $g(t)$ be integer polynomials with no non-constant factors over $\mathbb{Z}[t]$. Then there exists c (independent of t) such that if p divides both $f(t)$ and $g(t)$, then $p|c$. In particular, $f(t)$ and $g(t)$ have no common large prime divisors.*

This can be proved using Euclid's algorithm.

A.3 Calculating $|\mathcal{T}_N|$

We have

$$\sum_{t\in\mathcal{T}_N} 1 = \sum_{d=1}^{\log^l N} \mu(d) \sum_{\substack{D(t)\equiv 0(d^2)\\ t\in[N,2N]}} 1. \tag{A.3}$$

There are $(N/d^2)\nu(d) + O(\nu(d))$ solutions to $D(t) \equiv 0 \bmod d^2$ for $t \in [N, 2N]$. By Lemma A.2, $\nu(d) \ll d^{\epsilon}$ for square-free d. Thus

$$|\mathcal{T}_N| = \sum_{d=1}^{\log^l N} \mu(d)\left[\frac{N}{d^2}\nu(d) + O(\nu(d))\right] = N\sum_{d=1}^{\log^l N} \frac{\mu(d)\nu(d)}{d^2} + O(\log^{l(1+\epsilon)} N). \tag{A.4}$$

As $\nu(d) \ll d^{\epsilon}$ for square-free d,

$$\left|\prod_{p<\log^l N}\left(1-\frac{\nu(p)}{p^2}\right) - \sum_{d=1}^{\log^l N}\frac{\mu(d)\nu(d)}{d^2}\right| \ll \sum_{d=\log^l N}^{\infty}\frac{d^{\epsilon}}{d^2} \ll \frac{1}{\log^{l(1-\epsilon)} N}. \tag{A.5}$$

Therefore,

$$|\mathcal{T}_N| = N\prod_{p<\log^l N}\left(1-\frac{\nu(p)}{p^2}\right) + O\left(\frac{N}{\log^{l(1-\epsilon)} N}\right) + O(\log^{l(1-\epsilon)} N). \tag{A.6}$$

We may take the product over all primes with negligible cost as

$$1 - \prod_{p\geqslant\log^l N}\left(1-\frac{\nu(p)}{p^2}\right) \ll \sum_{n\geqslant\log^l N}\frac{n^{\epsilon}}{n^2} \ll \frac{1}{\log^{l(1-\epsilon)} N}. \tag{A.7}$$

We have shown the following.

LEMMA A.4. *We have $\mathcal{T}_N = \{t \in [N, 2N] : d^2 \nmid D(t)$ for $2 \leqslant d \leqslant \log^l N\}$, and*

$$|\mathcal{T}_N| = N\prod_{p}\left(1-\frac{\nu(p)}{p^2}\right) + O\left(\frac{N}{\log^{l(1-\epsilon)} N}\right). \tag{A.8}$$

A.4 Estimating $\mathcal{T}_{\text{sqfree}}$

Assuming the ABC conjecture, Granville [Gra98, Theorem 1] proved that number of $t \in [N, 2N]$ such that $D(t)$ is square-free is

$$|\mathcal{T}_{\text{sqfree}}| = N\prod_{p}\left(1-\frac{\nu(p)}{p^2}\right) + o(N). \tag{A.9}$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Again, if the degree of $D(t)$ is at most 3, the ABC conjecture is not required. The family has a positive percentage of t giving $D(t)$ square-free (as we are assuming no square divides $D(t)$ for all t, no $\nu(p) = p^2$, hence the product can be bounded away from zero).

A.5 Evaluation of $|\mathcal{T}_N - \mathcal{T}_{\text{sqfree}}|$ and applications

By Equations (A.8) and (A.9), as $\mathcal{T}_{\text{sqfree}} \subset \mathcal{T}_N$, we have $|\mathcal{T}_N - \mathcal{T}_{\text{sqfree}}| = o(N)$.

We have proved

$$\sum_{\substack{t\in[N,2N]\\ D(t)\ \text{sqfree}}} S(t) = \sum_{t\in\mathcal{T}_N} S(t) + O\Bigg(\sum_{t\in\mathcal{T}} S(t)\Bigg)$$

$$= \sum_{d=1}^{\log^l N} \mu(d) \sum_{\substack{D(t)\equiv 0(d^2)\\ t\in[N,2N]}} S(t) + O\Bigg(\sum_{t\in\mathcal{T}} S(t)\Bigg). \tag{A.10}$$

We use arithmetic progressions to handle the piece with $d \leqslant \log^l N$, and Cauchy–Schwartz to handle $t \in \mathcal{T}$:

$$\sum_{t\in\mathcal{T}} S(t) \ll \Bigg(\sum_{t\in\mathcal{T}} S^2(t)\Bigg)^{1/2} \Bigg(\sum_{t\in\mathcal{T}} 1\Bigg)^{1/2} \ll \Bigg(\sum_{t\in[N,2N]} S^2(t)\Bigg)^{1/2} o(\sqrt{N}). \tag{A.11}$$

If we can show $\sum_{t=N}^{2N} S^2(t) = O(N)$, then the error term is negligible as $N \to \infty$.

A.6 Conditions implying $|\mathcal{F}| = c_{\mathcal{F}}N + o(N)$, $c_{\mathcal{F}} > 0$

Assume that no square divides $D(t)$ for all t. The number of $t \in [N, 2N]$ with $D(t)$ not divisible by d^2, $d \leqslant \log^l N$, is $N\prod_p(1 - \nu(p)/p^2) + o(N)$. Let $D(t) = \prod_i D_i^{r_i}(t)$, $D_i(t)$ irreducible. By multiple applications of Lemma A.3, there exists c such that for all t, there is no prime $p > c$ which divides two of the $D_i(t)$. Thus, if $D(t)$ is divisible by p^2 for a large prime, one of the factors is divisible by p^2. As there are finitely many factors, it is sufficient to bound by $o(N)$ the number of $t \in [N, 2N]$ with $p^2|D(t)$ for a large prime for irreducible $D(t)$.

Let $|\mathcal{F}|$ equal the number of $t \in [N, 2N]$ with $D(t)$ square-free. Let $c_{\mathcal{F}} = \prod_{p\leqslant\log^l N}(1-\nu(p)/p^2)$. We have seen that extending the product to all primes costs $O(1/\log^{l(1-\epsilon)}N)$. Thus, we need only bound $c_{\mathcal{F}}$ away from zero.

Let $D(t) = a_k t^k + \cdots + a_0$ with discriminant δ. For $p\nmid a_k\delta$, $\nu(p) \leqslant k$.

Let $\mathcal{P}$ be the set of primes dividing $a_k\delta$ and all primes at most $\sqrt{k}$. The contribution from $p \notin \mathcal{P}$ is bounded away from zero. Therefore, if $\nu(p) < p^2$ for $p|a_k\delta$ and $p \leqslant \sqrt{k}$, then $c_{\mathcal{F}} > 0$.

If $D(t)$ is divisible by a square for all t, the above arguments fail. Let P be the largest product of primes such that for all t, $P^2|D(t)$. By changing variables $\tau \to P^m t + t_0$, for m sufficiently large, $D(\tau)$ is divisible by fixed powers of $p|P$, depending only on $D(t_0)$. Thus, instead of sieving to $D(t)$ square-free, we sieve to $D(\tau)$ square-free except for primes dividing P.

Let δ_τ denote the new discriminant. As the discriminant is a product over the differences of the roots, t_0 does not change the discriminant, and P^m rescales by a power of P. Thus, $\delta_\tau = P^M\delta$. Further, the new leading coefficient is $P^{mk}a_k$. Thus, for $p\nmid P$, our previous arguments are still applicable, except that we are no longer sieving over $p|P$. We have shown the following.

THEOREM A.5 (Conditions on $D(t)$ implying $|\mathcal{F}| = c_{\mathcal{F}}N + o(N)$). *Assume no square divides $D(t)$ for all t. Let $\mathcal{P}$ be the set of primes dividing $a_k\delta$ and all primes at most $\sqrt{k}$. If for all $p \in \mathcal{P}$, $\nu(p) \leqslant p^2 - 1$, then $|\mathcal{F}| = c_{\mathcal{F}}N + o(N)$, $c_{\mathcal{F}} > 0$. If for all t, $B^2|D(t)$ (there exists $p \in \mathcal{P}$, $\nu(p) = p^2$),*

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

let P be the product of all primes either in $\mathcal{P}$ or dividing B. By changing variables to $\tau = P^m t + t_0$ for m large and sieving to $D(\tau)$ square-free except for $p|P$ (where for all t, the power of $p|P$ dividing $D(t)$ is constant), we again obtain $|\mathcal{F}| = c_{\mathcal{F}} N + o(N)$, $c_{\mathcal{F}} > 0$. In this case, $c_{\mathcal{F}}$ no longer includes factors from $p|P$.

If all irreducible factors of $D(t)$ have degree at most 3, these results are unconditional; if there is an irreducible factor with degree at least 4, these results are conditional and a consequence of the ABC or square-free sieve conjecture.

Further, let $\mathcal{T} = \{t \in [N, 2N] : \exists d > \log^l N \text{ with } d^2 | D(t)\}$. Then $\mathcal{T} = o(N)$.

Appendix B. Handling the conductors $C(t)$

For many families of elliptic curves, by sieving to a positive percentage subsequence of t we obtain a sub-family where the conductors are a monotone polynomial in t. In particular, we prove this for all rational surfaces.

Tate's algorithm (see [Cre92, pp. 49–52]) allows us to calculate the conductor $C(t)$ for an elliptic curve E_t over $\mathbb{Q}$:

$$C(t) = \prod_{p|\Delta(t)} p^{f_p(t)}, \tag{B.1}$$

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 (see [Sil86]).

Let $\Delta(t) = d\Delta_1(t)\Delta_2(t)$, where $(\Delta_2(t), c_4(t)) = 1$ and $\Delta_1(t)$ is the product of powers of irreducible polynomials dividing $\Delta(t)$ and $c_4(t)$. By possibly changing d, we may take $\Delta_i(t)$ primitive. Let $D_i(t)$ be the product of all irreducible polynomials dividing $\Delta_i(t)$, $D(t) = D_1(t)D_2(t)$.

For t with $D(t)$ square-free except for small primes, $C(t) = D_1^2(t)D_2(t)$ if $\Delta(t)$ has no irreducible polynomial factor occurring at least 12 times (except for corrections from the small primes). Hence, while $f_p(t)$ may vary, the product of $p^{f_p(t)}$, except for a finite set of primes, is well behaved.

Let

$$\mathcal{P}_0 = \{p : p \leqslant \deg \Delta(t)\} \cup \{p : p|cd\}, \quad P_0 = \prod_{p \in \mathcal{P}_0} p. \tag{B.2}$$

The idea is that while for such p, $f_p(t)$ may vary, by changing variables from t to $P_0^m t + t_1$ for some enormous m, for $p \in \mathcal{P}_0$, $f_p(P_0^m t + t_1) = f_p(t_1)$. Thus, for this subsequence and these primes, $f_p(t)$ is constant.

We need two preliminary results. First, given a finite set of primes $\mathcal{P}_0$, we may find an m and a t_1 such that for those primes, $f_p(P_0^m t + t_1)$ is constant. Second, applying Lemma A.3, given two polynomials with no non-constant factors over $\mathbb{Q}$, there is a finite set of primes $\mathcal{P}_2$ such that if there exists t such that there exists p dividing both polynomials, then $p \in \mathcal{P}_2$.

B.1 $f_p(t)$, $p \in \mathcal{P}_0$

Consider the original family of elliptic curves

$$E_t : y^2 + a_1(t)xy + a_3(t)y = x^3 + a_2(t)x^2 + a_4(t)x + a_6(t). \tag{B.3}$$

Assume $\Delta(t)$ is not identically zero. Choose t_1 such that for all $t \geqslant t_1$, $\Delta(t) \neq 0$. Apply Tate's algorithm to E_{t_1}. If the initial equation was non-minimal for p, we change coordinates by $T(0, 0, 0, p)$ (see [Cre92]) and restart. After finitely many passes, Tate's algorithm terminates.

In determining $f_p(t_1)$, assume that we passed through Tate's algorithm $L_{t_1}(p)$ times. For each prime p, after possibly many coordinate changes, one of the following conditions held: $p \nmid \Delta$, $p \nmid c_4$, $p^2 \nmid a_6$, $p^3 \nmid b_8$, $p^3 \nmid b_6$, $p \nmid w(a_2, a_4, a_6)$, $p \nmid xa_3^2(a_3) + 4xa_6(a_6)$, $p \nmid xa_4^2(a_4) - 4xa_2(a_2)xa_6(a_6)$, $p^4 \nmid a_4$,

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

$p^6 \nmid a_6$; and every function is polynomial in a_i. Thus, after possibly many coordinate changes, some polynomial (with integer coefficients) of a_i is not divisible by either p, p^2, p^3, p^4 or p^6.

Consider $\tau = P_0^m t + t_1$. For enormous m, $f_p(\tau) = f_p(t_1)$ for $p \in \mathcal{P}_0$ because in Tate's algorithm, we only need the values modulo a power of p. We have

$$a_i(\tau) = a_i(P_0^m t + t_1) = P_0^m t \widehat{a}_i(P_0^m t) + a_i(t_1) = \widetilde{a}_i(t) + a_i(t_1). \tag{B.4}$$

If m is sufficiently large, we can ignore $\widetilde{a}_i(t)$ in all equivalence checks, as for these powers of p, $\widetilde{a}_i(t) \equiv 0$. Let

$$\begin{aligned} n_t(p) &= \mathrm{ord}(p, \Delta(t)) \\ n &= \max_{p \in \mathcal{P}_0} n_{t_1}(p) \\ L &= \max_{p \in \mathcal{P}_0} L_{t_1}(p). \end{aligned} \tag{B.5}$$

We prove $f_p(\tau) = f_p(t_1)$ for large m. How large must m be? Excluding lines 42–65, on each pass through Tate's algorithm we sometimes divide our coefficients by powers of p: up to p^2 on lines 26 and 30, up to p^3 on line 34, up to p^4 on line 69, and p^{12} on line 80. Over-estimating, we divide by at most $p^{2\cdot 2+1\cdot 3+1\cdot 4+1\cdot 12} = p^{23}$.

For lines 42–65, we have a loop which can be executed at most $n+4$ times. We constantly divide by increasing powers of p; the largest power is the last time through the loop, which is at most $p^{2(n+6)}$. As we pass through this loop at most $n+4$ times, we divide by at most $p^{2n^2+20n+48}$.

Thus, on each pass we have divisions by at most $p^{2n^2+20n+48+23}$. As we loop through the main part of Tate's algorithm at most L times, we have divisions by at most $p^{(2n^2+20n+71)L}$. If $m > (2n^2 + 20n + 71)L$, then for all t, none of the $\widetilde{a}_i(t) = P_0^m t \widehat{a}_i(t)$ terms affect any congruence. Significantly smaller choices of m work: many of the divisions (for example, from lines 42–65) arise only once.

B.2 Rational surfaces I

B.2.1 *Preliminaries.* Recall that an elliptic surface $y^2 = x^3 + A(t)x + B(t)$ is rational if and only if one of the following cases is true:

1) $0 < \max\{3 \deg A(t), 2 \deg B(t)\} < 12$;
2) $3 \deg A(t) = 2 \deg B(t) = 12$ and $\mathrm{ord}_{t=0}\, t^{12} \Delta(t^{-1}) = 0$.

See [RS98, pp. 46–47] for more details.

Assume that we are in case 1. No non-constant polynomial of degree 11 or more divides $\Delta(t)$; however, a 12th or higher power of a prime might divide $\Delta(t)$. Let $k = \deg \Delta(t)$, and write

$$\begin{aligned} \Delta(t) &= d\Delta_1(t)\Delta_2(t) \\ c_4(t) &= c\gamma_1(t)\gamma_2(t) \\ \mathcal{P}_0 &= \{p : p \leqslant \deg \Delta(t)\} \cup \{p : p | cd\}, \quad P_0 = \prod_{p \in \mathcal{P}_0} p \end{aligned} \tag{B.6}$$

where $\Delta_1(t)$ through $\gamma_2(t)$ are primitive polynomials, $\Delta_1(t)$ and $\gamma_1(t)$ are divisible by the same non-constant irreducible polynomials, and $\Delta_2(t)$ and $c_4(t)$ are not both divisible by any non-constant polynomial.

Let $D_i(t)$ be the product of all non-constant irreducible polynomials dividing $\Delta_i(t)$, and similarly for $c_i(t)$. Let $D(t) = D_1(t)D_2(t) = \alpha_\kappa t^\kappa + \cdots + \alpha_0$ $(\kappa \leqslant k)$, $c(t) = c_1(t)c_2(t)$.

Apply Lemma A.3 to $c(t)$ and $D_2(t)$. Thus there exists c' such that if there exists t where p divides both polynomials, then $p|c'$. Let $\mathcal{P}_2$ be the prime divisors of c' not in $\mathcal{P}_0$ and let $\mathcal{P}_1$ be the

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

prime divisors of $\alpha_\kappa \cdot \text{Discriminant}(D(t))$ not in $\mathcal{P}_0$. Define

$$\mathcal{P} = \bigcup_{i=1}^{2} \mathcal{P}_i, \quad P = \prod_{p \in \mathcal{P}} p. \tag{B.7}$$

Note that every prime in $\mathcal{P}$ is greater than k and not in $\mathcal{P}_0$.

As the product of primitive polynomials is primitive, $D(t)$ is primitive. For any prime, either $D(t)$ mod p is a constant not divisible by p or a non-constant polynomial of degree at most k. In the second case, as there are at most k roots to $D(t) \equiv 0 \bmod p$, we find that given a $p > k$, there exists t_p such that $D(t_p) \not\equiv 0 \bmod p$. By the Chinese remainder theorem, there exists $t_0 \equiv t_p \bmod p$ for all $p \in \mathcal{P}$.

B.2.2 *Calculating the conductor.* For all $p \in \mathcal{P}$, $D(Pt + t_0) \equiv D(t_0) \not\equiv 0 \bmod p$. As $\mathcal{P}$ and $\mathcal{P}_0$ are disjoint, this implies that $D(Pt + t_0)$ is minimal for all $p \in \mathcal{P}$, as $\mathcal{P}_0$ contains the factors of d, 2 and 3. Moreover, $f_p(Pt + t_0) = 0$ for $p \in \mathcal{P}$.

By changing variables again, from t to $P_0^m t + t_1$, we can determine the powers of $p \in \mathcal{P}_0$ in the conductor. Combining the two changes, we send t to $\tau = P(P_0^m t + t_1) + t_0$.

Originally we had $\Delta(t) = d\Delta_1(t)\Delta_2(t)$. Now we have $\Delta(\tau) = d\Delta_1(\tau)\Delta_2(\tau)$. It is possible that $D_1(\tau)D_2(\tau)$ is no longer primitive; however, if there is a common prime divisor p, p divides $\alpha_\kappa(P \cdot P_0^m)^\kappa$, implying $p \in \mathcal{P}_0 \sqcup \mathcal{P}$.

We sieve to $D(\tau)$ square-free for $p \notin \mathcal{P}_0 \sqcup \mathcal{P}$. As $\mathcal{P}_0 \sqcup \mathcal{P}$ contains all primes less than k, as well as the prime divisors of P_0, P, α_κ and $\text{Discriminant}(\Delta(t))$, we can perform the sieving. Note that the discriminants of $\Delta(t)$ and $\Delta(\tau)$ differ by a power of $P \cdot P_0^m$. Thus, away from these primes, $D(\tau) \equiv 0 \bmod p^2$ has at most $k < p^2$ roots, and we may sieve to a positive percentage of t. The sieving is unconditional if each irreducible factor of $D(\tau)$ is of degree at most 3.

Now, $D(\tau)$ is divisible by fixed powers of primes in $\mathcal{P}_0$ and never divisible by primes in $\mathcal{P}$. Thus there exists c_1, c_2 with factors in $\mathcal{P}_0$ such that $\widetilde{D}(\tau) = (D_1(\tau)/c_1)(D_2(\tau)/c_2)$ is not divisible by any $p \in \mathcal{P}_0 \sqcup \mathcal{P}$. We sieve to $\widetilde{D}(\tau)$ square-free; for $p \notin \mathcal{P}_0 \sqcup \mathcal{P}$, this is the same as $D(\tau)$ not divisible by p^2.

We need to determine $f_p(\tau)$ for $p \in \mathcal{P}_0$, $p \in \mathcal{P}$, and $p \notin \mathcal{P}_0 \sqcup \mathcal{P}$.

By our previous arguments, if m is sufficiently large, $f_p(\tau) = f_p(Pt_1 + t_0)$ for $p \in \mathcal{P}_0$.

If $p \in \mathcal{P}$ then $p \notin \mathcal{P}_0$. Modulo p, $\Delta(\tau) = \Delta(P(P_0 t + t_1) + t_0) \equiv \Delta(t_0) \not\equiv 0$. Thus, for these p, $f_p(\tau) = 0$.

Assume $p \notin \mathcal{P}_0 \sqcup \mathcal{P}$. The leading term of $dD(\tau)$ is $d\alpha_\kappa(P \cdot P_0^m)^\kappa$. By construction, p does not divide the leading coefficient of $\Delta(\tau)$, as $\mathcal{P}_0 \sqcup \mathcal{P}$ contains the prime divisors of d, α_k, P and P_0. If we sieve to $\widetilde{D}(\tau)$ square-free for $p \notin \mathcal{P}_0 \sqcup \mathcal{P}$, then as the degree of $\Delta(\tau)$ is at most 10, the curve is minimal for such p. Thus, $f_p(\tau)$ is 1 if $p | D_2(\tau)$ and 2 if $p | D_1(\tau)$.

Thus, we have shown the following.

THEOREM B.1. *With all quantities as above, for $\widetilde{D}(\tau)$ square-free, the conductors are*

$$C(\tau) = \prod_{p \in \mathcal{P}_0} p^{f_p} \cdot \left(\frac{|D_1(\tau)|}{c_1}\right)^2 \frac{|D_2(\tau)|}{c_2}. \tag{B.8}$$

For sufficiently large τ, $C(\tau)$ is a monotone increasing polynomial (we may drop the absolute values), and a positive percentage of τ yield $\widetilde{D}(\tau)$ square-free.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

B.3 Rational surfaces II

We consider what could go wrong in our proof if we are in case 2, where $3\deg A(t) = 2\deg B(t) = 12$ and $\mathrm{ord}_{t=0}\, t^{12}\Delta(t^{-1}) = 0$.

Thus, $\Delta(t)$ is a degree 12 polynomial, and we need to worry about minimality issues. As before, we have

$$\begin{aligned} \Delta(t) &= -2^4(2^2A^3(t) + 3^3B^2(t)) = d\Delta_1(t)\Delta_2(t) \\ c_4(t) &= c\gamma_1(t)\gamma_2(t) \\ \mathcal{P}_0 &= \{p : p \leqslant \deg\Delta(t)\} \cup \{p : p|cd\}, \quad P_0 = \prod_{p\in\mathcal{P}_0} p. \end{aligned} \tag{B.9}$$

There are three cases:

- $\Delta(t)$ not divisible by a 12th power;
- $(\alpha t+\beta)^{12}|\Delta(t)$, $(\alpha t+\beta)\nmid c_4(t)$;
- $(\alpha t+\beta)^{12}|\Delta(t)$, $(\alpha t+\beta)|c_4(t)$.

These cases are handled in a similar fashion as before; see [Mil02] for the calculations.

B.4 Generalizations

The previous arguments are applicable to any family where $\deg\Delta(t) \leqslant 12$ (which can include some non-rational families). It is straightforward to generalize these arguments for all families.

B.5 Summary

We summarize our sieving and conductor results as follows.

Theorem B.2 (Conductors and cardinalities for families). *For a one-parameter family with* $\deg\Delta(t) \leqslant 12$, *which includes all rational families, by sieving to a positive percentage subsequence we obtain a family with conductors given by a monotone polynomial. Further, by Theorem A.5, after changing variables to* $\tau = P^m t + t_0$, *a positive percentage of* $t \in [N, N]$ *give* $D(\tau)$ *square-free except for primes* $p|P$, *where the power of such* p *dividing* $D(\tau)$ *is independent of* t. *If all the irreducible factors of* $\Delta(t)$ *are degree* 3 *or less, the sieving is unconditional; for degree* 4 *and higher, the sieving is a consequence of the ABC or square-free sieve conjecture.*

Appendix C. Sums of test functions at primes

We list several standard sums of test functions over primes. $\widehat{F}$, $\widehat{f_i}$ are even Schwartz functions with compact support, $\varphi(m)$ is the Euler phi-function.

All statements below are straightforward applications of partial summation and the Riemann hypothesis (or the GRH for Dirichlet L-functions if $m \neq 1$) to handle the prime sums (see, for example, [Mil02]); weaker error terms are obtainable by the prime number theorem.

Lemma C.1 (Sum of $\widehat{F}$ over primes). *We have*

$$\frac{1}{\log N}\sum_{p\equiv b(m)} \frac{\log p}{p}\widehat{F}\left(a\frac{\log p}{\log N}\right) = \frac{1}{2a\varphi(m)}F(0) + O\left(\frac{1}{\log N}\right). \tag{C.1}$$

Setting $m = 1$ and $a = 1, 2$ yields the following.

Corollary C.2. *We have*

$$\frac{1}{\log N}\sum_{p} \frac{\log p}{p}\widehat{F}\left(\frac{\log p}{\log N}\right) = \frac{1}{2}F(0) + O\left(\frac{1}{\log N}\right).$$

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Corollary C.3. *We have*

$$\frac{1}{\log N}\sum_p \frac{\log p}{p}\widehat{F}\left(2\frac{\log p}{\log N}\right) = \frac{1}{4}F(0) + O\left(\frac{1}{\log N}\right).$$

Lemma C.4. *We have*

$$4\sum_p \frac{\log^2 p}{\log^2 M}\frac{1}{p}\widehat{f}_1\widehat{f}_2\left(\frac{\log p}{\log M}\right) = 2\int_{-\infty}^{\infty}|u|\widehat{f}_1(u)\widehat{f}_2(u)\,du + O\left(\frac{1}{\log M}\right). \tag{C.2}$$

For $p \equiv b(m)$ we have the following.

Lemma C.5. *We have*

$$4\sum_{p\equiv b(m)} \frac{\log^2 p}{\log^2 M}\frac{1}{p}\widehat{f}_1\widehat{f}_2\left(\frac{\log p}{\log M}\right) = \frac{2}{\varphi(m)}\int_{-\infty}^{\infty}|u|\widehat{f}_1(u)\widehat{f}_2(u)\,du + O\left(\frac{1}{\log M}\right). \tag{C.3}$$

Lemma C.6. *Let $\mathcal{E}$ have rank r over $\mathbb{Q}(t)$ and assume Tate's conjecture for $\mathcal{E}$ (known if $\mathcal{E}$ is a rational surface). Then*

$$2\sum_p \frac{\log p}{\log X}\frac{1}{p}\widehat{f}\left(\frac{\log p}{\log X}\right)\frac{-A_{1,\mathcal{F}}(p)}{p} = rf(0) + o(1). \tag{C.4}$$

Finally, we constantly encounter sums such as

$$\sum_p \frac{\log p}{\log C(t)}\frac{1}{p^r}\widehat{f}\left(r\frac{\log p}{\log C(t)}\right)a_t^r(p), \tag{C.5}$$

where $r \in \{1,2\}$ and $\log C(t)$ is $k\log N + o(\log N)$.

By Hasse, $a_t^r(p) \leqslant (2\sqrt{p})^r$. The contribution S_l from $p \leqslant \log^l N$ is

$$S_l \ll \frac{1}{\log N}\sum_{p\leqslant \log^l N}\frac{\log p}{p^{r/2}}. \tag{C.6}$$

Clearly the larger contribution is from $r = 1$. By the prime number theorem, $\sum_{p\leqslant x}\log p \ll x$. By partial summation, $\sum_{p\leqslant x}(\log p/\sqrt{p}) \ll \sqrt{x}$. Thus

$$S_l \ll \frac{\sqrt{\log^l N}}{\log N}. \tag{C.7}$$

We have shown the following.

Lemma C.7 (Removing small primes). *The sums over primes $p \leqslant \log^l N$ in the explicit formula contribute $O(\log^{(l/2)-1} N)$. For $l < 2$, this is negligible.*

Appendix D. Handling the error terms in the two-level density

Following Rudnick and Sarnak [RS96] and Rubinstein [Rub98], we handle the error terms in the two-level density, assuming that we are able to prove the one-level density theorem with error terms. By the explicit formula (Equation (2.3)).

$$\sum_{j_i} F_i\left(\frac{\log N_E}{2\pi}\gamma_E^{(j_i)}\right) = \text{Good}_i + O((\log N_E)^{-1/2}), \tag{D.1}$$

where Good_i is the good part of the explicit formula, involving $\widehat{F}(0)$, $F(0)$, and sums of $a_E(p)$ and $a_E^2(p)$ for primes $p > \log N$.

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press

Multiplying and summing over i yields

$$\frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\prod_{i=1}^{2}\left[\sum_{j_i}F_i\left(\frac{\log N_E}{2\pi}\gamma_E^{(j_i)}\right)+O((\log N_E)^{-\frac{1}{2}})\right]=\frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}\prod_{i=1}^{2}\text{Good}_i. \tag{D.2}$$

Multiplying out the left-hand side yields terms such as

$$O\left[\frac{1}{|\mathcal{F}|}\sum_{E\in\mathcal{F}}(\log N_E)^{-(2-k)/2}\prod_{m=1}^{k}\sum_{j_{m_i}}F_i\left(\frac{\log N_E}{2\pi}\gamma_E^{(j_{m_i})}\right)\right]. \tag{D.3}$$

If each function F_i were positive, we could insert absolute values and move $(1/|\mathcal{F}|)\sum_{E\in\mathcal{F}}$ past the $\log^{-(2-k)/2}N_E$ factor. We assume that our family has been sieved, so that the conductors satisfy $\log N_E = c\log N + o(\log N)$.

There are three terms. If $k=0$ there is clearly no net contribution. For $k=1$ we have a one-level density, which is finite by assumption. No error hits the $k=2$ piece (this is the piece we want to calculate). Only the $k=1$ piece is troublesome for F_i not positive.

If F_i is not positive, we increase the above by replacing F_i with a positive function g_i such that g_i is an even Schwartz function whose Fourier transform is supported in the same interval as that of F_i and $g_i(x)\geqslant|F_i(x)|$. As the g_i satisfy the necessary conditions, we may apply the one-level density theorem to the g_i, obtaining a bounded quantity. Hitting this with $(\log N_E)^{-1/2}$, we see that there is a negligible contribution.

For a construction of g_i, see Rubinstein [Rub98, pp. 40–41] or Rudnick and Sarnak [RS96, pp. 302–304].

We have shown the following.

Theorem D.1 (Handling the error terms). *If we are able to do the one-level density calculations, then we may ignore the error terms in the two-level density.*

Note that the error need not be $O(\log^{-1/2}N)$; $o(1)$ also works.

Acknowledgements

The author wishes to thank Harald Helfgott, Henryk Iwaniec, Nick Katz, Wenzhi Luo and Peter Sarnak for many enlightening conversations.

References

BEW98 B. Berndt, R. Evans and K. Williams, *Gauss and Jacobi sums*, Canadian Mathematical Society Series of Monographs and Advanced Texts, vol. 21 (Wiley-Interscience, New York, 1998).

BS66 B. Birch and N. Stephens, *The parity of the rank of the Mordell–Weil group*, Topology **5** (1966), 295–299.

BS63 B. Birch and H. Swinnerton-Dyer, *Notes on elliptic curves. I*, J. reine angew. Math. **212** (1963), 7–25.

BS65 B. Birch and H. Swinnerton-Dyer, *Notes on elliptic curves. II*, J. reine angew. Math. **218** (1965), 79–108.

BCDT01 C. Breuil, B. Conrad, F. Diamond and R. Taylor, *On the modularity of elliptic curves over **Q**: wild 3-adic exercises*, J. Amer. Math. Soc. **14** (2001), 843–939.

Bru92 A. Brumer, *The average rank of elliptic curves I*, Invent. Math. **109** (1992), 445–472.

BH A. Brumer and R. Heath-Brown, *The average rank of elliptic curves V*, Preprint.

CFKRS02 J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein and N. C. Snaith, *Integral Moments of L-Functions*, Preprint (2002), math.NT/0206018.

Cre92 J. E. Cremona, *Algorithms for modular elliptic curves* (Cambridge University Press, 1992).

Dav80 H. Davenport, *Multiplicative number theory*, second edition, Graduate Texts in Mathematics, vol. 74 (Springer, New York, 1980), revised by H. Montgomery.

Del80 P. Deligne, *La conjecture de Weil. II*, Publ. Math. Inst. Hautes Études Sci. **52** (1980), 137–252.

Hej94 D. Hejhal, *On the triple correlation of zeros of the zeta function*, Internat. Math. Res. Notices (1994), 294–302.

Gra98 Granville, *ABC allows us to count squarefrees*, Internat. Math. Res. Notices **19** (1998), 991–1009.

HW95 G. Hardy and E. Wright, *An introduction to the theory of numbers*, fifth edition, Oxford Science Publications (Clarendon, Oxford, 1995).

Hel H. Helfgott, *On the distribution of root numbers in families of elliptic curves*, Preprint.

Hoo76 C. Hooley, *Applications of sieve methods to the theory of numbers* (Cambridge University Press, Cambridge, 1976).

ILS00 H. Iwaniec, W. Luo and P. Sarnak, *Low lying zeros of families of L-functions*, Publ. Math. Inst. Hautes Études Sci. **91** (2000), 55–131.

KS99a N. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues and monodromy*, Amer. Math. Soc. Colloq. Publ., vol. 45 (American Mathematical Society, Providence, RI, 1999).

KS99b N. Katz and P. Sarnak, *Zeros of zeta functions and symmetries*, Bull. Amer. Math. Soc. **36** (1999), 1–26.

Kna92 A. Knapp, *Elliptic curves* (Princeton University Press, Princeton, 1992).

Mes86 J. Mestre, *Formules explicites et minorations de conducteurs de variétés algébriques*, Compositio Math. **58** (1986), 209–232.

Mic95 P. Michel, *Rang moyen de familles de courbes elliptiques et lois de Sato–Tate*, Monatsh. Math. **120** (1995), 127–136.

Mil02 S. J. Miller, *One- and two-level densities for families of elliptic curves: evidence for the underlying group symmetries*, PhD thesis, Princeton University (2002), http://www.math.princeton.edu/∼sjmiller/thesis/thesis.pdf.

Mon73 H. Montgomery, *The pair correlation of zeros of the zeta function*, in *Analytic number theory*, *Proc. Symp. Pure Math.*, vol. 24 (American Mathematical Society, Providence, RI, 1973), 181–193.

Nag97 K. Nagao, $\mathbb{Q}(t)$*-rank of elliptic curves and certain limit coming from the local points*, Manuscripta Math. **92** (1997), 13–32.

Nag81 T. Nagell, *Introduction to number theory* (Chelsea, New York, 1981).

Riz O. Rizzo, *Average root numbers for a non-constant family of elliptic curves*, Preprint.

RS98 M. Rosen and J. Silverman, *On the rank of an elliptic surface*, Invent. Math. **133** (1998), 43–67.

Rub98 M. Rubinstein, *Evidence for a spectral interpretation of the zeros of L-functions*, PhD Thesis, Princeton University (1998), http://www.ma.utexas.edu/users/miker/thesis/thesis.html.

RS96 Z. Rudnick and P. Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Math. J. **81** (1996), 269–322.

Sil86 J. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106 (Springer, Berlin, 1986).

Sil94 J. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 151 (Springer, Berlin, 1994).

Sil98 J. Silverman, *The average rank of an algebraic family of elliptic curves*, J. reine angew. Math. **504** (1998), 227–236.

Tat65 J. T. Tate, *Algebraic cycles and poles of zeta functions*, in *Arithmetical Algebraic Geometry* (Harper and Row, New York, 1965), 93–110.

TW95 R. Taylor and A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. (2) **141** (1995), 553–572.

Was87 L. Washington, *Class numbers of the simplest cubic fields*, Math. Comp. **48** (1987), 371–384.

Wil95 A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2) **141** (1995), 443–551.

Steven J. Miller sjmiller@math.ohio-state.edu
Department of Mathematics, Ohio State University, Columbus, OH 43210, USA

https://doi.org/10.1112/S0010437X04000582 Published online by Cambridge University Press