Journal of Number Theory 130 (2010) 2341–2358

Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

Towards an 'average' version of the Birch and Swinnerton-Dyer conjecture[☆]

John Goes[a], Steven J. Miller[b,*]

[a] Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, IL 60680, United States
[b] Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267, United States

ARTICLE INFO

Article history:
Received 21 November 2009
Revised 23 April 2010
Communicated by David Goss

MSC:
primary 11M41
secondary 11G40, 15B52

Keywords:
Elliptic curves
Birch and Swinnerton-Dyer conjecture
Average rank
One-level density
Katz–Sarnak conjectures
Random matrix theory

ABSTRACT

Text. The Birch and Swinnerton-Dyer conjecture states that the rank of the Mordell–Weil group of an elliptic curve E equals the order of vanishing at the central point of the associated L-function $L(s,E)$. Previous investigations have focused on bounding how far we must go above the central point to be assured of finding a zero, bounding the rank of a fixed curve or on bounding the average rank in a family. Mestre (1986) [Mes] showed the first zero occurs by $O(1/\log\log N_E)$, where N_E is the conductor of E, though we expect the correct scale to study the zeros near the central point is the significantly smaller $1/\log N_E$. We significantly improve on Mestre's result by averaging over a one-parameter family of elliptic curves $\mathcal{E}$ over $\mathbb{Q}(T)$. We assume GRH, Tate's conjecture if $\mathcal{E}$ is not a rational surface, and either the ABC or the Square-Free Sieve Conjecture if the discriminant has an irreducible polynomial factor of degree at least 4. We find non-trivial upper and lower bounds for the average number of normalized zeros in intervals on the order of $1/\log N_E$ (which is the expected scale). Our results may be interpreted as providing further evidence in support of the Birch and Swinnerton-Dyer conjecture, as well as the Katz–Sarnak density conjecture from random matrix theory (as the number of zeros predicted by random matrix theory lies between our upper and lower bounds). These methods may be applied to additional families of L-functions.

[☆] This work was done at the 2009 SMALL Undergraduate Research Project at Williams College, funded by NSF Grant DMS0850577 and Williams College; it is a pleasure to thank them for their support. We would also like to thank Michael Greenblatt, Mark Watkins, the organizers and participants of the 2009 Young Mathematicians Conference at Ohio State and the referee for helpful discussions on an earlier draft. The second named author was also partly supported by NSF Grant DMS0855257.

[*] Corresponding author.
E-mail addresses: johnwgoes@gmail.com (J. Goes), Steven.J.Miller@williams.edu (S.J. Miller).

doi:10.1016/j.jnt.2010.04.002

Video. For a video summary of this paper, please click here or visit http://www.youtube.com/watch?v=3EVYPNi_LG0.

1. Introduction

The goal of this paper is to provide evidence towards the Birch and Swinnerton-Dyer conjecture in one-parameter families of elliptic curves. We briefly summarize our results, assuming the reader is familiar with the notation and subject. Afterwards we review the needed background material from elliptic curves and previous results in Section 2; for the convenience of the reader, we state all the conjectures assumed or discussed at various points in Appendix A. We then prove our theorems and discuss generalizations to other families of L-functions in Section 3, where we give explicit non-trivial upper and lower bounds.

The Birch and Swinnerton Dyer conjecture asserts that if E is an elliptic curve whose Mordell–Weil group $E(\mathbb{Q})$ has geometric rank r, then the associated completed L-function $\Lambda(s,E)$ has analytic rank r (i.e., it vanishes to order r at the central point). This is an exceptionally hard problem to investigate, theoretically and numerically. While there is some theoretical evidence when the rank is at most 1, the general case is intractable both theoretically and experimentally. For example, although we can construct elliptic curves with geometric rank exceeding 20, the largest known lower bound for the analytic rank of a $\Lambda(s,E)$ is only 3.[1]

We consider the following natural question. Let E be an elliptic curve with geometric rank r, and assume the Generalized Riemann Hypothesis (GRH). The Birch and Swinnerton-Dyer conjecture predicts that there should be r zeros at the central point. *How far must we go along the critical line before we are assured of seeing r zeros?*

If N_E denotes the conductor of the elliptic curve, we expect the correct scale for zeros near the central point to be of size $1/\log N_E$. Miller [Mil3] investigated the first few zeros above the central point for the family of all elliptic curves as well as one-parameter families of small rank over $\mathbb{Q}(T)$. His results are consistent with the low zeros being of height on the order of $1/\log N_E$; however, the first few zeros are higher than the $N_E \to \infty$ scaling limits predicted by the independent model of random matrix theory. The data suggests that, for finite conductors, better agreement is obtained by modeling these zeros with the interaction model (which involves Jacobi ensembles). Determining the correct corresponding random matrix ensemble involves understanding the discretization of the central values of L-functions and the lower order terms in the one-level density. In his thesis Duc Khiem Huynh [Huy] successfully modeled the first zero of the family of quadratic twists of a fixed elliptic curve, and current work by the second named author and Eduardo Dueñez, Duc Khiem Huynh, Jon Keating and Nina Snaith is investigating the case of a general one-parameter family [DHKMS1, DHKMS2].

The best theoretical result on the first zero above the central point is due to Mestre. Assuming the Generalized Riemann Hypothesis, Mestre [Mes] bounded the analytic rank of E by $O(\log N_E/\log\log N_E)$ and showed its first zero above the central point is at most $B/\log\log N_E$. While this is significantly larger than what we expect the truth to be, namely $O(1/\log N_E)$, it has the advantage of holding for all elliptic curves.

In this note we show that we may reduce the window on the critical line to something of the expected order if we average over a one-parameter family of elliptic curves. Specifically, consider a one-parameter family $\mathcal{E}: y^2 = x^3 + A(T)x + B(T)$ of geometric rank r over $\mathbb{Q}(T)$, with $A(T), B(T) \in \mathbb{Z}[T]$.

[1] The number of terms needed for the computation is on the order of the square-root of the conductor of E, which grows rapidly in families. While it is possible to numerically show that the first r Taylor coefficients of $\Lambda(s,E)$ are close to zero for many E's with geometric rank r, in general these computations can only provide evidence. The exception is when we have formulas for the derivatives as a known quantity times a rational, in which case we can convert these calculations to proofs of vanishing. See http://web.math.hr/~duje/tors/rk28.html for an example by N. Elkies of an elliptic curve with geometric rank at least 28.

For each $t \in \mathbb{Z}$ we may specialize and obtain an elliptic curve $E_t : y^2 = x^3 + A(t)x + B(t)$ with conductor $N_t := N_{E_t}$. By Silverman's specialization theorem [Sil2], for all t sufficiently large each elliptic curve E_t has geometric rank at least r. Assuming standard conjectures, Helfgott [He] proved that for a generic family the sign of the functional equation is 1 half the time and -1 the other half. It is believed that a generic curve in a generic family has analytic rank as small as possible consistent with all constraints. In our case, as the rank must be at least r if the Birch and Swinnerton-Dyer conjecture is true, we expect that in the limit half the curves will have analytic rank r and the other half $r+1$, for an average rank of $r + \frac{1}{2}$.

We take our family to be $\mathcal{F}_R := \{\Lambda(s, E_t): R \leqslant t \leqslant 2R\}$ with $R \to \infty$, though we often abuse notation and use $\mathcal{F}_R$ to denote t in $[R, 2R]$. There are two ways to normalize the zeros of $\Lambda(s, E_t)$ near the central point: (1) globally, using $\frac{\log N}{2\pi} := \frac{1}{R}\sum_{t\in\mathcal{F}_R}\frac{\log N_t}{2\pi}$; (2) locally, using $\frac{\log N_t}{2\pi}$. It is significantly easier to use the global rescaling; however, as each elliptic curve can be considered independent of the family, it is more correct to use the local rescaling (*in this case, due to the technicalities that arise we must add some additional restrictions on which $t \in [R, 2R]$ are in the family*).

Before stating our main result, we must first introduce some notation. All conjectures are stated in full in Appendix A.

Definition 1.1 *(Sieved family).* Let $\mathcal{E} : y^2 = x^3 + A(T)x + B(T)$ be a one-parameter family of elliptic curves over $\mathbb{Q}(T)$ with discriminant $\Delta(T)$, let $D(T)$ be the product of the irreducible polynomial factors of the discriminant, and let B be the largest square dividing $D(t)$ for all integers t. For a fixed c, t_0, our family is the set of all $t = ct' + t_0$ (with $t \in [R, 2R]$) such that $D(ct' + t_0)$ is square-free except for primes $p|B$ where the power of such $p|D(t)$ is independent of t. In [Mil2] it is shown that for any one-parameter family, there is a choice of c and t_0 such that the number of such t is $c_{\mathcal{E}}R + o(R)$ for some $c_{\mathcal{E}} > 0$ if every irreducible polynomial factor of $\Delta(T)$ has degree at most 3 (if not, the claim is true if we assume either the ABC or Square-Free Sieve Conjecture). We let $\mathcal{F}'_R$ denote the sieved family.

Definition 1.2 *(Average number of zeros in a family).* Let $\mathcal{E} : y^2 = x^3 + A(T)x + B(T)$ be a one-parameter family of elliptic curves over $\mathbb{Q}(T)$ with specialized curves E_t with conductors N_t. Assume GRH and write the non-trivial zeros of $\Lambda(s, E_t)$ as $\frac{1}{2} + i\gamma_{t,j}$, and set

$$\frac{\log N}{2\pi} := \frac{1}{R}\sum_{t=R}^{2R}\frac{\log N_t}{2\pi}. \tag{1.1}$$

The average number of zeros with imaginary part at most τ (in absolute value) under the global and local renormalizations are defined to be

$$\begin{aligned} Z^{(\mathrm{global})}_{\mathrm{ave},\mathcal{E},R}(\tau) &:= \frac{1}{R}\sum_{t=R}^{2R}\#\left\{j:\ \gamma_{t,j}\frac{\log N}{2\pi} \in [-\tau, \tau]\right\}, \\ Z^{(\mathrm{local})}_{\mathrm{ave},\mathcal{E},R}(\tau) &:= \frac{1}{|\mathcal{F}'_R|}\sum_{\substack{t=R\\ t\in\mathcal{F}'_R}}^{2R}\#\left\{j:\ \gamma_{t,j}\frac{\log N_t}{2\pi} \in [-\tau, \tau]\right\}, \end{aligned} \tag{1.2}$$

with $\mathcal{F}'_R$ as in Definition 1.1.

The Birch and Swinnerton-Dyer conjecture implies that, for families where half the curves have even and half have odd sign,

$$Z^{(\mathrm{global})}_{\mathrm{ave},\mathcal{E},R}(0) = Z^{(\mathrm{local})}_{\mathrm{ave},\mathcal{E},R}(0) \geqslant r + \frac{1}{2}.$$

Our main results are upper and lower bounds for how many normalized zeros there are on average in the interval $[-\tau, \tau]$, in particular, how small we may take τ and be assured on average that there are $r + \frac{1}{2}$ zeros in the interval.

Theorem 1.3. *Let $\mathcal{E}$ be a one-parameter family of elliptic curves of geometric rank r over $\mathbb{Q}(T)$; if $\mathcal{E}$ is not a rational surface (see Remark* A.1 *for a definition) then assume Tate's conjecture. Additionally, if we are using the local renormalization of the zeros we must assume either the ABC or the Square-Free Sieve Conjecture if the discriminant has an irreducible polynomial factor of degree at least* 4.

Let σ be chosen such that we can compute the one-level density (defined in Section 2.3*) for even Schwartz test functions ϕ with* $\operatorname{supp}(\widehat{\phi}) \subset (-\sigma, \sigma)$; *see Theorem* 2.3 *for details on what σ are permissible for a given family.*

Then

- Lower bounds for the average number of normalized zeros in $[-\tau, \tau]$. *Let the notation be as in Definition* 1.2, *and assume GRH. Let h be any even, twice continuously differentiable function supported on* $[-1, 1]$ *and monotonically decreasing on* $[0, 1]$. *For fixed $\tau > 0$ let $f(y) = h(2y/\sigma)$, $g(y) = (f * f)(y)$ (the convolution of f with itself), and let $\phi(x)$ equal the Fourier transform of $g(y) + (2\pi\tau)^{-2} g''(y)$. Note* $\operatorname{supp}(\widehat{\phi}) \subset (-\sigma, \sigma)$ *and $\phi(x)$ is non-negative for $|x| < \tau$ and non-positive for $|x| > \tau$. Then*

$$Z^{(\mathrm{global})}_{\mathrm{ave},\mathcal{E},R}(\tau),\, Z^{(\mathrm{local})}_{\mathrm{ave},\mathcal{E},R}(\tau) \geqslant \left(r + \frac{1}{2}\right) + \frac{\widehat{\phi}(0)}{\phi(0)} + O\left(\frac{\log\log R}{\phi(0)\log R}\right), \tag{1.3}$$

where $\widehat{\phi}(0)/\phi(0)$ depends on the fixed τ:

$$\frac{\widehat{\phi}(0)}{\phi(0)} = \frac{(\int_0^1 h(u)^2\,du) + (\frac{1}{\sigma\tau\pi})^2(\int_0^1 h(u)h''(u)\,du)}{\sigma(\int_0^1 h(u)\,du)^2}. \tag{1.4}$$

If we let $\tau_{\mathrm{BSD}}(\sigma)$ denote the value of τ such that we are assured of at least $r + \frac{1}{2}$ zeros on average (as $R \to \infty$) in $[-\tau, \tau]$ given that we can compute the one-level density for test functions whose Fourier transform is supported in $(-\sigma, \sigma)$, then

$$\tau_{\mathrm{BSD}}(\sigma) \leqslant \frac{1}{\pi}\left(-\frac{\int_0^1 h(u)^2\,du}{\int_0^1 h(u)h''(u)\,du}\right)^{-1/2}\frac{1}{\sigma} := \frac{1}{\pi C(h)\sigma}. \tag{1.5}$$

This should be compared to the predictions from the Birch and Swinnerton-Dyer and parity conjectures for a generic family, which predict $\tau_{\mathrm{BSD}}(\sigma) = 0$. In particular, taking

$$h(x) = \begin{cases} (1 - x^2)(1 - 0.233428x^2 + 0.0189588x^4) & \text{if } |x| \leqslant 1, \\ 0 & \text{otherwise} \end{cases} \tag{1.6}$$

yields

$$\tau_{\mathrm{BSD}}(\sigma) \leqslant \frac{1}{\pi C(h)\sigma}, \tag{1.7}$$

where $C(h) \approx 0.63662$ (which is approximately $2/\pi$); note $1/\pi C(h)\sigma$ is approximately $1/2\sigma$. In the arguments below we use $2/\pi$ for brevity without reminding the reader that the numerical calculation is only close to the above.

- Upper bounds for the average number of normalized zeros in $[-\tau,\tau]$. *Let ψ be a twice continuously differentiable even Schwartz test function with* $\mathrm{supp}(\widehat{\psi}) \subset (-\sigma,\sigma)$, $\psi(x) \geqslant 0$ *for all x, and* $\psi(x)$ *monotonically decreasing on* $[0,\tau)$. *Then*

$$Z_{\mathrm{ave},\mathcal{E},R}^{(\mathrm{global})}(\tau),\ Z_{\mathrm{ave},\mathcal{E},R}^{(\mathrm{local})}(\tau) \leqslant \left(r+\frac{1}{2}\right) + \frac{(r+\frac{1}{2})(\psi(0)-\psi(\tau))+\widehat{\psi}(0)}{\psi(\tau)} + O\left(\frac{\log\log R}{\psi(0)\log R}\right). \tag{1.8}$$

If we consider the interval $(-\frac{1}{2\sigma},\frac{1}{2\sigma})$ from the lower bound, taking $\psi(x)=(\frac{\sin x\pi\sigma}{x\pi\sigma})^2$ yields the average number of normalized zeros in the limit in this interval is at most $(r+\frac{1}{2}+\frac{1}{\sigma})/\psi(1/2\sigma)=\frac{\pi^2}{4}(r+\frac{1}{2}+\frac{1}{\sigma})$.

- Random matrix theory prediction. *Let $\mathcal{E}$ be a generic one-parameter family of elliptic curves of rank r over $\mathbb{Q}(T)$ with half of the specialized functional equations even and half odd. Assuming the Katz–Sarnak density conjecture, as $R\to\infty$ the average number of normalized zeros in $[-\tau,\tau]$ is $(r+\frac{1}{2})+2\tau$; more precisely, random matrix theory predicts*

$$\lim_{R\to\infty} Z_{\mathrm{ave},\mathcal{E},R}^{(\mathrm{global})}(\tau),\ \lim_{R\to\infty} Z_{\mathrm{ave},\mathcal{E},R}^{(\mathrm{local})}(\tau) = r+\frac{1}{2}+2\tau. \tag{1.9}$$

In particular, setting $\tau=\frac{1}{2\sigma}$ yields a prediction of $r+\frac{1}{2}+2\cdot\frac{1}{2\sigma}$ normalized zeros in the limit on average.

In summary, the number of normalized zeros on average as $R\to\infty$ in the interval $(-\frac{1}{2\sigma},\frac{1}{2\sigma})$ satisfies

$$r+\frac{1}{2} \leqslant \lim_{R\to\infty} Z_{\mathrm{ave},\mathcal{E},R}^{(\mathrm{global})}\left(\frac{1}{2\sigma}\right),\ \lim_{R\to\infty} Z_{\mathrm{ave},\mathcal{E},R}^{(\mathrm{local})}\left(\frac{1}{2\sigma}\right) \leqslant \frac{\pi^2}{4}\left(r+\frac{1}{2}+\frac{1}{\sigma}\right), \tag{1.10}$$

and this interval contains the prediction from random matrix theory, $r+\frac{1}{2}+\frac{1}{\sigma}$.

Remark 1.4. We obtained our upper bound for $\tau_{\mathrm{BSD}}(\sigma)$ by setting $\widehat{\phi}(0)/\phi(0)=0$. The important item to note is that $\tau_{\mathrm{BSD}}(\sigma)$ (or any τ) is inversely proportional to the support σ. In other words, the larger we may take σ, the more we may concentrate ϕ near the central point and thus the smaller the window. Random matrix theory predicts we may take σ arbitrarily large, which would imply we may take τ arbitrarily small and thus prove the Birch and Swinnerton-Dyer conjecture on average.

2. Background material and previous results

2.1. Elliptic curves

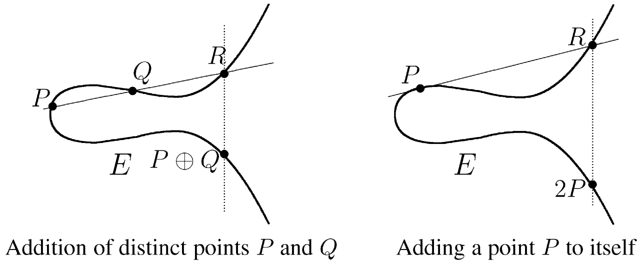
We quickly review the needed background material on elliptic curves; the reader familiar with the notation and theory may safely skip this subsection. See [Kn,Kob,Sil1,ST] for proofs, as well as the survey [Yo1].

Let E be an elliptic curve over $\mathbb{Q}$, say $y^2=x^3+ax+b$ with $a,b\in\mathbb{Z}$, and set

$$E(\mathbb{Q}) := \{(x,y)\in\mathbb{Q}^2:\ y^2=x^3+ax+b\}. \tag{2.1}$$

We can define addition of two elements of $E(\mathbb{Q})$ as follows (see Fig. 1). If $P=(x_1,y_1)$ and $Q=(x_2,y_2)$ are in $E(\mathbb{Q})$, then the line $y=mx+b$ connecting them has rational coordinates.[2] Substituting this expression for y into the elliptic curve, we find $(mx+b)^2=x^3+ax+b$. This is a cubic in x

[2] We assume the two points are distinct; if they are the same, the argument below must be slightly modified.



Addition of distinct points P and Q

Adding a point P to itself

Fig. 1. The addition law on an elliptic curve. In the second example the line is tangent to E at P.

with rational coefficients. By construction two of its roots are x_1 and x_2, both rational numbers. Thus the third root, say x_3, must also be rational. Set $R(P,Q) = (x_3, \sqrt{x_3^3 + ax_3 + b})$ and $\widetilde{R}(P,Q) = (x_3, -\sqrt{x_3^3 + ax_3 + b})$. If we define addition by $P \oplus Q = \widetilde{R}(P,Q)$, then this (plus adding a 'point at infinity') turns $E(\mathbb{Q})$ into a finitely generated abelian group. We write $E(\mathbb{Q})$ as $\mathbb{Z}^r \oplus \mathbb{T}$, where $\mathbb{T}$ is a torsion group[3] and r is called the geometric rank of the curve.

Given an elliptic curve E as above, we may associate an L-function as follows. Assume $y^2 = x^3 + ax + b$ is a globally minimal Weierstrass equation for $E/\mathbb{Q}$ with discriminant $\Delta = -16(4a^3 + 27b^2)$ and conductor N_E. Set

$$a_E(p) := p - \#\{(x,y) \in (\mathbb{Z}/p\mathbb{Z})^2 \colon\ y^2 \equiv x^3 + ax + b \bmod p\}. \tag{2.2}$$

Note that the $a_E(p)$'s encode local data, specifically the number of solutions modulo p. Hasse proved $|a_E(p)| \leqslant 2\sqrt{p}$, and we define the L-function by

$$L(s,E) := \prod_{p|\Delta}\left(1 - \frac{a_E(p)}{\sqrt{p}}p^{-s}\right)^{-1} \prod_{p\nmid\Delta}\left(1 - \frac{a_E(p)}{\sqrt{p}}p^{-s} + p^{-2s}\right)^{-1}; \tag{2.3}$$

we have included the factors of $\sqrt{p}$ so that the completed L-function has a functional equation from s to $1-s$ and not $2-s$:

$$\Lambda(s,E) := \left(\frac{\sqrt{N}}{2\pi}\right)^s \Gamma\left(s + \frac{1}{2}\right) L(s,E) = \epsilon_E \Lambda(1-s,E), \tag{2.4}$$

where $\epsilon_E \in \{1,-1\}$ is the sign of the functional equation. The following work of Wiles [Wi], Taylor and Wiles [TW] and Breuil, Conrad, Diamond, and Taylor [BCDT], we may associate a weight 2 modular form f to any elliptic curve E, where the level of f equals the conductor N_E of E. We have $\Lambda(s,f) = \Lambda(s,E)$; in particular, the completed L-function converges for all s. We call the order of vanishing of $\Lambda(s,E)$ at $s = 1/2$ the analytic rank of E.

The Birch and Swinnerton-Dyer conjecture [BS-D1,BS-D2] states[4] that the order of vanishing of $\Lambda(s,E)$ at the central point $s = 1/2$ equals the rank of the Mordell–Weil group $E(\mathbb{Q})$, or that the analytic rank equals the geometric rank. Sadly, we are far from being able to prove this, though the evidence for the conjecture is compelling, especially in the case of complex multiplication and rank at most 1 [Bro,CW,GKZ,GZ,Kol1,Kol2,Ru]. In addition there is much suggestive numerical evidence for the

[3] Mazur [Ma] proved that torsion group is one of the following: $\mathbb{Z}/N\mathbb{Z}$ for $N \in \{1,2,\ldots,10,12\}$ or $\mathbb{Z}/2 \times \mathbb{Z}/2N\mathbb{Z}$ for $N \in \{1,2,3,4\}$.

[4] There is a more precise form of the conjecture which relates the leading term in the Taylor expansion to the period integral, regulator, Tamagawa numbers and the Tate–Shafarevich group, but this version is not needed for our purposes.

conjecture; for example, for elliptic curves with modest geometric rank r, numerical approximations of the first $r-1$ Taylor coefficients are consistent with these coefficients vanishing.

2.2. Explicit formula

One powerful tool for investigating the Birch and Swinnerton-Dyer conjecture is the explicit formula (see [RS] for a proof for a general L-function, or [Mil1] for the calculation for elliptic curves), which connects the zeros of an L-function to the Fourier coefficients.

Theorem 2.1. *Let ϕ be an even, twice continuously differentiable test function whose Fourier transform*

$$\widehat{\phi}(y) := \int_{-\infty}^{\infty} \phi(x) e^{-2\pi ixy}\, dx \tag{2.5}$$

has compact support, and denote the non-trivial zeros of $\Lambda(s,E)$ by $\frac{1}{2} + i\gamma_{E;j}$ (under the Generalized Riemann Hypothesis, each $\gamma_{E;j} \in \mathbb{R}$). Then

$$\begin{aligned}\sum_{\gamma_{E;j}} \phi\left(\gamma_{E;j}\frac{\log N_E}{2\pi}\right) = \widehat{\phi}(0) + \phi(0) - 2\sum_p \frac{a_E(p)\log p}{p\log N_E}\widehat{\phi}\left(\frac{\log p}{\log N_E}\right) \\ - 2\sum_p \frac{a_E^2(p)\log p}{p^2\log N_E}\widehat{\phi}\left(\frac{2\log p}{\log N_E}\right) + O\left(\frac{\log\log N_E}{\log N_E}\right).\end{aligned} \tag{2.6}$$

Using the explicit formula, Mestre proved the following theorem.[5]

Theorem 2.2. *(See Mestre [Mes].) Assuming the Generalized Riemann Hypothesis*:

(1) *The order of vanishing at the central point is $O(\log N_E / \log\log N_E)$.*
(2) *There is an absolute constant B such that the first zero above the central point occurs before $B/\log\log N_E$.*

From the functional equation, however, we expect the first zero above the central point to be on the order of $1/\log N_E$, and not $1/\log\log N_E$. Thus Mestre's result is significantly larger than what we expect the truth to be; however, it holds for *any* elliptic curve. The situation is very different if instead we consider families of elliptic curves. By averaging the explicit formula over the family and exploiting cancelation in the sums of the Fourier coefficients $a_E(p)$, it is possible to prove (on average) significantly better results.

Numerous studies have been concerned with bounding the average rank in families. We list some of the frequently studied families below (note that, for technical reasons, often one has to do some sieving and remove some curves in order to make certain sums tractable). These results are obtained by averaging the explicit formula over some family $\mathcal{F}_R$, where R is a parameter localizing the conductors, and sending $R \to \infty$.

- The family of all elliptic curves: $y^2 = x^3 + Ax + B$, and $\mathcal{F}_R = \{(A,B)\colon\ |A| \leqslant R^2,\ |B| \leqslant R^3\}$ (or something along these lines).
- One parameter families over $\mathbb{Q}(T)$: $y^2 = x^3 + A(T)x + B(T)$, with $A(T), B(T) \in \mathbb{Z}[T]$ and either $\mathcal{F}_R = \{t\colon\ R \leqslant t \leqslant 2R\}$ or a sub-family of this where the conductors are given by a polynomial.

[5] Mestre actually proved more, as his results hold for any weight k cuspidal newform, and not just elliptic curves (which correspond to weight 2 cuspidal newforms).

- Quadratic (or higher) twists of a fixed elliptic curve: $dy^2 = x^3 + ax + b$, with $\mathcal{F}_R = \{d\colon\ d \leqslant R$ a fundamental discriminant$\}$.

The current record belongs to M. Young [Yo2], who showed the average rank in the family of all elliptic curves is bounded by $25/14 \approx 1.79$; results for one-parameter families and quadratic twist families are significantly worse. For a sample of the literature, see [BMSW,Bru,BM,CPRW,DFK,Gao,Go, GM,H-B,Kow1,Kow2,Mi,Mil2,RSi,RuSi,Sil3,Yo2,ZK] (especially the surveys [BMSW,Kow1,RuSi]).

2.3. The one-level density

For a family $\mathcal{F}_R$ of L-functions ordered by conductor (with $R \to \infty$), the averaged explicit formula is called the one-level density. Specifically, let ϕ be an even Schwartz test function whose Fourier transform is supported in $(-\sigma, \sigma)$, and denote the zeros of $L(s, f)$ by $1/2 + i\gamma_{f,j}$ (under GRH each $\gamma_{f,j} \in \mathbb{R}$). Let N_f denote the analytic conductor of $L(s, f)$. We define the one-level density by

$$D_{\mathcal{F}_R}(\phi) := \frac{1}{|\mathcal{F}_R|} \sum_{f \in \mathcal{F}_R} \sum_j \phi\left(\gamma_{f;j} \frac{\log N_f}{2\pi}\right). \tag{2.7}$$

This statistic has been fruitfully used by many researchers to study the zeros of elliptic curves L-functions (as well as other families of L-functions) near the central point.

Unlike the n-level correlations, which are the same for any cuspidal newform arising from an automorphic representation (see [Hej,Mon,RS]), the one-level density for a family of L-functions depends on the symmetry of the family. Katz and Sarnak [KS1,KS2] conjecture that families of L-functions correspond to classical compact groups; specifically, the behavior as the conductors tend to infinity of zeros (respectively values) of L-functions is well modeled by the limit as the matrix size tends to infinity of roots (respectively values) of characteristic values of random matrices.[6] They conjecture that

$$\lim_{R \to \infty} D_{\mathcal{F}_R}(\phi) = \int \phi(x) W_{G(\mathcal{F})}(x)\, dx, \tag{2.8}$$

where $G(\mathcal{F})$ indicates unitary, symplectic or orthogonal (possibly SO(even) or SO(odd)) symmetry; this has been observed in numerous families. Note by Parseval's theorem that

$$\int \phi(x) W_{G(\mathcal{F})}(x)\, dx = \int \widehat{\phi}(y) \widehat{W}_{G(\mathcal{F})}(y)\, dy. \tag{2.9}$$

Let $I(u)$ be the characteristic function of $[-1, 1]$. Katz and Sarnak prove the Fourier transforms of the one-level densities of the classical compact groups are

$$\widehat{W}_{\mathrm{SO(even)}}(u) = \delta(u) + \frac{1}{2} I(u),$$
$$\widehat{W}_{\mathrm{SO}}(u) = \delta(u) + \frac{1}{2},$$
$$\widehat{W}_{\mathrm{SO(odd)}}(u) = \delta(u) - \frac{1}{2} I(u) + 1,$$

[6] These conjectures are a natural outgrowth of observed similarities between behavior of L-functions and matrix ensembles. While random matrix theory first arose in statistics problems in the early 1900s (see for example [Wis]), it blossomed in the 1950s when it was successfully applied to describe the energy levels of heavy nuclei. Its connections to number theory were first noticed by Montgomery [Mon] and Dyson in the 1970s in studies of the pair correlation of zeros of $\zeta(s)$. See [FM] for a survey on the development of the subject and some of the connections between the two fields.

$$\widehat{W}_{\mathrm{USp}}(u) = \delta(u) - \frac{1}{2}I(u),$$
$$\widehat{W}_{\mathrm{U}}(u) = \delta(u). \tag{2.10}$$

For functions whose Fourier transforms are supported in $[-1,1]$, the three orthogonal densities are indistinguishable, though 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 $[-1,1]$.[7]

For families of elliptic curves with rank, it is useful to consider additional subgroups of the classical compact groups above. We consider the $N \to \infty$ scaling limits of matrices of the form

$$\begin{pmatrix} I_{r,r} & \\ & g \end{pmatrix},$$

where $I_{r,r}$ is the $r \times r$ identity matrix and g is an $N \times N$ orthogonal matrix (drawn from either the full orthogonal family or one of the split families, namely even or odd). These matrices have r forced eigenvalues at 1 (or r eigenangles at 0) for each g; thus as we vary g in one of the three families we obtain the same one-level densities as before *except* for an additional factor of r. Explicitly,

$$\widehat{W}_{r;\mathrm{SO(even)}}(u) = \delta(u) + \frac{1}{2}I(u) + r,$$
$$\widehat{W}_{r;\mathrm{SO}}(u) = \delta(u) + \frac{1}{2} + r,$$
$$\widehat{W}_{r;\mathrm{SO(odd)}}(u) = \delta(u) - \frac{1}{2}I(u) + 1 + r. \tag{2.11}$$

For our elliptic curve families, we must evaluate the average over $\mathcal{F}_R$ or $\mathcal{F}'_R$ of (2.6). Note that almost all of the conductors will be a bounded power of R for $t \in [R, 2R]$. If we rescale each elliptic curve E's zeros by the correct local factor, namely $(\log N_E)/2\pi$, we have

$$\begin{aligned} D_{\mathcal{F}_R}^{\mathrm{local}}(\phi) &= \frac{1}{|\mathcal{F}_R|}\sum_{E\in\mathcal{F}_R}\sum_{\gamma_{E;j}}\phi\left(\gamma_{E;j}\frac{\log N_E}{2\pi}\right) \\ &= \widehat{\phi}(0) + \phi(0) - 2\frac{1}{|\mathcal{F}_R|}\sum_{E\in\mathcal{F}_R}\sum_p \frac{a_E(p)\log p}{p\log N_E}\widehat{\phi}\left(\frac{\log p}{\log N_E}\right) \\ &\quad - 2\frac{1}{|\mathcal{F}_R|}\sum_{E\in\mathcal{F}_R}\sum_p \frac{a_E^2(p)\log p}{p^2\log N_E}\widehat{\phi}\left(\frac{2\log p}{\log N_E}\right) + O\left(\frac{\log\log R}{\log R}\right). \end{aligned} \tag{2.12}$$

The difficulty with this expression is that, as the conductors are varying, we cannot easily pass the sum over the family through the test function to the Fourier coefficients $a_E(p)$ and $a_E(p)^2$. By sieving it is possible to surmount these technical details; this is the main result in [Mil2].

[7] One can also distinguish between the various orthogonal groups by looking at the 2-level density, as these three ensembles have distinct behavior for arbitrarily small support; see for instance [Mil2]. If $n \geqslant 3$, the determinant expansions for the n-level density are hard to work with; in fact, in Gao's thesis [Gao] he is able to compute the number theory and random matrix theory results for greater support than he can show agreement. In place of the determinant formulas, one can also use expansions from [HM]; though these hold for smaller support, they are sometimes easier for comparisons.

If instead we rescale the zeros of each elliptic curve E's L-function by the global factor, namely

$$\frac{\log N}{2\pi} = \frac{1}{|\mathcal{F}_R|} \sum_{t \in \mathcal{F}_R} \frac{\log N_E}{2\pi}, \tag{2.13}$$

then we find

$$\begin{aligned} D_{\mathcal{F}_R}^{\text{global}}(\phi) &= \frac{1}{|\mathcal{F}_R|} \sum_{E \in \mathcal{F}_R} \sum_{\gamma_{E;j}} \phi\left(\gamma_{E;j} \frac{\log N}{2\pi}\right) \\ &= \widehat{\phi}(0) + \phi(0) - 2\frac{1}{|\mathcal{F}_R|} \sum_{E \in \mathcal{F}_R} \sum_p \frac{a_E(p) \log p}{p \log N} \widehat{\phi}\left(\frac{\log p}{\log N}\right) \\ &\quad - 2\frac{1}{|\mathcal{F}_R|} \sum_{E \in \mathcal{F}_R} \sum_p \frac{a_E^2(p) \log p}{p^2 \log N} \widehat{\phi}\left(\frac{2 \log p}{\log N}\right) + O\left(\frac{\log \log N}{\log N}\right). \end{aligned} \tag{2.14}$$

The analysis is significantly easier here, as now we can pass the summation over the family past the test function and exploit cancelation in sums of the Fourier coefficients $a_E(p)$ and $a_E(p)^2$.

We quote the best known results for general one-parameter families.

Theorem 2.3. *(See Miller, Theorem 7.8 of [Mil1] or Theorem 5.8 of [Mil2].) Notation*:

- *Let $\mathcal{E}$ be a one-parameter family of elliptic curves of geometric rank r over $\mathbb{Q}(T)$.*
- *Let ϕ be a twice continuously differentiable function*[8] *with* $\mathrm{supp}(\widehat{\phi}) \subset (-\sigma, \sigma)$.
- *Consider the sieved family* (*see Definition* 1.1), *and denote the degree of the conductor polynomial by m.*
- *Let G denote either* SO, SO(even) *or* SO(odd).

Assume

- *If $\mathcal{E}$ is not a rational surface* (*see Remark* A.1 *for a definition*) *then assume Tate's conjecture.*
- *If the discriminant has an irreducible polynomial factor of degree at least* 4, *assume either the ABC or the Square-Free Sieve Conjecture.*

Then

$$D_{\mathcal{F}_R}^{\text{local}}(\phi) = \int \widehat{\phi}(y) \widehat{W}_{r;G}(y)\, dy = \left(r + \frac{1}{2}\right)\phi(0) + \widehat{\phi}(0) + O\left(\frac{\log \log R}{\log R}\right) \tag{2.15}$$

provided $\sigma < \min(1/2, 2/3m)$; *a similar result holds for* $D_{\mathcal{F}_R}^{\text{global}}(\phi)$ (*without the assumptions that $\mathcal{E}$ satisfies Tate's hypothesis and without assuming either the ABC or Square-Free Sieve Conjecture*).

Remark 2.4. We briefly discuss some consequences and generalizations of the above theorem.

- Similar statements hold for quadratic twist families and the family of all elliptic curves.
- The above result provides evidence that the zeros of one-parameter families of rank r over $\mathbb{Q}(T)$ are modeled by the scaling limits of orthogonal matrices with r independent eigenvalues of 1.

[8] While the theorem was proved under the assumption that ϕ is Schwartz, a careful analysis of the argument reveals it suffices that ϕ be twice differentiable.

- As $\operatorname{supp}(\widehat{\phi}) \subset (-1, 1)$, the three orthogonal groups have indistinguishable one-level densities. We can see which group correctly models our family by studying the 2-level density, which requires us to understand the distribution of signs of the functional equations in our family.

3. Proof of Theorem 1.3

3.1. Preliminaries

Before proving Theorem 1.3, we first prove general results for the upper and lower bounds in a window of variable size for a general family of L-functions. Theorem 1.3 then follows immediately from Theorem 3.1, Theorem 2.3 and the constructions of test functions satisfying the necessary conditions, which are given below.

Theorem 3.1. *Let $\mathcal{F}_R$ denote a family of L-functions, and let $Z^{(\text{global})}_{\text{ave},\mathcal{F},R}(\tau)$, $Z^{(\text{local})}_{\text{ave},\mathcal{F},R}(\tau)$ be defined as in Definition* 1.2. *Let $\phi(x)$ and $\psi(x)$ be twice continuously differentiable functions with Fourier transform supported in $(-\sigma, \sigma)$. Assume for both normalizations of zeros that there are constants a and b such that one has*

$$D_{\mathcal{F}_R}(\phi) = a\phi(0) + b\widehat{\phi}(0) + O\left(\frac{\log\log R}{\log R}\right), \tag{3.1}$$

as well as the corresponding formula for $D_{\mathcal{F}_R}(\psi)$ with replacing ϕ by ψ in (3.1). If $\phi(x) \geqslant 0$ for $|x| \leqslant \tau$ and $\phi(x) \leqslant 0$ whenever $|x| \geqslant \tau$, and if $\phi(x)$ is largest when $x = 0$, then

$$Z^{(\text{global})}_{\text{ave},\mathcal{F},R}(\tau), Z^{(\text{local})}_{\text{ave},\mathcal{F},R}(\tau) \geqslant a + b\frac{\widehat{\phi}(0)}{\phi(0)} + O\left(\frac{\log\log R}{\phi(0)\log R}\right), \tag{3.2}$$

while if $\psi(x) \geqslant 0$ for all x and is monotonically decreasing on $(0, \tau)$, then

$$Z^{(\text{global})}_{\text{ave},\mathcal{F},R}(\tau), Z^{(\text{local})}_{\text{ave},\mathcal{F},R}(\tau) \leqslant a + \frac{a(\psi(0) - \psi(\tau)) + b\widehat{\psi}(0)}{\psi(\tau)} + O\left(\frac{\log\log R}{\psi(0)\log R}\right). \tag{3.3}$$

Proof. We give the proof for the local rescaling; the global case follows analogously. As $\phi(x)$ is non-positive for $|x| \geqslant \tau$, the contribution to the one-level density from the scaled zeros as large or larger than τ in absolute value is non-positive; thus if we remove these contributions then the one-level density gives the lower bound

$$\frac{1}{|\mathcal{F}_R|}\sum_{f\in\mathcal{F}_R}\sum_{|\gamma_{f;j}|\leqslant\tau}\phi(\widetilde{\gamma}_{f;j}) \geqslant a\phi(0) + b\widehat{\phi}(0) + O\left(\frac{\log\log R}{\log R}\right). \tag{3.4}$$

As ϕ is maximized at 0, we increase the left-hand side above by replacing $\phi(\widetilde{\gamma}_{f;j})$ with $\phi(0)$; doing so and dividing by $\phi(0)$ yields the claimed bound for $Z^{(\text{local})}_{\text{ave},\mathcal{F},R}(\tau)$. The upper bound is proved analogously. □

Remark 3.2. These results are of course not of interest unless we are able to construct ϕ and ψ satisfying the conditions in Theorem 3.1. For one-parameter families of elliptic curves of rank r over $\mathbb{Q}(T)$, we have $a = r + \frac{1}{2}$ and $b = 1$.

Remark 3.3. For test functions whose Fourier transform is supported in $(-1, 1)$, all known one-level densities of families of L-functions are in the form of Theorem 3.1, and thus our results are immediately applicable. For some families where the support exceeds $(-1, 1)$ (such as families of cuspidal newforms of square-free level split by sign of the functional equation), a little more work is needed as the functional form of the one-level density is different.[9] For ease of exposition in this paper we confine ourselves to the $(-1,1)$ case.

3.2. Proof of Theorem 1.3

The main step in the proof of Theorem 1.3 is showing that our result is non-vacuous by constructing ϕ and ψ with the claimed properties. Our construction of ϕ is almost surely similar to the construction implicit in Mestre's work [Mes]; see also Hughes and Rudnick [HR].

Proof of the lower bound in Theorem 1.3. We give the lower bound for the number of zeros in $[-\tau,\tau]$ by constructing a good test function ϕ. As our results depend on the support of $\widehat{\phi}$ (which is finite), it is convenient to normalize our test function and express everything in terms of h, which we take to be an even, twice continuously differentiable function supported on $(-1,1)$ and monotonically decreasing on $[0,1)$. For fixed $\sigma,\tau > 0$ let $f(y) = h(2y/\sigma)$, $g(y) = (f * f)(y)$ (the convolution[10] of f with itself), and let $\phi(x)$ equal the Fourier transform of $g(y) + (2\pi\tau)^{-2}g''(y)$. We must show (i) $\operatorname{supp}(\widehat{\phi}) \subset (-\sigma,\sigma)$ and (ii) $\phi(x)$ is non-negative for $|x| < \tau$ and non-positive for $|x| > \tau$.

The proof of (i) follows from standard properties of convolution. Specifically, as $\operatorname{supp}(f) \subset (-\sigma/2,\sigma/2)$, we have $\operatorname{supp}(g) \subset (-\sigma,\sigma)$.[11] As the support of g'' is contained in the support of g and $\widehat{\phi}(y) = g(y) + (2\pi\tau)^{-2}g''(y)$, the support of $\widehat{\phi}$ is contained in $(-\sigma,\sigma)$ as claimed.

For (ii), the Fourier transform of $g''(y)$ is $-(2\pi y)^2\widehat{g}(y)$ (the Fourier transform converts differentiation to multiplication by $2\pi ix$ in our normalization). Further $g = f * f$ implies $g'' = f * f''$. Combining the above, we find[12] the Fourier transform of $\widehat{\phi}(y) = g(y) + (2\pi\tau)^{-2}g''(y)$ is $\phi(x) = \widehat{g}(x)\cdot(1-(x/\tau)^2)$.

To complete the proof, we must show

$$\frac{\widehat{\phi}(0)}{\phi(0)} = \frac{(\int_0^1 h(u)^2\,du) + (\frac{1}{\sigma\tau\pi})^2(\int_0^1 h(u)h''(u)\,du)}{\sigma(\int_0^1 h(u)\,du)^2}. \tag{3.5}$$

By construction we have

$$\frac{\widehat{\phi}(0)}{\phi(0)} = \frac{g(0) + (2\pi\tau)^{-2}g''(0)}{\widehat{g}(0)}. \tag{3.6}$$

Since g is even and monotonically decreasing near the origin (as g has a maximum at 0), $g''(0) < 0$. Thus larger values of τ should decrease the ratio above, at the cost of increasing the size of our window.

From our construction, as h and f are even we have

$$g(0) = \int_{-\sigma/2}^{\sigma/2} f(t)^2\,dt = 2\int_0^{\sigma/2} h\left(\frac{2t}{\sigma}\right)dt = \sigma\int_0^1 h(u)^2\,du \tag{3.7}$$

and

[9] For the family of Dirichlet characters of prime conductor, the one-level density is known to be $\widehat{\phi}(0)$ for support is known up to $(-2,2)$, and thus is of the desired form.

[10] The convolution is defined by $(A * B)(x) = \int_{-\infty}^{\infty} A(t)B(x-t)\,dt$.

[11] We may interpret the relation between f and g as follows. Let X be a random variable with density f supported in $(-\sigma/2,\sigma/2)$. Then $g = f * f$ is the density of $X + X$, and is supported in $(-\sigma,\sigma)$.

[12] As ϕ and $\widehat{\phi}$ are even, the Fourier transform of the Fourier transform is the original function $\phi(x)$; if ϕ were not even, we would have to replace $\phi(x)$ with $\phi(-x)$.

$$\begin{aligned}
g''(0) &= \int_{-\sigma/2}^{\sigma/2} f(t) f''(t)\,dt \\
&= 2\int_{0}^{\sigma/2} f(t) f''(t)\,dt \\
&= \frac{8}{\sigma^2}\int_{0}^{\sigma/2} h\left(\frac{2t}{\sigma}\right) h''\left(\frac{2t}{\sigma}\right) dt \quad \left(\text{since } f(t) = h\left(\frac{2t}{\sigma}\right),\ f''(t) = \frac{4}{\sigma^2} h\left(\frac{2t}{\sigma}\right)\right) \\
&= \frac{4}{\sigma}\int_{0}^{1} h(u) h''(u)\,du. \qquad (3.8)
\end{aligned}$$

As the Fourier transform of a convolution is the product of the Fourier transforms, a straightforward calculation yields

$$\widehat{g}(0) = \widehat{f}(0)\cdot\widehat{f}(0) = \sigma^2\left(\int_0^1 h(u)\,du\right)^2. \qquad (3.9)$$

Collecting the above equalities, after some elementary algebra we can express the ratio $\widehat{\phi}(0)/\phi(0)$ in terms of h as

$$\frac{\widehat{\phi}(0)}{\phi(0)} = \frac{(\int_0^1 h(u)^2\,du) + (\frac{1}{\sigma\tau\pi})^2(\int_0^1 h(u)h''(u)\,du)}{\sigma(\int_0^1 h(u)\,du)^2}. \qquad (3.10)$$

If we set this ratio equal to zero (i.e., if we choose τ so that the numerator vanishes) then we find[13] that on average there are at least $r + \frac{1}{2}$ normalized zeros in the band $(-\frac{1}{\pi C(h)\sigma}, \frac{1}{\pi C(h)\sigma})$, where

$$C(h) = \left(-\frac{\int_0^1 h(u)^2\,du}{\int_0^1 h(u)h''(u)\,du}\right)^{1/2}. \qquad \square \qquad (3.11)$$

Proof of the upper bound in Theorem 1.3. The proof is similar to that of the lower bound; in particular, once we construct a function ψ with the desired properties then the claim follows immediately from straightforward algebra.

We are thus again reduced to constructing a function with the specified properties. For convenience we construct a ψ which is not Schwartz, but which is twice differentiable; a careful analysis of the proof of Theorem 2.3 shows that this suffices, and thus such a ψ is sufficient for our purposes.

Consider the function $\psi(x) = (\frac{\sin x\pi\sigma}{x\pi\sigma})^2$ with a compactly supported Fourier transform given by

$$\widehat{\psi}(y) = \begin{cases} \frac{1}{\sigma}(1 - \frac{|y|}{\sigma}) & \text{if } y \in (-\sigma, \sigma), \\ 0 & \text{if } y \notin (-\sigma, \sigma); \end{cases} \qquad (3.12)$$

[13] In obvious notation, we have $\int_0^1 h^2 \geqslant -(\pi\sigma\tau_{\text{critical}})^{-2}\int_0^1 hh''$. We see $\int_0^1 hh'' \leqslant 0$, and thus $\tau_{\text{critical}} \geqslant (-\int_0^1 h^2/\int_0^1 hh'')^{-1/2}(\pi\sigma)^{-1}$.

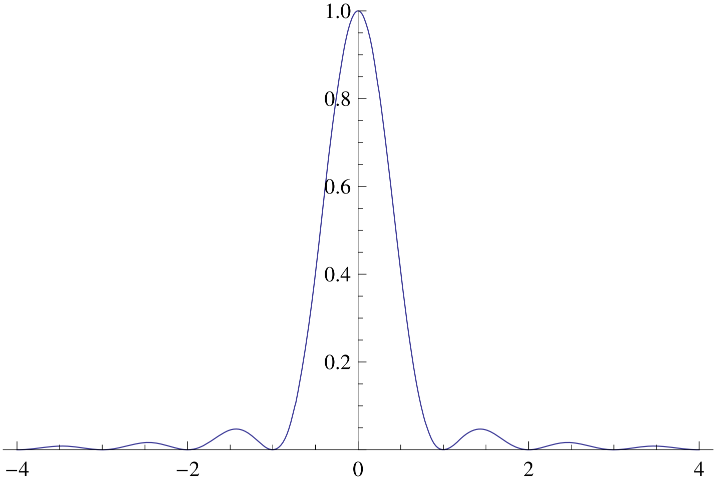


Fig. 2. Plot of $\psi(x) = (\frac{\sin(x\pi\sigma)}{x\pi\sigma})^2$ for $\sigma = 1$.

see Fig. 2 for a plot. Away from the origin, the derivative is given by

$$\psi'(x) = \frac{2\sin(\sigma\pi x)}{\sigma\pi x^2}\left(\cos(\sigma\pi x) - \frac{\sin(\sigma\pi x)}{\sigma\pi x}\right). \tag{3.13}$$

It is easy to see that the global maximum is at $x = 0$ and that $\psi(x)$ is decreasing up to $x = 1/\sigma$, proving the claim for any $\tau \leqslant 1/\sigma$ (though the bound worsens as τ approaches $1/\sigma$ as $\psi(1/\sigma) = 0$). □

Proof of the random matrix theory prediction in Theorem 1.3. We assume the conjectures from random matrix theory hold for any even test function, and not just Schwartz test functions. We therefore take $\phi(x)$ to be the characteristic function of the interval $[-\tau, \tau]$, which has Fourier transform equal to $\frac{\sin(2\pi\tau y)}{2\pi\tau y} \cdot 2\tau$. Using such a test function simply counts all normalized zeros in our family that are in $[-\tau, \tau]$ (there is no weighting as ϕ is identically 1 in this interval). Thus the predicted average number of such zeros in this interval as $R \to \infty$ is

$$\begin{aligned}\int_{-\infty}^{\infty} \widehat{\phi}(y)\widehat{W}_{r;\mathrm{SO}}(y)\,dy &= \int_{-\infty}^{\infty} \widehat{\phi}(y)\left(\delta(y) + \frac{1}{2} + r\right)dy \\ &= \left(r + \frac{1}{2}\right)\phi(0) + \widehat{\phi}(0) \\ &= r + \frac{1}{2} + 2\tau. \qquad \square \end{aligned} \tag{3.14}$$

3.3. Explicit upper and lower bounds

We conclude by determining the upper and lower bounds from Theorem 1.3 for the average number of normalized zeros in given intervals as $R \to \infty$.

We first consider the lower bound, which means we must maximize $C(h)$ (as it is in the denominator for τ, the larger $C(h)$ the smaller the window). As the optimal choice of h (in a given class of functions) is only slightly better than similar h, we do not spend too much time on determining the truly best h. Consider the family of functions given by

$$h_n(x) = \left(1 - x^2\right)\left(1 + a_2 x^2 + \cdots + a_{2i} x^{2i} + \cdots + a_{2n} x^{2n}\right). \tag{3.15}$$

We set $a_0 = 1$ as the maximum is to occur at $x = 0$, and since the ratio is invariant under rescaling the a_i's, we might as well take $a_0 = 1$. Note that each $a_{2i+1} = 0$ as our function is even. We chose h_n of this form as this forces h_n to be even and to vanish at ±1. We have

$$C(h_n) = \left(-\frac{\int_0^1 h_n(u)^2\,du}{\int_0^1 h_n(u) h_n''(u)\,du}\right)^{1/2}. \tag{3.16}$$

The optimum value of the square-root appears to be $2/\pi$. For example, when $n = 2$ we must compute

$$\begin{aligned}
&\max_{a_2,a_4}\left(-\frac{\frac{8}{15} + \frac{16a_2}{105} + \frac{8a_2^2}{315} + \frac{16a_4}{315} + \frac{16a_2a_4}{693} + \frac{8a_4^2}{1287}}{-\frac{4}{3} - \frac{8a_2}{15} - \frac{44a_2^2}{105} - \frac{8a_4}{35} - \frac{8a_2a_4}{15} - \frac{52a_4^2}{231}}\right)^{1/2} \\
&\quad = \max_{a_2,a_4}\left(\frac{6006 + 286a_2^2 + 572a_4 + 70a_4^2 + 52a_2(33 + 5a_4)}{39(385 + 121a_2^2 + 66a_4 + 65a_4^2 + 154a_2(1 + a_4))}\right)^{1/2}.
\end{aligned} \tag{3.17}$$

Using Mathematica we find the optimal values are $a_2 \approx -.233428$ and $a_4 \approx .0189588$, which leads to $C(h) \approx 0.63662$; as $2/\pi \approx 0.63662$, this suggests the optimal value of $C(h)$ might be $2/\pi$. This yields the window $(-\frac{1}{2\sigma}, \frac{1}{2\sigma})$ in which we have on average (as $R \to \infty$) $r + \frac{1}{2}$ zeros.

Remark 3.4. As we expect the true answer to be a window of size 0 (i.e., letting $\sigma = \infty$), it is not worthwhile to find the true optimum above merely to save a bit in a few decimal places. The purpose of this analysis is to show that we do see the correct number of zeros on average in the limit in a window of size proportional to $1/\sigma$; the actual value of the proportionality constant, while interesting, is in some sense immaterial as we believe the density conjecture holds for arbitrary σ.

We list some approximate values for $C(h)$ for other obvious candidates, which are all less than the 0.63662 (which is approximately $2/\pi$) found above.

- $h(x) = (1 - x^2)^2$ has $C(h) \approx 0.57735$ (with the quantity inside the square-root looking like 1/3); if we take just $(1 - x^2)$ we get $C(h) = \sqrt{2/5} \approx 0.632456$.
- $h(x) = \exp(-1/(1 - x^2))$ has $C(h) \approx 0.570024$.
- $h(x) = \exp(-.754212/(1 - x^2))$ has $C(h) \approx 0.575629$ (the value of .754212 was obtained by searching for optimal test functions among $\exp(-a/(1 - x^2))$).

We now turn to finding explicit upper bounds for the average number of normalized zeros in $[-\tau, \tau]$ as $R \to \infty$. We continue to analyze the candidate function $\psi(x) = (\frac{\sin(\pi\sigma x)}{\pi\sigma x})^2$ (see Fig. 2 for a plot). We have freedom in terms of how we rate our approximation; for example, we can decrease the upper bound if we simultaneously decrease the size of the interval.

A natural value to take for the size of our interval is the optimal interval found in the lower bound analysis, namely τ is of the order $1/\sigma$. From (1.7), if we take $2/\pi$ for $C(h)$ then we take $\tau = 1/2\sigma$. As

$$\widehat{\psi}(y) = \begin{cases} \frac{1}{\sigma}(1 - \frac{|y|}{\sigma}) & \text{if } |y| \leqslant \sigma, \\ 0 & \text{otherwise}, \end{cases} \tag{3.18}$$

we have $\widehat{\psi}(0) = 1/\sigma$, $\psi(0) = 1$ and $\psi(1/2\sigma) = 4/\pi^2 \approx 0.405285$. Thus after some algebra (see (1.8) and the lines immediately following this) we see that the average number of normalized zeros in the interval $(-\frac{1}{2\sigma}, \frac{1}{2\sigma})$ is at most $\frac{\pi^2}{4}(r + \frac{1}{2} + \frac{1}{\sigma})$.

Appendix A. Standard conjectures

At various points in the paper we assume the following conjectures.

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

Birch and Swinnerton-Dyer conjecture. *(See [BS-D1,BS-D2].) Let E be an elliptic curve of geometric rank r over $\mathbb{Q}$ with Mordell–Weil group $E(\mathbb{Q}) = \mathbb{Z}^r \oplus \mathbb{T}$. Then the analytic rank (the order of vanishing of the completed L-function at the critical point) equals the geometric rank.*

Tate's conjecture for elliptic surfaces. *(See [Ta].) Let $\mathcal{E}/\mathbb{Q}$ be an elliptic surface and $L_2(\mathcal{E}, s)$ be the L-series attached to $H^2_{\acute{e}t}(\mathcal{E}/\overline{\mathbb{Q}}, \mathbb{Q}_l)$. $L_2(\mathcal{E}, s)$ has a meromorphic continuation to $\mathbb{C}$ and $-\mathrm{ord}_{s=1} L_2(\mathcal{E}, s) =$* 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* $\mathrm{Re}(s) = 1$.

Remark A.1. Tate's conjecture is known for rational elliptic surfaces. An elliptic surface $\mathcal{E} : y^2 = x^3 + A(T)x + B(T)$ 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 $\mathrm{ord}_{T=0} T^{12} \Delta(T^{-1}) = 0$.

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

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 Conjecture (see [Gr]):

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 $o(N)$ [Ho, Chapter 4].

We use the Square-Free Sieve to handle the variations in the conductors. If our evaluation of the logarithm 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.

The following conjecture is used only to interpret some of our results (unless we are calculating the 2-level density to distinguish the three orthogonal candidate groups).

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 can fail (there are families with constant $j(E_t)$ where all curves have the same sign, as well as more exotic examples). Helfgott [He] has related the Restricted Sign Conjecture to the Square-Free Sieve Conjecture and standard conjectures on sums of Moebius:

Polynomial Moebius conjecture. *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 shows the square-free sieve and polynomial Moebius imply the Restricted Sign Conjecture for many families; this is also discussed in [Mil1]. More precisely, let $M(t)$ be the product of the irreducible polynomials dividing $\Delta(t)$ and not $c_4(t)$.

Theorem *(Equidistribution of sign in a family). (See [He].) 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.*

Supplementary material

The online version of this article contains additional supplementary material. Please visit doi:10.1016/j.jnt.2010.04.002.

References

[BMSW] B. Bektemirov, B. Mazur, W. Stein, M. Watkins, Average ranks of elliptic curves: Tension between data and conjecture, Bull. Amer. Math. Soc. (N.S.) 44 (2007) 233–254.

[BS-D1] B. Birch, H. Swinnerton-Dyer, Notes on elliptic curves. I, J. Reine Angew. Math. 212 (1963) 7–25.

[BS-D2] B. Birch, H. Swinnerton-Dyer, Notes on elliptic curves. II, J. Reine Angew. Math. 218 (1965) 79–108.

[BCDT] C. Breuil, B. Conrad, F. Diamond, R. Taylor, On the modularity of elliptic curves over Q: wild 3-adic exercises, J. Amer. Math. Soc. 14 (4) (2001) 843–939.

[Bro] M.L. Brown, Heegner Modules and Elliptic Curves, Lecture Notes in Math., vol. 1849, Springer-Verlag, 2004.

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

[BM] A. Brumer, O. McGuinness, The behavior of the Mordell–Weil group of elliptic curves, Bull. Amer. Math. Soc. (N.S.) 23 (1990) 375–382.

[CW] J. Coates, A. Wiles, On the conjecture of Birch and Swinnerton-Dyer, Invent. Math. 39 (3) (1977) 223–251.

[CPRW] J.B. Conrey, A. Pokharel, M.O. Rubinstein, M. Watkins, Secondary terms in the number of vanishings of quadratic twists of elliptic curve L-functions, in: Ranks of Elliptic Curves and Random Matrix Theory, in: London Math. Soc. Lecture Note Ser., vol. 341, Cambridge University Press, Cambridge, 2007, pp. 215–232.

[DFK] C. David, J. Fearnley, H. Kisilevsky, On the vanishing of twisted L-functions of elliptic curves, Experiment. Math. 13 (2) (2004) 185–198.

[DHKMS1] E. Dueñez, D.K. Huynh, J. Keating, S.J. Miller, N. Snaith, The lowest eigenvalue in Jacobi ensembles and Painlevé VI, preprint.

[DHKMS2] E. Dueñez, D.K. Huynh, J. Keating, S.J. Miller, N. Snaith, Models for zeros at the central point in families of elliptic curves, in preparation.

[FM] F.W.K. Firk, S.J. Miller, Nuclei, primes and the random matrix connection, Symmetry 1 (2009) 64–105, doi:10.3390/sym1010064.

[Gao] P. Gao, N-level density of the low-lying zeros of quadratic Dirichlet L-functions, PhD thesis, University of Michigan, 2005.

[Go] D. Goldfeld, Conjectures on elliptic curves over quadratic fields, in: Number Theory, Carbondale, in: Lecture Notes in Math., vol. 751, Springer-Verlag, 1979, pp. 108–118.

[GM] F. Gouvêa, B. Mazur, The square-free sieve and the rank of elliptic curves, J. Amer. Math. Soc. 4 (1991) 1–23.

[Gr] A. Granville, ABC allows us to count squarefrees, Int. Math. Res. Not. 19 (1998) 991–1009.

[GKZ] B.H. Gross, W. Kohnen, D.B. Zagier, Heegner points and derivatives of L-series. II, Math. Ann. 278 (1–4) (1987) 497–562.

[GZ] B.H. Gross, D.B. Zagier, Heegner points and derivatives of L-series, Invent. Math. 84 (2) (1986) 225–320.

[H-B] D.R. Heath-Brown, The average rank of elliptic curves IV, Duke Math. J. 122 (3) (2004) 591–623.

[Hej] D. Hejhal, On the triple correlation of zeros of the zeta function, Int. Math. Res. Not. 7 (1994) 294–302.

[He] H.A. Helfgott, On the behaviour of root numbers in families of elliptic curves, preprint, http://arxiv.org/abs/math/0408141, 2004.

[Ho] C. Hooley, Applications of Sieve Methods to the Theory of Numbers, Cambridge University Press, Cambridge, 1976.

[HM] C. Hughes, S.J. Miller, Low-lying zeros of L-functions with orthogonal symmetry, Duke Math. J. 136 (1) (2007) 115–172.

[HR] C. Hughes, Z. Rudnick, Linear statistics of low-lying zeros of L-functions, Q. J. Math. 54 (2003) 309–333.

[Huy] D.K. Huynh, Elliptic curve L-functions of finite conductor and random matrix theory, PhD thesis, University of Bristol, 2009.

[KS1] N. Katz, P. Sarnak, Random Matrices, Frobenius Eigenvalues and Monodromy, Amer. Math. Soc. Colloq. Publ., vol. 45, Amer. Math. Soc., Providence, 1999.

[KS2] N. Katz, P. Sarnak, Zeros of zeta functions and symmetries, Bull. Amer. Math. Soc. (N.S.) 36 (1999) 1–26.

[Kn] A. Knapp, Elliptic Curves, Princeton University Press, Princeton, 1992.

[Kob] N. Koblitz, Introduction to Elliptic Curves and Modular Forms, Springer-Verlag, 1993.

[Kol1] V.A. Kolyvagin, The Mordell–Weil and Shafarevich–Tate groups for Weil elliptic curves, Izv. Akad. Nauk SSSR Ser. Mat. 52 (6) (1988) 1154–1180, 1327; translation in Math. USSR-Izv. 33 (3) (1989) 473–499.

[Kol2] V.A. Kolyvagin, Finiteness of $E(Q)$ and $Shah(E, Q)$ for a subclass of Weil curves, Izv. Akad. Nauk SSSR Ser. Mat. 52 (3) (1988) 522–540, 670–671; translation in Math. USSR-Izv. 32 (3) (1989) 523–541.

[Kow2] E. Kowalski, On the rank of quadratic twists of elliptic curves over function fields, Int. J. Number Theory (2006).

[Kow1] E. Kowalski, Elliptic curves, rank in families and random matrices, in: J.B. Conrey, D.W. Farmer, F. Mezzadri, N.C. Snaith (Eds.), Ranks of Elliptic Curves and Random Matrix Theory, in: London Math. Soc. Lecture Note Ser., vol. 341, 2007.

[Ma] B. Mazur, Rational isogenies of prime degree, Invent. Math. 44 (2) (1978) 129–162.

[Mes] J. Mestre, Formules explicites et minorations de conducteurs de variétés algébriques, Compos. Math. 58 (1986) 209–232.

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

[Mil1] S.J. Miller, 1- and 2-level densities for families of elliptic curves: Evidence for the underlying group symmetries, PhD thesis, Princeton University, 2002, http://www.williams.edu/go/math/sjmiller/public_html/math/thesis/SJMthesis_Rev2005.pdf.

[Mil2] S.J. Miller, 1- and 2-level densities for families of elliptic curves: Evidence for the underlying group symmetries, Compos. Math. 140 (4) (2004) 952–992.

[Mil3] S.J. Miller, Investigations of zeros near the central point of elliptic curve L-functions, Experiment. Math. 15 (3) (2006) 257–279, with an appendix by E. Dueñez.

[Mon] H. Montgomery, The pair correlation of zeros of the zeta function, in: Analytic Number Theory, in: Proc. Sympos. Pure Math., vol. 24, Amer. Math. Soc., Providence, 1973, pp. 181–193.

[RSi] M. Rosen, J. Silverman, On the rank of an elliptic surface, Invent. Math. 133 (1998) 43–67.

[Ru] K. Rubin, The one-variable main conjecture for elliptic curves with complex multiplication, in: L-Functions and Arithmetic, Durham, 1989, in: London Math. Soc. Lecture Note Ser., vol. 153, Cambridge University Press, Cambridge, 1991, pp. 353–371.

[RuSi] K. Rubin, A. Silverberg, Ranks of elliptic curves, Bull. Amer. Math. Soc. (N.S.) 39 (2002) 455–474.

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

[Sil1] J. Silverman, The Arithmetic of Elliptic Curves, Grad. Texts in Math., vol. 106, Springer-Verlag, New York, 1986.

[Sil2] J. Silverman, Advanced Topics in the Arithmetic of Elliptic Curves, Grad. Texts in Math., vol. 151, Springer-Verlag, Berlin, New York, 1994.

[Sil3] J. Silverman, The average rank of an algebraic family of elliptic curves, J. Reine Angew. Math. 504 (1998) 227–236.

[ST] J. Silverman, J. Tate, Rational Points on Elliptic Curves, Springer-Verlag, New York, 1992.

[Ta] J. Tate, Algebraic cycles and the pole of zeta functions, in: Arithmetical Algebraic Geometry, Harper and Row, New York, 1965, pp. 93–110.

[TW] R. Taylor, A. Wiles, Ring-theoretic properties of certain Hecke algebras, Ann. of Math. 141 (1995) 553–572.

[Wi] A. Wiles, Modular elliptic curves and Fermat's last theorem, Ann. of Math. 141 (1995) 443–551.

[Wis] J. Wishart, The generalized product moment distribution in samples from a normal multivariate population, Biometrika 20A (1928) 32–52.

[Yo1] M.P. Young, Basics of elliptic curves, talk at the American Institute of Mathematics, June 1, 2006, http://www.aimath.org/conferences/ntrmt/talks/BasicsofEllipticCurves.pdf.

[Yo2] M.P. Young, Low-lying zeros of families of elliptic curves, J. Amer. Math. Soc. 19 (1) (2006) 205–250.

[ZK] D. Zagier, G. Kramarz, Numerical investigations related to the L-series of certain elliptic curves, J. Indian Math. Soc. 52 (1987) 51–69.