SOME HEURISTICS ABOUT ELLIPTIC CURVES

MARK WATKINS

Abstract. We give some heuristics for counting elliptic curves with certain properties. In particular, we re-derive the Brumer-McGuinness heuristic for the number of curves with positive/negative discriminant up to X, which is an application of lattice-point counting. We then introduce heuristics that allow us to predict how often we expect an elliptic curve E with even parity to have $L(E,1)=0$. We find that we expect there to be about $c_1 X^{19/24}(\log X)^{3/8}$ curves with $|\Delta|<X$ with even parity and positive (analytic) rank; since Brumer and McGuinness predict $cX^{5/6}$ total curves, this implies that asymptotically almost all even parity curves have rank 0. We then derive similar estimates for ordering by conductor, and conclude by giving various data regarding our heuristics and related questions.

1. Introduction

We give some heuristics for counting elliptic curves with certain properties. In particular, we re-derive the Brumer-McGuinness heuristic for the number of curves with positive/negative discriminant up to X, which is an application of lattice-point counting. We then introduce heuristics (with refinements from random matrix theory) that allow us to predict how often we expect an elliptic curve E with even parity to have $L(E,1)=0$. It turns out that we roughly expect that a curve with even parity has $L(E,1)=0$ with probability proportional to the square root of its real period, and, since we have an upper bound of size $1/\Delta^{1/12}$ on the real period, this leads us to the prediction that almost all curves with even parity should have $L(E,1)\neq 0$. By the conjecture of Birch and Swinnerton-Dyer, this says that almost all such curves have rank 0.

We then make similar heuristics when enumerating by conductor. The first task here is simply to count curves with conductor up to X, and for this we use heuristics involving how often large powers of primes divide the discriminant. Upon making this estimate, we are then able to imitate the argument we made previously, and thus derive an asymptotic for the number of curves with even parity and $L(E,1)=0$ under the ordering by conductor. We again get the heuristic that almost all curves with even parity should have $L(E,1)\neq 0$.

We then make a few remarks regarding how often curves should have nontrivial isogenies and/or torsion under different orderings, and then present some data regarding average ranks and the proportion of rank 2 curves. In particular, we give new evidence that the proportion of rank 2 curves goes to zero; this involves a careful "random" sampling of curves whose conductor is larger than previously considered, and we require an analysis of the variation of the real period to ensure that our sample is not overly biased. We conclude by giving data for the Mordell-Weil lattice distribution of rank 2 curves, and speculating about symmetric power L-functions.

2. The Brumer-McGuinness Heuristic

First we re-derive the Brumer-McGuinness heuristic [7] for the number of elliptic curves whose absolute discriminant is less than a given bound X; the technique here is essentially lattice-point counting, and we derive our estimates via the assumption that these counts are well-approximated by the areas of the given regions.

Conjecture 2.1. *[Brumer-McGuinness] The number $A_\pm(X)$ of elliptic curves over $\mathbf{Q}$ whose minimal (integral) discriminant has absolute value less than X is asymptotically given by (splitting into positive and negative discriminant)*

$$A_\pm(X) \sim \frac{\alpha_\pm}{\zeta(10)} X^{5/6} \quad \textit{where} \quad \alpha_\pm = \frac{\sqrt{3}}{10}\int_{\pm 1}^{\infty} \frac{dx}{\sqrt{x^3 \mp 1}}.$$

As indicated by Brumer and McGuinness, the identity $\alpha_- = \sqrt{3}\alpha_+$ was already known to Legendre, and is related to complex multiplication. These constants can be expressed in terms of Beta integrals

$$B(u,v) = \int_0^1 x^{u-1}(1-x)^{v-1}\,dx = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)}$$

as $\alpha_+ = \frac{1}{3}B(1/2,1/6)$ and $\alpha_- = B(1/2,1/3)$.

Recall that every elliptic curve over $\mathbf{Q}$ has a unique integral minimal model

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \quad \text{with} \quad a_1, a_3 \in \{0,1\} \quad \text{and} \quad |a_2| \le 1.$$

Fix one of the 12 choices of (a_1, a_2, a_3). Since these are all bounded by 1 the discriminant is thus approximately $-64a_4^3 - 432a_6^2$. So we essentially wish to count the number of (a_4, a_6)-lattice-points with $|64a_4^3 + 432a_6^2| \le X$, noting that Brumer and McGuinness divide the curves according to the sign of the discriminant. The lattice-point count for $a_1 = a_2 = a_3 = 0$ is given by

$$\sum\sum_{0<-64a_4^3-432a_6^2<X} 1 \quad + \sum\sum_{-X<-64a_4^3-432a_6^2<0} 1.$$

We estimate this lattice-point count by the integral $\int\int_U du_4du_6$ for the region U given by $|64u_4^3 + 432u_6^2| < X$. After splitting into two parts based upon the sign of the discriminant and performing the u_4-integration, we get

$$\frac{2}{(64)^{1/3}}\int_0^\infty \Big[(-432u_6^2)^{1/3} - (-X - 432u_6^2)^{1/3}\Big]\,du_6 + \\ + \frac{2}{(64)^{1/3}}\int_0^\infty \Big[(X - 432u_6^2)^{1/3} - (-432u_6^2)^{1/3}\Big]\,du_6,$$

where the factor of 2 comes from the sign of u_6. Changing variables $u_6 = w\sqrt{X/432}$ and multiplying by 12 for the number of (a_1, a_2, a_3)-choices we get

$$\frac{24}{(64)^{1/3}}\frac{X^{5/6}}{\sqrt{432}}\int_0^\infty \Big[(w^2+1)^{1/3} - (w^2)^{1/3}\Big]\,dw + \\ + \frac{24}{(64)^{1/3}}\frac{X^{5/6}}{\sqrt{432}}\int_0^\infty \Big[(w^2)^{1/3} - (w^2-1)^{1/3}\Big]\,dw.$$

These integrals are probably known, but I am unable to find a reference. The integrals respectively simplify[1] to

$$\frac{3}{5}\int_1^\infty \frac{dx}{\sqrt{x^3-1}} = \frac{1}{5}B(1/2,1/6) \quad \text{and} \quad \frac{3}{5}\int_{-1}^\infty \frac{dx}{\sqrt{x^3+1}} = \frac{3}{5}B(1/2,1/3).$$

This counts all models of curves; if we eliminate non-minimal models, for which we have $p^4|c_4$ and $p^{12}|\Delta$ for some prime p, we expect to accrue an extra factor[2] of $1/\zeta(10)$. From this, we get the conjecture of Brumer and McGuinness of above.

3. Counting curves of even parity whose central L-value vanishes.

See [51, §15-16] for the definitions of the conductor N and L-function $L(E,s)$ of an elliptic curve E. Due to work of Wiles [60, 55] and others [20, 12, 4], we know that elliptic curves over $\mathbf{Q}$ are modular, and this implies that the completed L-function $\Lambda(E,s) = \Gamma(s)(\sqrt{N}/2\pi)^s L(E,s)$ extends to an entire function and satisfies the functional equation $\Lambda(E,s) = \pm\Lambda(E,2-s)$. When the plus sign occurs, we say that E has even parity.

We now try to count elliptic curves E with even parity for which $L(E,1)=0$. Throughout this section, E shall be a curve with even parity, and we shall order curves by discriminant. Via the conjectural 50-50 Principle, we expect that, under any reasonable ordering, half of the elliptic curves should have even parity.[3] In particular, we predict that there are asymptotically $A_\pm(X)/2$ curves with even parity and positive/negative discriminant up to X.

Our main tool shall be random matrix theory, which gives a heuristic for predicting how often $L(E,1)$ is small. We could alternatively derive a cruder heuristic by assuming the order of the Shafarevich-Tate group is a random square integer in a given interval, but random matrix theory has the advantage of being able to predict a more explicit asymptotic. Our principal heuristic is the following:

Heuristic 3.1. *The number $R(X)$ of elliptic curves $E/\mathbf{Q}$ with even parity and $L(E,1)=0$ and minimal absolute discriminant less than X is given asymptotically by $R(X) \sim cX^{19/24}(\log X)^{3/8}$ for some constant $c>0$.*

In particular, note that we get the prediction that almost all curves with even parity have $L(E,1) \neq 0$ under this ordering.

3.1. Random matrix theory.

Originally developed in mathematical statistics by Wishart [61] in the 1920s and then in mathematical physics (especially the spectra of highly excited nuclei) by Wigner [59], Dyson, Mehta, and others (particularly [38]), random matrix theory [40] has now found some applications in number theory, the earliest being the oft-told story of Dyson's remark to Montgomery regarding the pair-correlation of zeros of the Riemann ζ-function. Based on substantial numerical evidence, random matrix theory appears to give reasonable models for the distribution of L-values in families, though the issue of what constitutes a proper family is

---

[1] As N. D. Elkies indicated to us, we can write $I(a) = \int_0^\infty \left[(t^2+a)^{1/3} - (t^2)^{1/3}\right] dt$, differentiate under the integral sign, then substitute $t^2 + a = ax^3$, and finally re-integrate to obtain $I(1)$.

[2] Note that some choices of (a_1, a_2, a_3) necessarily have odd discriminant, but the other choices compensate to give the proper Euler factors at 2 (and 3). A more direct way of getting the $1/\zeta(10)$ factor is to note that non-minimality at p occurs when c_4/p^4 and c_6/p^6 satisfy the Connell congruences [11] we mention below. A referee points out that Brumer [6, Lemma 4.3] obtains this same $1/\zeta(10)$ factor via sieving.

[3] A referee says that a preprint of Helfgott [30] discusses this question for parametrised families.

a delicate one (see particularly [13, §3], where the notion of family comes from the ability to calculate moments of L-functions rather than from algebraic geometry).

The family of quadratic twists of a given elliptic curve $E : y^2 = x^3 + Ax + B$ is given by $E_d : y^2 = x^3 + Ad^2x + Bd^3$ for squarefree d. The work (most significantly a monodromy computation) of Katz and Sarnak [33] regarding families of curves over function fields implies that when we restrict to quadratic twists with even parity, we should expect that the L-functions are modelled by random matrices with even orthogonal symmetry. This means that local statistics involving the distribution of spacings of zeros of the L-functions should be the same as the statistics concerning the distribution of the spacings of the eigenangles of matrices taken randomly from SO(even) with respect to Haar measure. Furthermore, the distribution of the special values of such L-functions should be related to the distribution of the evaluations at 1 of the characteristic polynomials of such matrices. An argument based upon frequency of small L-values and discretisation (similar to the below) then gives that the number of d with E_d of even parity and $|d| < D$ and $L(E_d, 1) = 0$ is given by $c_E D^{3/4}(\log D)^{b_E}$ where b_E takes on four possible values (see [19]) depending on the splitting behaviour of the cubic polynomial $x^3 + Ax + B$, while c_E has yet to be determined explicitly. Various data have been given by Rubinstein [15] to lend credence to this guess. It can be noted that the 3/4-exponent can already be suspected from work of Waldspurger [57], which relates[4] to E a modular form of weight 3/2 whose dth coefficient $c(d)$ is such that $c(d)^2$ proportional to $L(E_d, 1)$. In particular, the Ramanujan conjecture predicts that the dth coefficient is bounded by $d^{1/4+\epsilon}$, and so, assuming a reasonable distribution, the probability that it is zero is about 1 in $d^{1/4}$ — summing over d up to D then gives the crude heuristic (possibly due to Sarnak).

Though we have no exact function field analogue when considering the set of all elliptic curves of even parity, we brazenly assume (largely from looking at the sign in the functional equation) that the symmetry type is again orthogonal with even parity.[5] What this means is that we want to model properties of the L-function via random matrices taken from SO($2M$) with respect to Haar measure. Here we wish the mean density of zeros of the L-functions to match the mean density of eigenvalues of our matrices, and so, as in [34], we should take $2M \approx 2 \log N$. We suspect that the L-value distribution is approximately given by the distribution of the evaluations at 1 of the characteristic polynomials of our random matrices. At the crude level, this distribution is determined entirely by the symmetry type, while finer considerations are distinguished via arithmetic considerations.

With this assumption, via the moment conjectures of [34] and then using Mellin inversion, as $t \to 0$ we have (see (21) of [14]) that (here 3/8 is $\binom{-1/2}{2}$)

$$\text{Prob}[L(E, 1) \le t] \sim \alpha(E)t^{1/2}M^{3/8}. \tag{1}$$

This heuristic is stated for fixed $M \approx \log N$, but we shall also allow $M \to \infty$. It is not easy to understand this probability, as both the constant $\alpha(E)$ and the matrix-size M depend on E. We can take curves with $e^M \le N \le e^{M+1}$ to mollify

[4]We give a simplified statement here, not worrying about conditions regarding whether d is a square modulo $4N$ and the sign of d.

[5]In was pointed out to us by E. Kowalski that Katz [32, §12.8] has some results in the function field case when considering (say) generalised or usual Weierstrass families of elliptic curves. A referee also notes that there is also work of Miller [42] and Young [62] in this direction.

the impact of the conductor, but in order to average over a set of curves, we need to understand how $\alpha(E)$ varies. One idea is that $\alpha(E)$ separates into two parts, one of which depends on local structure (Frobenius traces) of the curve, and the other of which depends only upon the size of the conductor N. Letting G be the Barnes G-function (such that $G(z+1) = \Gamma(z)G(z)$ with $G(1) = 1$) and $M = \lfloor \log N \rfloor$ we have that

$$\alpha(E) = \alpha_R(M) \cdot \alpha_A(E)$$

with $\alpha_R(M) \to \hat{\alpha}_R = 2^{1/8}G(1/2)\pi^{-1/4}$ as $M \to \infty$ and

(2)

$$\alpha_A(E) = \prod_p F(p) = \prod_p \left(1 - \frac{1}{p}\right)^{3/8} \left(\frac{p}{p+1}\right) \left(\frac{1}{p} + \frac{L_p(1/p)^{-1/2}}{2} + \frac{L_p(-1/p)^{-1/2}}{2}\right)$$

where $L_p(X) = (1 - a_pX + pX^2)^{-1}$ when $p \nmid \Delta$ and $L_p(X) = (1 - a_pX)^{-1}$ otherwise; see (10) of [14] evaluated at $k = -1/2$, though that equation is wrong at primes that divide the discriminant — see (20) of [16], where Q should be taken to be 1. Note that the Sato-Tate conjecture [54] implies that a_p^2 is p on average, and this implies that the above Euler product converges.[6]

3.2. **Discretisation of the L-value distribution.** For precise definitions of the Tamagawa numbers, torsion group, periods, and Shafarevich-Tate group, see [51], though below we give a brief description of some of these. We let $\tau_p(E)$ be the Tamagawa number of E at the (possibly infinite) prime p, and write $\tau(E) = \prod_p \tau_p(E)$ for the Tamagawa product and $T(E)$ for the size of the torsion group. We also write $\Omega_{\rm re}(E)$ for the real period, and $S(E)$ for the size of the Shafarevich-Tate group when $L(E,1) \neq 0$, with $S(E) = 0$ when $L(E,1) = 0$.

We wish to assert that sufficiently small values of $L(E,1)$ actually correspond to $L(E,1) = 0$. We do this via the conjectural formula of Birch and Swinnerton-Dyer [2], which asserts that

$$L(E,1) = \Omega_{\rm re}(E) \cdot \frac{\tau(E)}{T(E)^2} \cdot S(E).$$

Our discretisation[7] will be that

$$L(E,1) < \Omega_{\rm re}(E) \cdot \frac{\tau(E)}{T(E)^2} \quad \text{implies} \quad L(E,1) = 0.$$

Note that we are only using that $S(E)$ takes on integral values, and do not use the (conjectural) fact that it is square.

Using (1), we estimate the number of curves with positive (for simplicity) discriminant less than X and even parity and $L(E,1) = 0$ via the lattice-point sum

$$W(X) = \sum_{\substack{c_4,\, c_6 \\ 0<c_4^3-c_6^2<1728X}} \sum_{\text{minimal}} \alpha_R(M)\alpha_A(E) \cdot \sqrt{\frac{\Omega_{\rm re}(E)\tau(E)}{T(E)^2}} \cdot M^{3/8}.$$

[6] A referee points out that Birch [1] has already shown (using the Selberg trace formula) that the Sato-Tate distribution holds when considering all curves over a fixed field $\mathbf{F}_p$, while Michel [41] has the best results in the case of one-parameter families.

[7] The precision of this discretisation might be the most debatable methodology we use. Indeed, we are essentially taking a "sharp cutoff", while it might be better to have a more smooth transition function. For this reason, we do not specify the leading constant in our final heuristic.

We need to introduce congruence conditions on c_4 and c_6 to make sure that they correspond to a minimal model of an elliptic curve. The paper [53] uses the work of Connell [11] in a different context to get that there are 288 classes of $(c_4 \bmod 576, c_6 \bmod 1728)$ that can give minimal models, and so we get a factor of $288/(576 \cdot 1728)$, assuming that each congruence class has the same impact on all the entities in the sum. Indeed, this independence (on average) of various quantities with respect to c_4 and c_6 is critical in our estimation of $W(X)$. There is also the question of non-minimal models, from which (as in the Brumer-McGuinness heuristic) we get a factor of $1/\zeta(10)$.

Guess 3.2. *The lattice-point sum $W(X)$ can be approximated as $X \to \infty$ by*

$$\hat{W}(X) = \frac{288}{(576 \cdot 1728)} \frac{1}{\zeta(10)} \cdot \hat{\alpha}_R \bar{\alpha}_A \beta(\sqrt{\tau}) \cdot \iint\limits_{1 \le \frac{u_4^3 - u_6^2}{1728} < X} \Omega_{\mathrm{re}}(E)^{1/2} \cdot (\log \Delta)^{3/8} du_4 du_6.$$

Here $\hat{\alpha}_R$ is the limit $2^{1/8}G(1/2)\pi^{-1/4}$ of $\alpha_R(M)$ as $M \to \infty$, while $\bar{\alpha}_A$ is a suitable average of the arithmetic factors $\alpha_A(E)$, and $\beta(\sqrt{\tau})$ is the average of the square root of the Tamagawa product. We have also approximated $\log N \approx \log \Delta$ and assumed the torsion is trivial; below we will give these heuristic justification (on average). Note that everything left in the integral is a smooth function of u_4 and u_6.

We shall first evaluate the integral in $\hat{W}(X)$ given these suppositions, and then try to justify the various assumptions that are inherent in this guess.[8] For convenience, we try to list all the heuristic assumptions we have made.

- Lattice-point sums are well-approximated by areal integrals.
- We have $\mathrm{Prob}[L(E,1) < t] \sim \alpha(E) t^{1/2} (\log N)^{3/8}$ via random matrix theory.
- We have $S(E) = \frac{L(E,1)}{\Omega_{\mathrm{re}}(E)} \frac{T(E)^2}{\tau(E)}$ by the BSD conjecture.
- There is independence amongst the arithmetic factor $\alpha_A(E)$, the Tamagawa products, and the real period.
- We can replace $\log N$ by $\log \Delta$, and torsion can be ignored.

3.3. **Evaluation of the integral.** Write E as $y^2 = 4x^3 - (c_4/12)x - c_6/216$, and put $e_1 > e_2 > e_3$ for the roots of the cubic polynomial on the right side. We have

$$1/\Omega_{\mathrm{re}} = \mathrm{agm}\big(\sqrt{e_1 - e_2}, \sqrt{e_1 - e_3}\big)/\pi.$$

We also have that $(e_1 - e_2)(e_1 - e_3)(e_2 - e_3) = \sqrt{\Delta/16}$ from the formula for the discriminant. We next write $e_1 - e_2 = \Delta^{1/6}\lambda$ and $e_2 - e_3 = \Delta^{1/6}\mu$ so that we have $\mu\lambda(\lambda+\mu) = 1/4$, while $e_1 = \frac{\Delta^{1/6}}{3}(\mu+2\lambda)$, $e_2 = \frac{\Delta^{1/6}}{3}(\mu-\lambda)$, and $e_3 = -\frac{\Delta^{1/6}}{3}(2\mu+\lambda)$. Thus we get

$$-c_6/864 = -e_1e_2e_3 = \frac{\Delta^{1/2}}{27}(\mu + 2\lambda)(\mu - \lambda)(2\mu + \lambda)$$

and

$$-c_4/48 = e_1e_2 + e_1e_3 + e_2e_3 = -\frac{\Delta^{1/3}}{3}(\mu^2 + \lambda\mu + \lambda^2).$$

[8] Note that our methods do not readily generalise to positive rank, as there is no apparent way to model the heights of points (and thus the regulator). A referee points out that Lang [36] gives some bounds, and perhaps suggests a distribution, but this seems insufficient for our purposes.

Changing variables in the $\hat{W}$-integral gives a Jacobian of $432/\Delta^{1/6}\sqrt{\mu^4+\mu}$ so that

$$\hat{W}(X) = \tilde{c}\int_1^X \int_0^\infty \frac{(\log\Delta)^{3/8}}{\sqrt{\Delta^{1/12}\,\mathrm{agm}(\sqrt{\lambda},\sqrt{\lambda+\mu})}}\frac{d\mu\, d\Delta}{\Delta^{1/6}\sqrt{\mu^4+\mu}},$$

for some constant $\tilde{c} > 0$, where $\lambda = (\sqrt{\mu^4+\mu} - \mu^2)/2\mu$. Thus the variables are nicely separated, and since the μ-integral converges, we do indeed get the asymptotic $\hat{W}(X) \sim cX^{19/24}(\log X)^{3/8}$. A similar argument can be given for curves with negative discriminant. This concludes our derivation of Heuristic 3.1, and now we turn to giving some reasons for our expectation that the arithmetic factors can be mollified by taking their averages.

3.4. **Expectations for arithmetic factors on average.** In the next section we shall explain (among other things) why we expect that $\log N \approx \log\Delta$ for almost all curves, and in section 5, we shall recall the classical parametrisations of $X_1(N)$ due to Fricke to indicate why we expect that the torsion size is trivial outside a sparse set of curves. Here we show how to compute the various averages (with respect to ordering by discriminant) of the square root of the Tamagawa product and the arithmetic factors $\alpha_A(E)$.

For both heuristics, we make the assumption that curves satisfying the discriminant bound $|\Delta| \le X$ behave essentially the same as those that satisfy $|c_4| \le X^{1/3}$ and $|c_6| \le X^{1/2}$. That is, we approximate our region by a big box. We write D for the absolute value of Δ, and consider how often high powers of primes divide D.

3.4.1. *Primes dividing the discriminant.* We wish to know how often a prime divides the discriminant to a high power. Fix a prime $p \ge 5$ with p a lot smaller than $X^{1/3}$. We estimate the probability that $p^k|\Delta$ by considering all p^{2k} choices of c_4 and c_6 modulo p^k, that is, we count the number of solutions $C(p^k)$ to the congruence $c_4^3 - c_6^2 = 1728\Delta \equiv 0 \pmod{p^k}$. This auxiliary curve $c_4^3 = c_6^2$ is singular at $(0,0)$ over $\mathbf{F}_p$, and has $(p-1)$ non-singular $\mathbf{F}_p$-solutions which lift to $p^{k-1}(p-1)$ points modulo p^k.

For p^k sufficiently small, our (c_4, c_6)-region is so large that we can show that the probability that $p^k|\Delta$ is $C(p^k)/p^{2k}$. We assume that big primes act (on average) in the same manner, while a similar heuristic can be given for $p = 2, 3$. Curves with $p^4|c_4$ and $p^6|c_6$ will not be given by their minimal model; indeed, we want to exclude these curves, and thus will multiply our probabilities by $\kappa_p = (1-1/p^{10})^{-1}$ to make them conditional on this criterion. For instance, the above counting of points says that there is a probability of $(p^2-p)/p^2$ that $p \nmid D$, and so upon conditioning upon minimal models we get $\kappa_p(1-1/p)$ for this probability.

What is the probability $P_m(p,k)$ that a curve given by a minimal model has multiplicative reduction at $p \ge 5$ and $p^k \| D$ for some $k > 0$? In terms of Kodaira symbols, this is the case of I_k. For multiplicative reduction we need that $p \nmid c_4$ and $p \nmid c_6$. These events are assumed independent and each has a probability $(1-1/p)$ of occurring. Upon assuming these conditions and working modulo p^k, there are $(p^k - p^{k-1})$ such choices for both c_4 and c_6, and of the resulting (c_4, c_6) pairs we noted above that $p^{k-1}(p-1)$ of them have $p^k|D$. So, given a curve with $p \nmid c_4$ and $p \nmid c_6$, we have a probability of $1/p^{k-1}(p-1)$ that $p^k|D$, which gives $1/p^k$ for the probability that $p^k\|D$. In symbols, we have that (for $p \ge 5$ and $k \ge 1$)

$$\mathrm{Prob}\Big[p^k\|(c_4^3 - c_6^2)\ \Big|\ p \nmid c_4, p \nmid c_6\Big] = 1/p^k.$$

Including the conditional probability for minimal models, we get

$$P_m(p,k) = (1-1/p^{10})^{-1}(1-1/p)^2/p^k \quad (\text{for } p \geq 5 \text{ and } k \geq 1). \tag{3}$$

Note that summing this over $k \geq 1$ gives $\kappa_p(1-1/p)/p$ for the probability for an elliptic curve to have multiplicative reduction at p.

What is the probability $P_a(p,k)$ that a curve given by a minimal model has additive reduction at $p \geq 5$ and $p^k \| D$ for some $k > 0$? We shall temporarily ignore the factor of $\kappa_p = (1-1/p^{10})^{-1}$ from non-minimal models and include it at the end. We must have that $p|c_4$ and $p|c_6$, and thus get that $k \geq 2$. For $k = 2,3,4$, which correspond to Kodaira symbols II, III, and IV respectively, the computation is not too bad: we get that $p^2\|D$ exactly when $p|c_4$ and $p\|c_6$, so that the probability is $(1/p)\cdot(1-1/p)/p = (1-1/p)/p^2$; for $p^3\|D$ we need $p\|c_4$ and $p^2|c_6$ and thus get $(1-1/p)/p\cdot(1/p^2) = (1-1/p)/p^3$ for the probability; and for $p^4\|D$ we need $p^2|c_4$ and $p^2\|c_6$ and so get $(1/p^2)\cdot(1-1/p)/p^2 = (1-1/p)/p^4$ for the probability. Note that the case $k = 5$ cannot occur. Thus we have (for $p \geq 5$) the formula $P_a(p,k) = (1-1/p^{10})^{-1}(1-1/p)/p^k$ for $k = 2,3,4$.

More complications occur for $k \geq 6$, where now we split into two cases depending upon whether additive reduction persists upon taking the quadratic twist by p. This occurs when $p^3|c_4$ and $p^4|c_6$, and we denote by $P_a^n(p,k)$ the probability that $p^k\|D$ in this subcase. Just as above, we get that $P_a^n(p,k) = (1-1/p^{10})^{-1}(1-1/p)/p^{k-1}$ for $k = 8,9,10$. These are respectively the cases of Kodaira symbols IV*, III*, and II*. For $k = 11$ we have $P_a^n(p,k) = 0$, while for $k \geq 12$ our condition of minimality implies that we should take $P_a^n(p,k) = 0$.

We denote by $P_a^t(p,k)$ the probability that $p^6|D$ with either $p^2\|c_4$ or $p^3\|c_6$. First we consider curves for which $p^7|D$, and these have multiplicative reduction at p upon twisting. In particular, these curves have $p^2\|c_4$ and $p^3\|c_6$, and the probability of this is $(1-1/p)/p^2\cdot(1-1/p)/p^3$. Consider $k \geq 7$. We then take c_4/p^2 and c_6/p^3 both modulo p^{k-6}, and get that $p^{k-6}\|(D/p^6)$ with probability $1/p^{k-6}$ in analogy with the above. So we get that $P_a^t(p,k) = (1-1/p^{10})^{-1}(1-1/p)^2/p^{k-1}$ for $k \geq 7$. This corresponds to the case of $\mathrm{I}_{k-6}^\star$.

Finally, for $p^6\|D$ (which is the case $\mathrm{I}_0^\star$) we get a probability of $(1/p^2)\cdot(1/p^3)$ for the chance that $p^2|c_4$ and $p^3|c_6$, and, since there are p points mod p on the auxiliary curve $(c_4/p^2)^3 \equiv (c_6/p^3)^2 \pmod p$, a conditional probability of $(p^2-p)/p^2$ that $p^6\|D$. So we get that $P_a^t(p,6) = (1-1/p^{10})^{-1}(1-1/p)/p^5$.

We now impose our current notation on the previous paragraphs, and naturally let $P_a^t(p,k) = 0$ and $P_a^n(p,k) = P_a(p,k)$ for $k \leq 5$. Our final result is that

$$P_a^n(p,k) = \begin{cases} (1-1/p^{10})^{-1}(1-1/p)/p^k & k = 2,3,4 \\ (1-1/p^{10})^{-1}(1-1/p)/p^{k-1} & k = 8,9,10 \end{cases} \tag{4}$$

$$P_a^t(p,k) = \begin{cases} (1-1/p^{10})^{-1}(1-1/p)/p^5 & k = 6 \\ (1-1/p^{10})^{-1}(1-1/p)^2/p^{k-1} & k \geq 7 \end{cases} \tag{5}$$

with $P_a^n(p,k)$ and $P_a^t(p,k)$ equal to zero for other k. We conclude by defining $P_0(p,k)$ to be zero for $k > 0$ and to be the probability $(1-1/p^{10})^{-1}(1-1/p)$ that $p \nmid D$ for $k = 0$. We can easily check that we really do have the required probability relation $\sum_{k=0}^{\infty}\bigl[P_m(p,k) + P_a^n(p,k) + P_a^t(p,k) + P_0(p,k)\bigr] = 1$, as: the cases of multiplicative reduction give $\kappa_p(1-1/p)/p$; the cases of Kodaira symbols II, III, and IV give $\kappa_p(1/p^2 - 1/p^5)$; the cases of Kodaira symbols IV*, III*, and II*

give $\kappa_p(1/p^7 - 1/p^{10})$; the cases of $\mathrm{I}_k^\star$ summed for $k \ge 1$ give $\kappa_p(1-1/p)/p^6$; the case of $\mathrm{I}_0^\star$ gives $\kappa_p(1-1/p)/p^5$; and the sum of these with $P_0(p,0) = \kappa_p(1-1/p)$ does indeed give us 1. We could do a similar (more tedious) analysis for $p = 2, 3$, but this would obscure our argument.

The heuristics we used in deriving these probabilities were that the curves with $|\Delta| \le X$ act like those in a big box with $|c_4| \le X^{1/3}$ and $|c_6| \le X^{1/2}$, and that the effect of large primes dividing the discriminant can be estimated in a similar manner as with the small primes.

3.4.2. *Tamagawa averages.* Given a curve of absolute discriminant D, we can now compute the expectation for its Tamagawa number. We consider primes $p|D$ with $p \ge 5$, and compute the local Tamagawa number $t(p)$; this can be done as in Algorithm 7.5.1 of [10] (with a corrected line 2 of step 3 in early printings).

When E has multiplicative reduction at p and $p^k \| D$, then $t(p) = k$ if $-c_6$ is square mod p, and else $t(p) = 1, 2$ depending upon whether k is odd or even. So the average of $\sqrt{t(p)}$ for this case is $\epsilon_m(k) = \frac{1}{2}(1+\sqrt{k}), \frac{1}{2}(\sqrt{2}+\sqrt{k})$ for k odd/even respectively.

When E has potentially multiplicative reduction at p with $p^k \| D$, for k odd we have $t(p) = 4, 2$ depending on whether $(c_6/p^3) \cdot (\Delta/p^k)$ is square mod p, and for k even we have $t(p) = 4, 2$ depending on whether Δ/p^k is square mod p. In both cases the average of $\sqrt{t(p)}$ is $\frac{1}{2}(\sqrt{2}+\sqrt{4})$. In the case of $\mathrm{I}_0^\star$ reduction where we have $p^6 \| D$, we have that $t(p) = 1, 2, 4$ corresponding to whether the cubic $x^3 - (27c_4/p^2)x - (54c_6/p^3)$ has $0, 1, 3$ roots modulo p. So the average of $\sqrt{t(p)}$ is

$$\frac{\sqrt{1}\big((p-1)(p+1)/3\big) + \sqrt{2}\big(p(p-1)/2\big) + \sqrt{4}\big((p-1)(p-2)/6\big)}{\big((p-1)(p+1)/3\big) + \big(p(p-1)/2\big) + \big((p-1)(p-2)/6\big)} = \frac{2}{3} + \frac{\sqrt{2}}{2} - \frac{1}{3p}.$$

in this case.

For the remaining cases, when $p^2 \| D$ or $p^{10} \| D$ we have $t(p) = 1$, while when $p^3 \| D$ or $p^9 \| D$ we have $t(p) = 2$. Finally, when $p^4 \| D$ we have $t(p) = 3, 1$ depending on whether $-6c_6/p^2$ is square mod p, and similarly when $p^8 \| D$ we have $t(p) = 3, 1$ depending on whether $-6c_6/p^4$ is square mod p, so that the average of $\sqrt{t(p)}$ in both cases is $\frac{1}{2}(1+\sqrt{3})$. We get that $\epsilon_a^n(k) = 1, \sqrt{2}, \frac{1}{2}(1+\sqrt{3}), \frac{1}{2}(1+\sqrt{3}), \sqrt{2}, 1$ for $k = 2, 3, 4, 8, 9, 10$, while

$$(6) \quad \epsilon_m(k) = \begin{cases} \frac{1}{2}(1+\sqrt{k}), & k \text{ odd} \\ \frac{1}{2}(\sqrt{2}+\sqrt{k}), & k \text{ even} \end{cases} \quad \text{and} \quad \epsilon_a^t(p,k) = \begin{cases} \frac{2}{3} + \frac{\sqrt{2}}{2} - \frac{1}{3p}, & k = 6 \\ \frac{1}{2}(\sqrt{2}+\sqrt{4}), & k \ge 7 \end{cases}$$

with $\epsilon_a^n(k)$ and $\epsilon_a^t(p,k)$ equal to zero for other k.

We define the expected square root of the Tamagawa number $K(p)$ at p by

$$(7) \quad K(p) = \sum_{k=0}^{\infty} \big[\epsilon_m(k) P_m(p,k) + \epsilon_a^n(k) P_a^n(p,k) + \epsilon_a^t(p,k) P_a^t(p,k) + P_0(p,k)\big]$$

and assume that all the primes act independently to get that the expected global[9] Tamagawa number is $\beta_\tau = \prod_p K(p)$. The convergence of this product follows from an analysis of the dominant $k = 0, 1, 2$ terms of (7), which gives a behaviour of $1 + O(1/p^2)$. So we get that the Tamagawa product is a constant on average,

[9]Note that the Tamagawa number at infinity is 1 when E has negative discriminant and else is 2, the former occurring approximately $\sqrt{3}/(1+\sqrt{3}) \approx 63.4\%$ of the time.

which we do not bother to compute explicitly (we would need to consider $p = 2, 3$ more carefully to get a precise value).

3.4.3. *Arithmetic averages.* To compute the average value of $\alpha_A(E) = \prod_p F(p)$ in (2) we similarly assume[10] that each prime acts independently; we then compute the average value for each prime by calculating the distribution of $F(p)$ when considering all the curves modulo p (including those with singular reduction, and again making the slight adjustment for non-minimal models). This gives some constant for the average $\bar{\alpha}_A$ of $\alpha_A(E)$, which we again do not compute explicitly. Note that $\prod_p F(p)$ converges if we assume the Sato-Tate conjecture [54] since then we have that a_p^2 is p on average.

4. Relation between conductor and discriminant

We now give heuristics for how often we expect the ratio between the absolute discriminant and the conductor to be large. The main heuristic we derive is:

Heuristic 4.1. *The number $B(X)$ of elliptic curves over $\mathbf{Q}$ whose conductor is less than X satisfies $B(X) \sim cX^{5/6}$ for some explicit constant $c > 0$.*

Remark 4.2. It must be noted that the data of Cremona [18] do not coincide with this heuristic — in fact, the growth seems almost linear in the conductor, for taking the curves with conductor 40000-130000 in his database and doing a log-log regression yields a best-fit exponent of 0.98, which is much closer to 1 than to 5/6. An upper bound of $B(X) \ll_\epsilon X^{1+\epsilon}$ is explicated in [23, §3.1].

Remark 4.3. We claim that the constant c here can be made explicit, but this would require a more careful analysis at $p = 2, 3$ than we wish to describe here.

To derive this heuristic, we estimate the proportion of curves with a given ratio of (absolute) discriminant to conductor. Since the conductor is often the squarefree kernel of the discriminant, by way of explanation we first consider the behaviour of $f(n) = n/\mathrm{sqfree}(n)$. The probability that $f(n) = 1$ is given by the probability that n is squarefree, which is classically known to be $1/\zeta(2) = 6/\pi^2$. Given a prime power p^m, to have $f(n) = p^m$ says that $n = p^{m+1}u$ where u is squarefree and coprime to p. The probability that $p^{m+1} \| n$ is $(1 - 1/p)/p^{m+1}$, and given this, the conditional probability that $\bigl(n/p^{m+1}\bigr)$ is squarefree is $(6/\pi^2) \cdot (1 - 1/p^2)^{-1}$. Extending this multiplicatively beyond prime powers, we get that

$$\mathrm{Prob}\bigl[n/\mathrm{sqfree}(n) = q\bigr] = \frac{6}{\pi^2} \prod_{p^m \| q} \frac{1/p^{(m+1)}}{(1 + 1/p)} = \frac{6}{\pi^2} \frac{1}{q} \prod_{p|q} \frac{1}{p+1}.$$

In particular, the average of $f(n)^\gamma$ exists for $\gamma < 1$; in our elliptic curve analogue, we will require such an average for $\gamma = 5/6$. We note that it is an interesting question[11] to prove an asymptotic for $\sum_{n \le X} n/\mathrm{sqfree}(n)$.

---

[10] This argumentative technique can also be used to bolster our assumption that using Connell's conditions should be independent of other considerations.

[11] The saddle-point method as indicated in [56] or [9] might be applicable, but appears to involve quite careful estimation to achieve an asymptotic rather than a log-asymptotic. It was noted to us by G. Tenenbaum that [49] improves upon the result of [5], though the result is not that explicit.

4.1. **Derivation of the heuristic.** We keep the notation $D = |\Delta|$ and wish to compute the probability that $D/N = q$ for a fixed positive integer q. For a prime power p^v with $p \geq 5$, the probability that $p^v \| (D/N)$ is given by: the probability that E has multiplicative reduction at p and $p^{v+1} \| D$, that is $P_m(p, v+1)$; plus the probability that E has additive reduction at p and $p^{v+2} \| D$, that is $P_a(p, v+2)$; and the contribution from $P_0(p, v)$, which is zero for $v > 0$ and for $v = 0$ is the probability that p does not divide D. So, writing $v = v_p(q)$, we get that (with a similar modified formula for $p = 2, 3$)

$$\text{Prob}\big[D/N = q\big] = \prod_p \big[P_m(p, 1+v) + P_a(p, 2+v) + P_0(p, v)\big]. \tag{8}$$

We emphasise that this probability is (as in the last section) with respect to curve-ordering by discriminant, and (as previously) we have assumed that the primes act independently, that curves with $|\Delta| \leq X$ act like those in a big box, and that the effect of large primes is similar to that from small primes. Writing $\alpha = \alpha_+ + \alpha_-$ from Conjecture 2.1 we have

$$\begin{aligned} \sum_{E:N_E \leq X} 1 = \sum_{q=1}^{\infty} \sum_{\substack{E:N\leq X \\ D/N=q}} 1 &\approx \sum_{q=1}^{\infty} \sum_{E:D\leq qX} \text{Prob}\big[D/N = q\big] \\ &\sim \sum_{q=1}^{\infty} \alpha \cdot (qX)^{5/6} \cdot \text{Prob}\big[D/N = q\big], \end{aligned} \tag{9}$$

and if this last sum converges, we then get Heuristic 4.1.

To show the last sum in (9) does indeed converge, we get an upper bound for the probability in (8). We have that $P_m(p, v+1) \leq 1/p^{v+1}$ and $P_a(p, v+2) \leq 2/p^{v+1}$, which implies

$$\hat{f}(q) = \text{Prob}\big[D/N = q\big] \leq \frac{1}{q} \prod_{p|q} \frac{3}{p}.$$

We then estimate

$$\sum_{q=1}^{\infty} q^{5/6} \hat{f}(q) \leq \sum_{q=1}^{\infty} \frac{1}{q^{1/6}} \prod_{p|q} \frac{3}{p} = \prod_p \left(1 + \sum_{l=1}^{\infty} \frac{3/p}{(p^l)^{1/6}}\right) \leq \prod_p \left(1 + \frac{3/p}{p^{1/6} - 1}\right),$$

and the last product is convergent upon comparison to $\zeta(7/6)^3$. Thus we shown that the last sum in (9) converges, so that Heuristic 4.1 follows.

We can note that Fouvry, Nair, and Tenenbaum [26] have shown that the number of minimal models with $D \leq X$ is at least $cX^{5/6}$, and that the number of curves with $D \leq X$ with Szpiro ratio $\frac{\log D}{\log N} \geq \kappa$ is no more than $c_\epsilon X^{1/\kappa+\epsilon}$ for every $\epsilon > 0$.

4.2. **Dependence of D/N and the Tamagawa product.** We assume that D/N should be independent of the real period, but the Tamagawa product and D/N should be somewhat related.[12] We compute the expected square root of the Tamagawa product when $D/N = q$. As with (8) and using the ϵ defined in (6), we find that this is given by

$$\eta(q) = \prod_p \frac{\big[\epsilon_m(v_1) P_m(p, v_1) + \epsilon_a^n(v_2) P_a^n(p, v_2) + \epsilon_a^t(p, v_2) P_a^t(p, v_2) + P_0(p, v)\big]}{\big[P_m(p, v_1) + P_a(p, v_2) + P_0(p, v)\big]},$$

[12]The size of the torsion subgroup should also be related to D/N, but in the next section we argue that curves with nontrivial torsion are sufficiently sparse so as to be ignored.

where $v_1 = v+1$ and $v_2 = v+2$ and $v = v_p(q)$.

4.3. **The comparison of $\log\Delta$ with $\log N$.** We now want to compare $\log\Delta$ with $\log N$, and explicate the replacement therein in Guess 3.2. In order to bound the effect of curves with large D/N, we note that

$$\mathrm{Prob}\bigl[D/N \geq Y\bigr] = \sum_{q\geq Y} \hat{f}(q) \leq \sum_{q\geq Y} \frac{1}{q}\prod_{p|q}\frac{3}{p},$$

and use Rankin's trick (that is, bounding the characteristic function of $q \geq Y$ by $(q/Y)^{1-\alpha}$ for a parameter $0<\alpha<1$ that will be chosen optimally), so that for any $0<\alpha<1$ we have (using $p^\alpha - 1 \geq \alpha\log p$ in the penultimate step, and then the prime number theorem to bound $\sum_p \frac{1}{p\log p} \ll 1$)

$$\begin{aligned}\mathrm{Prob}\bigl[D/N \geq Y\bigr] &\leq \sum_{q=1}^{\infty}\Bigl(\frac{q}{Y}\Bigr)^{1-\alpha} \times \frac{1}{q}\prod_{p|q}\frac{3}{p} = \frac{Y^\alpha}{Y}\prod_p\Bigl(1+\frac{3}{p^{1+\alpha}}+\frac{3}{p^{1+2\alpha}}+\cdots\Bigr)\\ &= \frac{Y^\alpha}{Y}\prod_p\Bigl(1+\frac{3/p}{p^\alpha-1}\Bigr) \ll \frac{Y^\alpha}{Y}\exp\Bigl(\sum_p\frac{\tilde{c}/p}{\alpha\log p}\Bigr) \ll \frac{e^{c\sqrt{\log Y}}}{Y}\end{aligned}$$

for some constants $\tilde{c}, c$, by taking $\alpha = 1/\sqrt{\log Y}$ (this result is stronger than needed).

However, a more pedantic derivation of Guess 3.2 does not simply allow replacing $\log N$ by $\log\Delta$, but requires analysis (assuming $\Omega_{\mathrm{re}}(E)$ to be independent of q) of

$$\frac{\hat{\alpha}_R\bar{\alpha}_A}{3456\,\zeta(10)}\cdot \iint_{\sqrt{X}\leq\frac{u_4^3-u_6^2}{1728}\leq X} \Omega_{\mathrm{re}}(E)\cdot\Biggl[\sum_{q<\Delta}\eta(q)(\log\Delta/q)^{3/8}\cdot\mathrm{Prob}\bigl[D/N=q\bigr]\Biggr]\,du_4\,du_6.$$

The above estimate on the tail of the probability and a simple bound on $\eta(q)$ in terms of the divisor function shows that we can truncate the q-sum at Y with an error of $O_\epsilon(1/Y^{1-\epsilon})$ (for all $\epsilon>0$), and choosing (say) $Y = e^{\sqrt{\log X}}$ gives us that $\log(\Delta/q) \sim \log\Delta$ (note that we restricted to $\Delta > \sqrt{X}$). So the bracketed term becomes the desired

$$\sum_{q<Y}\eta(q)(\log\Delta)^{3/8}\cdot\mathrm{Prob}\bigl[D/N=q\bigr] \sim \beta(\sqrt{\tau})(\log\Delta)^{3/8},$$

upon noting that the q-part of the sum converges to $\beta(\sqrt{\tau})$ as $Y\to\infty$.

4.4. **Counting curves with vanishing L-value.** We now estimate the number of elliptic curves E with even parity and $L(E,1)=0$ when ordered by conductor.

Heuristic 4.4. *Let $\tilde{R}(X)$ be the number of elliptic curves E with even parity and conductor less than X and $L(E,1)=0$. Then $\tilde{R}(X)\sim cX^{19/24}(\log X)^{3/8}$ for some constant $c>0$.*

From Guess 3.2 we get that the number of even parity curves with $0<\Delta<qX$ and $D/N=q$ and $L(E,1)=0$ is given by

$$\hat{W}(qX)\cdot\bigl(\eta(q)/\beta(\sqrt{\tau})\bigr)\cdot\mathrm{Prob}\bigl[D/N=q\bigr],$$

and we sum this over all q. As we argued above, the tail of the sum does not affect the asymptotic (and so we can take $\log\Delta\sim\log N$ in $\hat{W}$), and again we get that the q-sum converges. This then gives the desired asymptotic for the number of even parity curves with conductor less than X and vanishing central L-value (upon arguing similarly for curves with negative discriminant).

4.5. **Relations to other work.** It is proposed by Hindry [31, Conj. 5.4 & 5.5] that a theorem of Brauer-Siegel type might hold for elliptic curves; that is, it should be that nonvanishing values of $L(E,1)$ are bounded quite far away (say $1/\log N$) from 0. This would say the product of the regulator and $\#Ш$ cannot be too small. We view this as unlikely; already in the rank 0 case we can see no reason why there should not be infinitely many curves with trivial Shafarevich-Tate group. Indeed, having $\#Ш = 1$ should be approximately as common as having positive rank according to the above discretisation methodology. The main difference between the elliptic curve case and that for number fields is that the latter deals with L-values at the edge of the critical strip, while our interest is in central values.

We might also point out that a guesstimate of $X^{19/24}$ curves of rank 2 up to X can also come from a couple of different methods. One method is to consider the (conjectural) BSD formula

$$S(E) = \frac{L(E,1)}{\Omega_{\mathrm{re}}(E)}\frac{T(E)^2}{\tau(E)} = \begin{cases} \#Ш(E) & \text{when } E \text{ has rank } 0,\\ 0 & \text{otherwise,}\end{cases}$$

and note that the torsion and Tamagawa contributions are small compared to the reciprocal of the real period.[13] A generalisation of the Lindelöf hypothesis implies that $L(E,1)$ is bounded above by something like $\log N$; this is also small compared to $1/\Omega_{\mathrm{re}}$, and so we view $S(E)$ as possibly taking values from 0 up to about $1/\Omega_{\mathrm{re}}$. Since $S(E)$ should be an integral square, this gives a crude probability of 1 in $\sqrt{1/\Omega_{\mathrm{re}}}$ of an curve of even parity having a vanishing central value. Summing over curves (which *inter alia* uses that $1/\Omega_{\mathrm{re}}$ is typically about $\Delta^{1/12}$), this gives[14] a rough count of $X^{19/24}$.

A different method to obtain $X^{19/24}$ is to estimate the number of integral points on the variety $y^2 = x^3 + Axz^2 + Bz^3$ for various ranges of (A,B,x,y,z). Though highly speculative, especially for larger ranks where arithmetic considerations may dominate, this gives an upper bound of size $X^{(21-r)/24}$ for the number of curves of rank r, yielding the asserted $X^{19/24}$ for $r = 2$. This will be discussed further in a forthcoming paper with A. Granville.

5. Torsion and isogenies

We can also count curves that have a given torsion group or isogeny structure. For instance, an elliptic curve with a 2-torsion point can be written as an integral model in the form $y^2 = x^3 + ax^2 + bx$ where $\Delta = 16b^2(a^2 - 4b)$; thus, by lattice-point counting, we estimate about $\sqrt{X}$ curves with absolute discriminant less than X. The effect on the conductor can perhaps more easily be seen by using the Fricke parametrisation $c_4 = (t+16)(t+64)T^2$ and $c_6 = (t-8)(t+64)^2T^3$ of curves with a rational 2-isogeny, and then substituting $t = p/q$ and $V = T/q$ to get $c_4 = (p+16q)(p+64q)V^2$ and $c_6 = (p-8q)(p+64q)^2V^3$ so that $\Delta = p(p+64q)^3q^2V^6$. The summation over the twisting parameter V just multiplies our estimate by a constant, while ABC-estimates imply that there should be no more than $X^{2/3+\epsilon}$ coprime pairs (p,q) with the squarefree kernel of $pq(p+64q)$ smaller than X in

[13] This can be made precise; below we note that $1/\Omega_{\mathrm{re}} \gg \Delta^{1/12}$, while the torsion is bounded and the Tamagawa product is bounded by a divisor function.

[14] This is vaguely related to the Sarnak estimate of $D^{3/4}$ for the count of vanishings in families of quadratic twists, but only relies upon the size of the real period.

absolute value. So we get the heuristic that almost all curves have no 2-torsion, even under ordering by conductor. Indeed, the exceptional set is so sparse that we can ignore it in our calculations. A similar argument applies for other isogenies, and more generally for splitting of division polynomials. Also, the results of Duke [22] for exceptional primes are applicable here, albeit with a different ordering.

6. Experiments

We wish to provide some experimental data for the above heuristics. However, it is difficult to distinguish numerically between 19/24 and 5/6 in the predictions

$$R(X) \sim cX^{19/24}(\log X)^{3/8} \quad \text{and} \quad A_\pm(X) \sim c' X^{5/6}.$$

Therefore, we instead try to refute the "null hypothesis", namely that there should be a positive proportion of rank 2 curves. In particular, the two large datasets of Brumer-McGuinness [7] and Stein-Watkins [53] for curves of prime conductor up to 10^8 and 10^{10} show little drop in the proportion of rank 2 curves, and an even smaller drop in the observed average (analytic) rank. These results led some to speculate that the average rank might (asymptotically) be greater than 0.5, with a positive proportion of elliptic curves having rank 2 or more.

Brumer and McGuinness considered about 310700 curves with prime conductor less than 10^8 and found an average rank of about 0.978, while Stein and Watkins extended this to over 11 million curves with prime conductor up to 10^{10} and found an average rank of about 0.964. Both datasets are expected to be nearly exhaustive[15] amongst curves with prime conductor up to the given limit. To extend the data in a computationally feasible manner, we chose a selection of curves with prime conductor of size 10^{14}. It is non-trivial to get a good dataset, since we must account for congruence conditions on the elliptic curve coefficients and the variation of the size of the real period.

6.1. Average analytic rank for curves with prime conductor near 10^{14}.

As in [53], we divided the (c_4, c_6) pairs into 288 congruence classes with $(\tilde{c}_4, \tilde{c}_6) = \bigl(c_4 \bmod 576, c_6 \bmod 1728\bigr)$. Many of these classes force the prime 2 to divide the discriminant, and thus do not produce any curves of prime conductor. For each class $(\tilde{c}_4, \tilde{c}_6)$, we took the 10000 parameter selections

$$(c_4, c_6) = \bigl(576(1000 + i) + \tilde{c}_4, 1728(100000 + j) + \tilde{c}_6\bigr) \text{ for } (i, j) \in [1..10] \times [1..1000],$$

and then of these 2880000 curves, took the 89913 models that had prime discriminant (note that all the discriminants are positive). This gives us good distribution across congruence classes, and while the real period does not vary as much as possible, below we will attempt to understand how this affects the average rank.

It then took a few months to compute the (suspected) analytic ranks for these curves. We got about 0.937 for the average rank. We then did a similar experiment for curves with negative discriminant given by

$$(c_4, c_6) = \bigl(576(-883 + i) + \tilde{c}_4, 1728(100000 + j) + \tilde{c}_6\bigr) \text{ for } (i, j) \in [1..10] \times [1..1000],$$

took the subset of 89749 curves with prime conductor, and found the average rank to be about 0.869. This discrepancy between positive and negative discriminant is also in the Brumer-McGuinness and Stein-Watkins datasets, and indeed was noted

[15] This is one reason to take prime conductor curves; we also have $|\Delta| = N$ with few exceptions.

in [7].[16] We do not average the results from positive and negative discriminant; the Brumer-McGuinness Conjecture 2.1 implies that the split is not 50-50.

In any case, our results show a substantial drop in the average rank, which, at the very least, indicates that the average rank is not constant in the range we considered. The alternative statistic of frequency of positive rank for curves with even parity also showed a significant drop. For prime positive discriminant curves it was 44.1% for Brumer-McGuinness and 41.7% for Stein-Watkins, but only 36.0% for our dataset — for negative discriminant curves of prime conductor, these numbers are 37.7%, 36.4%, and 31.3%.

6.2. **Variation of real period.** Our random sampling of curves with prime conductor of size 10^{14} must account for various properties of the curves if our results are to possess legitimacy. Above we speculated that the real period plays the most significant rôle, and so we wish to understand how our choice has affected it. Indeed, as was noted to us by X.-F. Roblot, the variation of real period when enumerating in a large (c_4, c_6)-box is quite different than when enumerating by discriminant. However, while this discrepancy with the distribution of the real period may be the weakest link in our experiment, we can still make a reasonable comparison between datasets, due to our assertion that only the size of the real period should matter.

To judge the effect that variation of the real period might have, we did some comparisons with the Stein-Watkins database. First consider curves of positive prime discriminant, and write E as $y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6$ and $e_1 > e_2 > e_3$ for the real roots of the cubic. We looked at curves with even parity and considered the frequency of positive rank as a function of the root quotient $t = \frac{e_1 - e_2}{e_1 - e_3}$, noting that[17] $\Omega_{\rm re}\Delta^{1/12} = \frac{2^{1/3}\pi(t-t^2)^{1/6}}{\mathrm{agm}(1,\sqrt{t})}$. The curves we considered all had $0.617 < t < 0.629$.

However, similar to when we considered curves ordered by conductor, before counting curves with extra rank, we should first simply count curves. Figure 1 indicates the distribution of root quotient t for the curves of prime (positive) discriminant and even parity from the Stein-Watkins database (more than 2 million curves meet the criteria). The x-axis is divided up into bins of size 1/1000; there are more than 100 times as many curves with $t < 0.001$ as there with $0.500 < t < 0.501$, with the most extremal dots not even appearing on the graph.

Next we plot the frequency of $L(E, 1) = 0$ as a function of the root quotient in Figure 2. Since there are only about 1000 curves in some of our bins, we do not get such a nice graph. Note that the left-most and especially the rightmost dots are much below their nearest neighbors, the graph slopes down in general, and drops more at the end. We see no evidence that our results should be overly biased. In particular, the frequency of $L(E, 1) = 0$ is 41.7% amongst all even parity curves of prime discriminant in the Stein-Watkins database, and is 42.8% for the 12324 such curves with $0.617 < t < 0.629$. The function plotted (labelled on the right axis) in Figure 2 is of $\Omega_{\rm re}\Delta^{1/12} = \frac{2^{1/3}\pi(t-t^2)^{1/6}}{\mathrm{agm}(1,\sqrt{t})}$ as a function of t, and note that this goes to 0 as $t \to 0, 1$; there is nothing canonical about the choice of our t-parameter, and we chose it more for convenience than anything else.

[16] "An interesting phenomenon was the systematic influence of the discriminant sign on all aspects of the arithmetic of the curve."

[17] The calculation follows as in the previous sections; via calculus, we can compute that this function is maximised at $t \approx .0388505246188$ with a maximum just below 4.414499094.

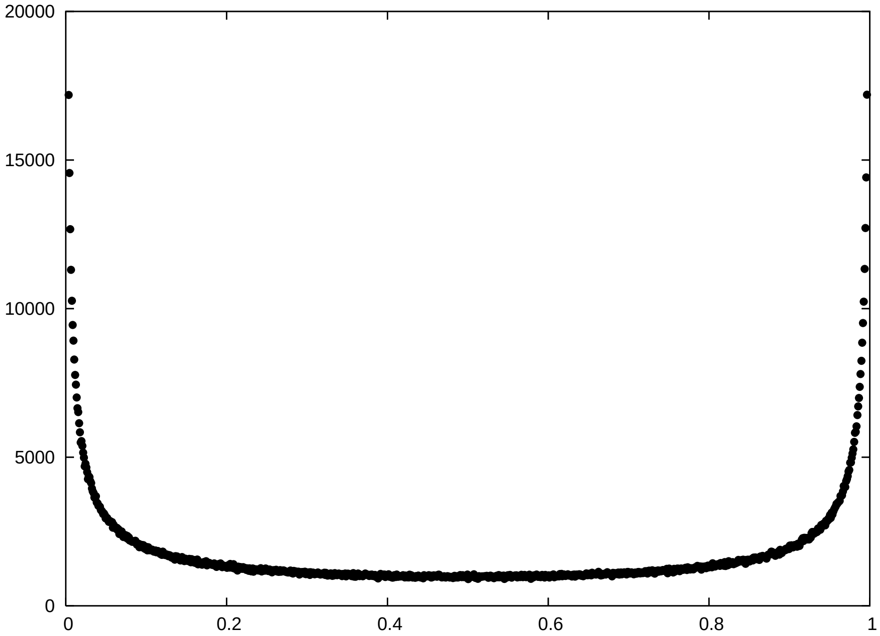

FIGURE 1. $\Delta > 0$: Curve distribution as a function of t

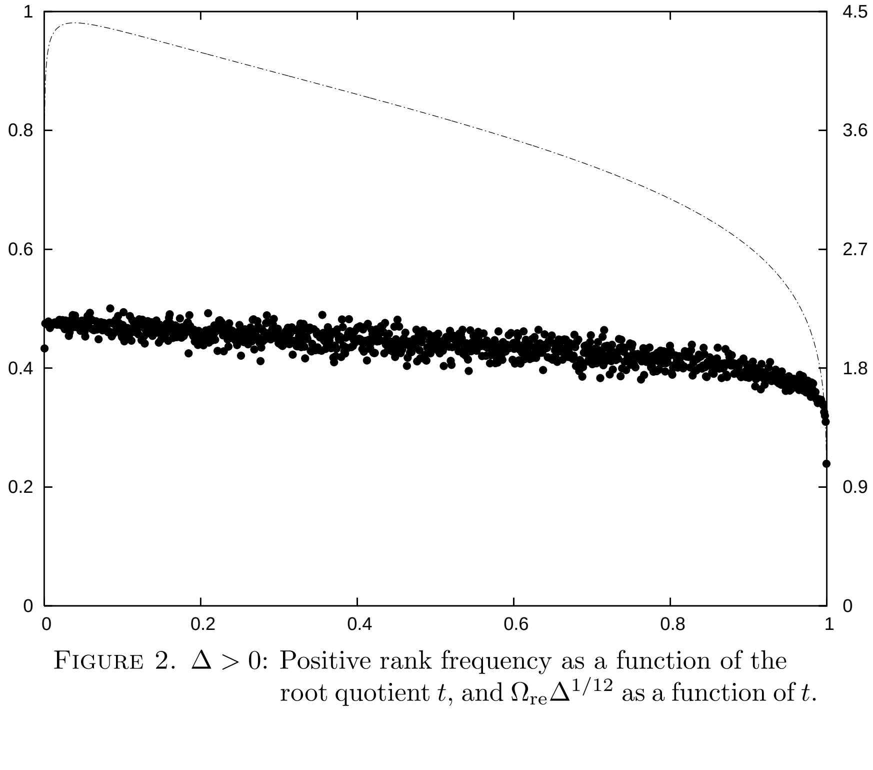

FIGURE 2. $\Delta > 0$: Positive rank frequency as a function of the root quotient t, and $\Omega_{\mathrm{re}}\Delta^{1/12}$ as a function of t.

Similar computations can be made in the case of negative discriminant, which we briefly discuss for completeness (again restricting to curves with even parity where appropriate). Let r be the real root of the cubic polynomial $4x^3+b_2x^2+2b_4x+b_6$, and $Z>0$ the imaginary part of the conjugate pair of nonreal roots. Letting $\tilde{r}=r+b_2/12$ and $c=\tilde{r}/Z$ we then have[18]

$$\Omega_{\mathrm{re}}|\Delta|^{1/12}=\frac{\pi\sqrt{2}}{(1+9c^2/4)^{1/12}\mathrm{agm}\Big(1,\sqrt{\frac{1}{2}+\frac{3c}{4\sqrt{1+9c^2/4}}}\Big)}.$$

We renormalise via taking $C=1/2+\arctan(c)/\pi$, and graph the distribution of curves versus C in Figure 3. The symmetry[19] of the graph might indicate that the coordinate transform is reasonable. All our curves have $0.555<C<0.557$.

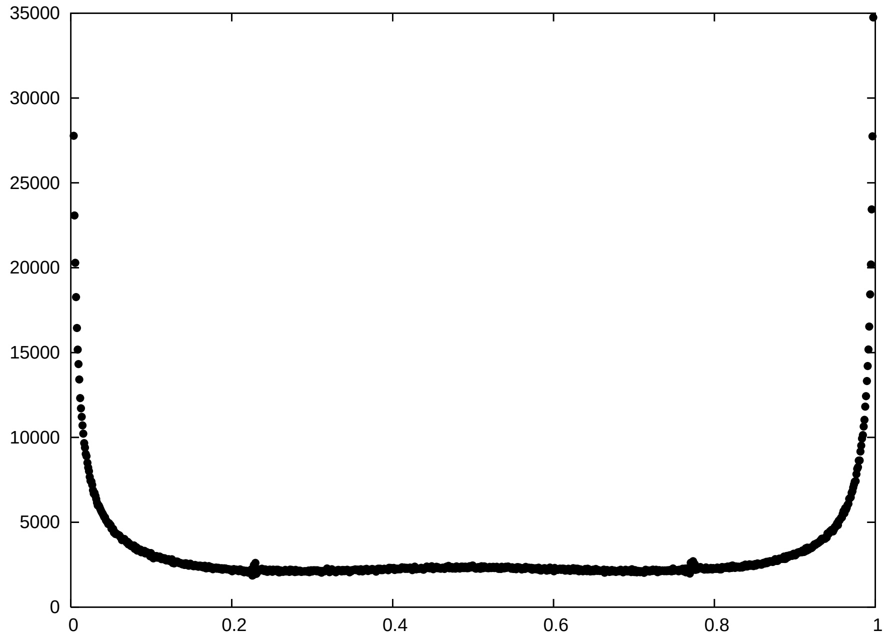

FIGURE 3. $\Delta<0$: Distribution of curves as a function of C

Next we plot the frequency of $L(E,1)=0$ as a function of the root quotient in Figure 4. Again we also graph the function $\Omega_{\mathrm{re}}|\Delta|^{1/12}$ on the right axis. Here the drop-off is more pronounced than with the curves of positive discriminant. Note the floating dot around $C=1/2$. Indeed the 100 closest curves with $C<1/2$ all have positive rank; this breaks down when crossing the 1/2-barrier. This is not particularly a mystery; these curves have $a_6=0$ and/or $b_6=1$, and thus have an obvious rational point. Recall that $C=1/2$ corresponds to $c=0=\tilde{r}$.

We again see no evidence that our results should be biased. In particular, the frequency of $L(E,1)=0$ is 36.4% amongst all even parity curves of negative prime discriminant in the Stein-Watkins database, and is 37.0% for the 4695 such curves with $0.555<C<0.557$.

[18] This is maximised at $c\approx -33.58515148525$, with the maximum a bit less than 8.82921518.

[19] The blotches around 0.22-0.23 and 0.77-0.78 appear to come from the fact that curves with a_4 small (in particular ± 1) tend to have C in these ranges (for our discriminant range), and this causes instability in the counting function.

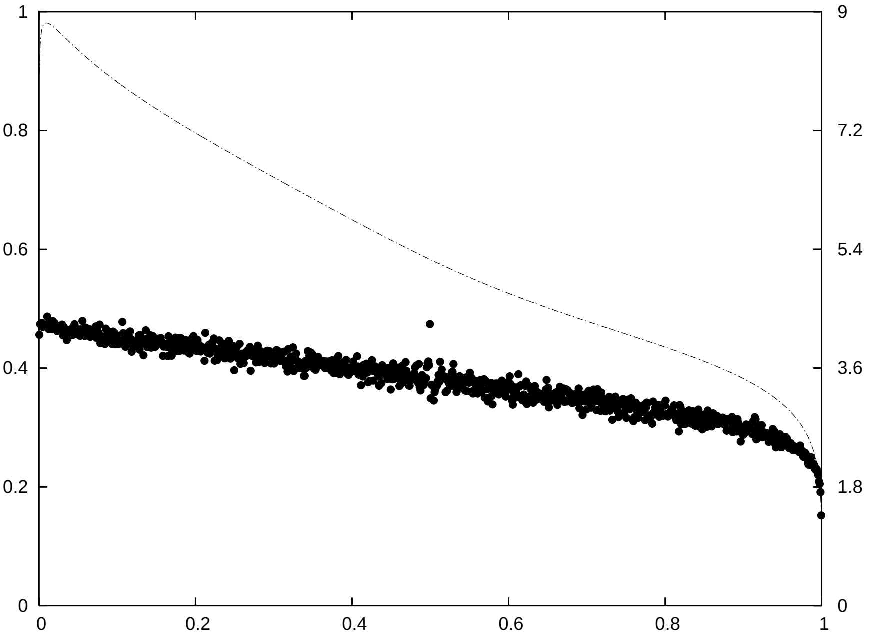

Figure 4. $\Delta < 0$: Positive rank frequency as a function of C, and $\Omega_{\rm re}|\Delta|^{1/12}$ as a function of C.

6.3. **Other considerations.** The idea that the "probability" that an even parity curve possesses positive rank should be proportional to $\sqrt{\Omega_{\rm re}}$ is perhaps overly simplistic — in particular, it is not borne out too precisely by the Stein-Watkins dataset. We consider positive prime discriminant curves with even parity; for those with $0.64 < \Omega_{\rm re} < 0.65$ we have 78784 curves of which 45.9% have positive rank, while of the 9872 with $0.32 < \Omega_{\rm re} < 0.325$ we have 36.0% with positive rank, for a ratio of 1.28, which is not too close to $\sqrt{2}$. One consideration here is that we have placed a discriminant limit on our curves, and there are curves with larger discriminant and $0.32 < \Omega_{\rm re} < 0.325$ that we have not considered. This, however, is extra-particular to the idea that only the real period should be of import.

One possibility is that curves with small discriminant and/or large real period have *smaller* probability of $L(E,1)=0$ that our estimate of $c\sqrt{\Omega_{\rm re}}$ would suggest — indeed, it might be argued (maybe due to arithmetic considerations, or perhaps explicit formulae for the zeros of L-functions) that curves with such small discriminant cannot realise their nominal expected frequency of positive rank. Unfortunately, we cannot do much to quantify these musings, as the effect would likely be in a secondary term, making it difficult to detect experimentally. Note also that a relative depression of rank for small discriminant curves would give a reason for the near-constant average rank observed by Brumer-McGuinness and Stein-Watkins.

6.4. **Mordell-Weil lattice distribution for rank 2 curves.** We have other evidence that curves of small discriminant might not behave quite as expected. We undertook to compute generators for the Mordell-Weil group for all 2143079 curves of (analytic) rank 2 of prime conductor less than 10^{10} in the Stein-Watkins database.[20] J. E. Cremona ran his `mwrank` programme [17] on all these curves, and it was successful in provably finding the Mordell-Weil group for 2114188 of these. For about 2500 curves, the search region was too big to find the 2-covering quartics via invariant methods, while around 8500 curves had a generator of large height that could not be found, and over 18000 had 2-Selmer rank greater than 2. We then used the `FourDescent` machinery of MAGMA [3] which reduced the number of problematic curves to 54. Of these, 19 have analytic Ш of 16.0 and we expect that either 3-descent or 8-descent [52] will complete (assuming GRH to compute the class group) the Mordell-Weil group verification; for the 35 other curves, there is likely a generator of height more than 225 which we did not attempt to find.[21]

We then looked at the distribution of the Mordell-Weil lattices obtained from the induced inner product from the height pairing; since all of our curves have rank 2, we get 2-dimensional lattices. We are not so interested in the size of the obtained lattices, but more so in their shape. Via the use of lattice reduction (which reduces to continued fractions in this case), given any two generators we can find the point P of smallest positive height on the curve. By normalising P to be the unit vector, we then get a vector in the upper-half-plane corresponding to another generator Q. Via the standard reduction algorithm, we can translate Q so that it corresponds to a point in the fundamental domain for the action of $\mathrm{SL}_2(\mathbf{Z})$. Finally, by replacing Q by $-Q$ if necessary, we can ensure that this point is in the right half of the fundamental domain (in other words, we must choose an embedding for our Mordell-Weil lattice). In this manner, for each rank 2 curve we associate a unique point $z = x + iy$ in the upper-half-plane with $x^2 + y^2 \geq 1$ and $0 \leq x \leq 1/2$.

With no other guidance, we might expect that the obtained distribution for the z is given by[22] the Haar measure $(dx\,dy)/y^2$. We find, however, that this is not borne out too well by experiment. In particular, we should expect that $\frac{1/2}{\pi/6} \approx 95.5\%$ of the curves should have $y \geq 1$, while the experimental result is about 93.5%. Furthermore, we should expect that the proportion of curves with $y \geq Y$ should die off like $1/Y$ as $y \to \infty$; however, we get that 35.6% of the curves have $y \geq 2$, only 9.7% have $y \geq 4$, while 1.97% have $y \geq 8$ and 0.35% have $y \geq 16$. The validity of the vertical distribution data might be arguable based upon concerns regarding the discriminant cutoff of our dataset, but the horizontal distribution is also skewed. If we consider only curves with $y \geq 1$, then we should get uniform distribution in the x-aspect; however, Table 1 shows that we do not have such uniformity.

TABLE 1. Horiztonal distribution of rank 2 lattices with $y \geq 1$

$0.00 \leq x < 0.05$	9.1%	$0.25 \leq x < 0.30$	10.0%
$0.05 \leq x < 0.10$	9.7%	$0.30 \leq x < 0.35$	10.1%
$0.10 \leq x < 0.15$	9.8%	$0.35 \leq x < 0.40$	10.4%
$0.15 \leq x < 0.20$	9.8%	$0.40 \leq x < 0.45$	10.6%
$0.20 \leq x < 0.25$	9.9%	$0.45 \leq x \leq 0.50$	10.6%

[20]We also computed the Mordell-Weil group for curves with higher ranks but do not describe the obtained data here.

[21]A bit more searching might resolve a few of the outstanding cases, but the extremal case of $[0, 0, 1, -237882589, -1412186639384]$ appears to have a generator of height more than 600, and thus other methods are likely to be needed in order to find it. Indeed, T. A. Fisher has recently [25] used 6-descent and 12-descent to find the remaining 35 points.

[22]Siegel [50] similarly uses Haar measure to put a natural measure on n-dimensional lattices of determinant 1.

We cannot say whether these unexpected results from the experimental data are artifacts of choosing curves with small discriminant; it is just as probable that our Haar-measure hypothesis concerning the lattice distribution is simply incorrect. At the suggestion of D. B. Zagier, we made a density plot of the ratio between the experimental and conjectural counts for bins in the fundamental domain — see Figure 5, where the axes are switched and the y-coordinate is plotted logarithmically. At the right edge of the graph, the conjectural amount ($\sim$1500 per bin) is typically 10 times the experimental amount — this increases to a factor of 100 for $y \sim 70$. Overall, our data seem to imply that the lattices are not as skinny and are less orthogonal than might be guessed.

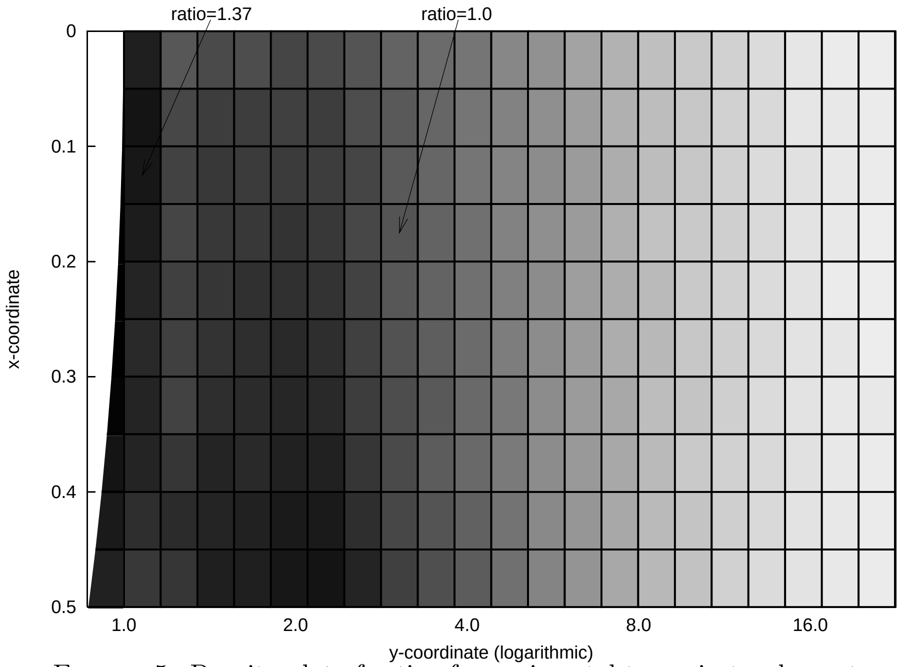

FIGURE 5. Density plot of ratio of experimental to conjectural counts

6.5. **Symmetric power L-functions.** Similar to questions about the vanishing of $L(E,s)$, we can ask about the vanishing of the symmetric power L-functions $L(\mathrm{Sym}^{2k-1}E,s)$. We refer the reader to [39] for more details about this, but mention that, due to conjectures of Deligne and more generally Bloch and Beĭlinson [45], we expect that we should have a formula similar to that of Birch and Swinnerton-Dyer, stating that $L(\mathrm{Sym}^{2k-1}E,k)(2\pi N)^{\binom{k}{2}}/\Omega_+^{\binom{k+1}{2}}\Omega_-^{\binom{k}{2}}$ should be rational with small denominator. Here, for k odd, Ω_+ is the real period and Ω_- the imaginary period, with this reversed for k even. As noted in [8], the order of vanishing of the central L-value should be related to the rank of the Griffiths group of the symmetric power variety.

Ignoring the contribution from the conductor, and crudely estimating that $\Omega_+ \approx \Omega_- \approx 1/\Delta^{1/12}$, an application of discretisation as before gives that the probability that $L(\mathrm{Sym}^{2k-1}E,s)$ has even parity and $L(\mathrm{Sym}^{2k-1}E,k) = 0$ is bounded above (cf. the ignoring of N) by $c(\log\Delta)^{3/8}\cdot\sqrt{1/\Delta^{k^2/12}}$. Again following the analogy of above, we can then get an upper bound $c_k(\epsilon)X^{5/6-k^2/24+\epsilon}$ (for every $\epsilon > 0$) for the number of curves with conductor less than X with even-signed symmetric $(2k-1)$st power and $L(\mathrm{Sym}^{2k-1}E,k)=0$. It could be argued that we should order curves according to the conductor of the symmetric power L-function rather than that of the curve, but we do not think such concerns are that relevant to our imprecise discussion. In particular, the above estimate predicts that there are finitely many curves with extra vanishing when $k\geq 5$ (that is, finitely many extra vanishings for the 9th symmetric power and beyond). It should be said that this heuristic will likely mislead us about curves with complex multiplication, for which the symmetric power L-function factors (it is imprimitive in the sense of the Selberg class), with each factor having a 50% chance of having odd parity. However, even ignoring CM curves, the data[23] of [39] find a handful of curves for which the 9th, 11th and even the 13th symmetric powers appear (up to 12 digits of precision) to have a central zero of order 2. We find this surprising, and casts some doubt about the validity of our methodology of modelling of vanishings.

6.6. **Quadratic twists of higher symmetric powers.** The techniques we used earlier in this paper have also been used to model vanishings in quadratic twist families, and we can extend the analyses to symmetric powers.

6.6.1. *Non-CM curves.* We fix a non-CM curve E and let E_d be its dth quadratic twist, taking d to be a fundamental discriminant. From an analogue of the Birch–Swinnerton-Dyer conjecture we expect to get a rational number with small denominator from the quotient[24]

$$\frac{L(\mathrm{Sym}^3E_d,2)(2\pi N_E)}{\Omega_{\mathrm{im}}(E_d)^3\Omega_{\mathrm{re}}(E_d).}$$

We have that $\Omega_{\mathrm{im}}(E_d)^3\Omega_{\mathrm{re}}(E_d)\approx\Omega_{\mathrm{im}}(E)^3/d^{3/2}\cdot\Omega_{\mathrm{re}}(E)/d^{1/2}$ (with the periods reversed when $d<0$), and so we expect the number of fundamental discriminants $|d|<D$ such that $L(\mathrm{Sym}^3E_d,s)$ has even parity with $L(\mathrm{Sym}^3E_d,2)=0$ to be given crudely (up to log-factors) by $\sum_{d<D}c/\sqrt{d^2}$. So we expect about $(\log D)^b$ quadratic twists with double zeros for the 3rd symmetric power; generalising predicts finitely many extra vanishings for higher (odd) powers.

We took the curves 11a:$[0,-1,1,0,0]$ and 14a:$[1,0,1,-1,0]$, and computed either $L(\mathrm{Sym}^3E_d,2)$ or $L'(\mathrm{Sym}^3E_d,2)$ for all fundamental discriminant d with $|d|<5000$. We did the same for 15a:$[1,1,1,0,0]$ for $|d|<4000$. We then looked at the number of vanishings (to 9 digits of precision). For 11a we found 58 double zeros and one triple zero (indicated by a star in Table 2) while for 14a we found 88 double zeros and three triple zeros, and 15a yielded 83 double zeros and two triple zeros. It is quite difficult to accrue much data, mostly due to the fast growth of the conductor; for elliptic curves, Rubinstein [15] does not compute $L(E_d,1)$ directly, but rather uses the Waldspurger formula (as made explicit in works such as [43]), and then just computes the coefficients of a modular form of weight 3/2.

[23]Data in Table 6 of [39] are inexact, as [18] had missed curves with $6^3|N$ and $90000\leq N\leq 10^5$.

[24]The contribution from the conductor actually comes from non-integral Tamagawa numbers from the Bloch-Kato exponential map, and in the case of quadratic twists, the twisting parameter d should not appear in the final expression.

TABLE 2. Fundamental d with $\operatorname*{ord}_{s=2} L(\mathrm{Sym}^3 E_d, s) \geq 2$

11a	−40 −52 −563 −824 −1007 −1239 −1460 −1668 −1799 −2207
	−2595 −2724 −2980 −3108 −3592 −4164 −4215 −4351 −4399
	12 69 152 181 232 273 364 401 412 421 444 476 488 652 669 696
	933 1101 1149 1401 1576 1596 1676 1884 1928 2348 2445 2616
	2632 3228 3293 3404 3720* 3793 4060 4093 4161 4481 4665 4953
14a	−31 −52 −67 −87 −91 −111 −203 −223 −255 −264 −271 −311 −327
	−367 −535 −552 −651 −759 −804 −831 −851 −852 −920 −1099 −1263
	−1267 −1335 −1524 −1547 −1567 −1623 −1679 −1707* −2047 −2235
	−2280 −2407 −2443 −2563 −2824 −2831 −3127 −3135 −3523 −4119
	−4179 −4191 −4323
	137 141* 229 233 281 345 469 473 492 497* 697 901 1065 1068 1353 1457
	1481 1513 1537 1793 1873 1905 2024 2093 2193 2265 2321 2589 2657 2668
	2732 2921 2981 2993 3437 3473 3529 4001 4124 4389 4488 4661 4817
15a	−11 −51 −71 −164 −219 −232 −292 −295 −323* −340 −356 −399 −519
	−580 −583 −584 −671 −763 −804 −851 −879 −943 −1012 −1060 −1151
	−1199 −1284 −1288 −1363 −1551 −1615 −1723 −1732 −2279 −2291
	−2379 −2395 −2407 −2571 −2632 −2635 −2756 −3396 −3588* −3832
	17 21 61 77 136 156 181 229 349 444 481 501 545 589 649 781 876 905
	924 949 1009 1144 1249 1441 1501 1580 1621 1804 1861 1921 2041 2089
	2109 2329 2581 2829 2840 2933 3001 3916

6.6.2. *CM curves.* Next we consider CM curves, for which we can compute significantly more data, but the modelling of vanishings is slightly different. Let E be a rational elliptic curve with CM, and ψ its Hecke character. We shall take ψ to be "twist-minimal" — this is not the same as the "canonical" character of Rohrlich [47], but rather we just take E to be a minimal (quadratic) twist. Indeed, we shall only consider 11 different choices of E, given (up to isogeny class) by 27a, 32a, 36a, 49a, 121a, 256a, 256b, 361a, 1849a, 4489a, and 26569a, noting that 27a/36a and 32a/256b are respectively cubic and quartic twist-pairs. In the tables herein, these can appear in a briefer format, such as 67^2 for 4489a.

We normalise the Hecke L-function $L(\psi, s)$ to have $s = 1$ be the center of the critical strip. For d a fundamental discriminant, we let ψ_d be the Hecke Grössencharacter ψ twisted by the quadratic Dirichlet character of discriminant d. Finally, note that the symmetric powers $L(\mathrm{Sym}^{2k-1}\psi, s)$ are just $L(\psi^{2k-1}, s)$, where we must take ψ^{2k-1} to be the primitive underlying Grössencharacter.

We then expect $L(\psi^3, 2)(2\pi)/\Omega_{\mathrm{im}}(E)^3$ to be rational with small denominator. We can then use discretisation as before to count the expected number of fundamental discriminants $|d| < D$ for which the L-function $L(\psi_d^3, s)$ has even parity but vanishes at the central point — since we have $\Omega_{\mathrm{im}}(E_d)^3 \approx \Omega_{\mathrm{im}}(E)^3/d^{3/2}$, we expect the number of discriminants d that yield even parity and $L(\psi_d^3, 2) = 0$ is crudely given by $\sum_{d<D} 1/\sqrt{d^{3/2}}$, so we should get about $D^{1/4}$ such discriminants up to D. Alternatively, we can note that ψ^3 corresponds to a weight 4 modular form, so that there is a weight 5/2 Shimura lift of it whose coefficients are related to $L(\psi_d^3, 2)$ via the Waldspurger correspondence; upon considering how often these coefficients should vanish, we obtain a similar heuristic.

For higher symmetric powers, we expect that $L(\psi^{2k-1}, k)(2\pi)^{k-1}/\Omega_+(E)^{2k-1}$ is rational with small denominator, and thus get that there should be finitely many quadratic twists of even parity with vanishing central value.

We took the above eleven CM curves and took their (fundamental) quadratic[25] twists up to 10^5. We must be careful to exclude twists that are isogenous to other twists. In particular, we need to define a **primitive** discriminant for a curve with CM by an order of the field K — this is a fundamental discriminant d such that $\mathrm{disc}(K)$ does not divide d, expect for $K = \mathbf{Q}(i)$ when $d > 0$ is additionally primitive when $8\|d$. Note also that 27a and 36a have the same symmetric cube L-function.

Table 3. Counts of double order zeros for primitive twists

	27a	32a	36a	49a	121a	256a	256b	361a	1849a	4489a	26569a
3rd	59	32	-	67	78	32	21	45	28	31	1
5th	3	1	5	2	1	2	2	0	0	0	0
7th	0	0	2	0	1	0	0	0	0	0	0

Table 3 lists our results for counts of central double zeros (to 32 digits) for the L-functions of the 3rd, 5th, and 7th symmetric powers.[26] Tables 4 and 5 list the primitive discriminants that yield the double zeros. The notable signedness can be explained via the sign of the functional equation.[27] We are unable to explain the paucity of double zeros for twists of 26569a; Liu and Xu have the latest results [37] on the vanishing of such L-functions, but their bounds are far from the observed data. Similarly, the last-listed double zero for 4489a at 67260 seems a bit odd. We stress, however, that we fully expect the asymptotic prediction of $cD^{1/4}(\log D)^b$ to be correct here, our suspicion being that the constant c for 26569a is rather small.

There appear to be implications vis-a-vis higher vanishings in some cases; for instance, except for 27a, in the thirteen cases that $L(\psi_d^5, s)$ has a double zero at $s = 3$ we have that $L(\psi_d, s)$ also has a double zero at $s = 1$. Similarly, the 7th symmetric power for the 27365th twist of 121a has a double zero, as does the 3rd symmetric power, while the L-function of the twist itself has a triple zero. Also, the 22909th twist of 36a has double zeros for its first, third, and fifth powers (note that 36a does not appear in Table 4 as the data are identical to that for 27a).

6.6.3. *Comparison between the CM and non-CM cases.* For the twist computations for the symmetric powers, we can go much further (about 20 times as far) in the CM case because the conductors do not grow as rapidly. For the 3rd symmetric power, the crude prediction is that we should have (asymptotically) many more extra vanishings for twists in the CM case than in the non-CM case, but this is not borne out by the data. Additionally, we have no triple zeros in the CM case (where the dataset is almost 100 times as large), while we already have six for the non-CM curves.[28] This is directly antithetical to our suspicion that there should be more extra vanishings in the CM case. As before, this might cast some doubt on our methodology of modelling of vanishings.

[25] The quartic twists of 32a and cubic twists of 27a/36a might also give interesting data — already in the early 1980s N. M. Stephens computed that the 7th symmetric power of the Hecke character for $y^2 = x^3 + 127x$ yields a double central vanishing. See [29] for related information.

[26] We found no even twists with $L(\psi_d^9, 5) = 0$, and no triple zeros appeared in the data.

[27] The local signs at $p = 2, 3$ involve wild ramification and are thus much more complicated (see [35, 58, 21, 24] for a theoretical description); thus there is no complete correlation.

[28] We similarly checked twists of the level 5 weight 4 cusp form, with no triple zeros up to 10^5.

TABLE 4. Primitive d with $\operatorname*{ord}_{s=2} L(\psi_d^3, s) = 2$

27a	172 524 1292 1564 1793 3016 4169 4648 6508 9149 9452 9560 10636 11137 12040 13784 14284 15713 17485 17884 22841 22909 22936 25729 27065 27628 29165 30392 34220 35749 38636 40108 41756 44221 47260 51512 54385 57548 58933 58936 58984 59836 59996 62353 64268 70253 74305 77320 77672 78572 84616 86609 86812 87013 92057 95861 96556 97237 99817
32a	−395 −5115 −17803 −25987 −58123 −60347 −73635 −79779 −84651 −99619 257 1217 2201 2465 14585 26265 45201 82945 4632 5336 5720 7480 9560 30328 30360 31832 38936 45848 69784 71832 83512 92312
49a	−79 −311 −319 −516 −856 −1007 −1039 −1243 −1391 −1507 −1795 −2024 −2392 −2756 −2923 −3527 −3624 −4087 −4371 −4583 −4727 −5431 −5524 −5627 −6740 −7167 −7871 −8095 −8283 −10391 −10628 −13407 −13656 −13780 −16980 −18091 −22499 −27579 −28596 −30083 −30616 −32303 −32615 −36311 −36399 −38643 −39127 −40127 −42324 −52863 −64031 −64399 −66091 −66776 −66967 −69647 −70376 −71455 −72663 −73487 −73559 −77039 −84383 −90667 −91171 −98655 −98927
11^2	12 140 632 1160 1208 1308 1704 1884 2072 2136 2380 2693 2716 3045 4120 4121 5052 5528 5673 5820 6572 7532 11053 11208 12277 12568 12949 13884 14844 15465 16136 18588 18885 19020 19884 24060 25788 27365 27597 28265 28668 29109 29573 32808 32828 35261 36552 37164 38121 38297 44232 44873 49512 49765 50945 52392 54732 55708 56076 56721 58460 59340 65564 66072 66833 71688 72968 79557 80040 80184 83388 84504 84620 84945 86997 87576 92460 95241
256a	401 497 2513 3036 3813 6933 6941 9596 9932 11436 14721 17133 17309 18469 21345 21749 26381 26933 28993 29973 30461 33740 51469 53084 62556 63980 67721 69513 73868 76241 81164 87697
256b	73 345 3521 5133 6693 7293 21752 25437 27113 34657 38485 41656 42433 44088 46045 75581 79205 83480 89737 93624 96193
19^2	44 60 1429 1793 3297 3340 3532 3837 3880 4109 5228 5628 7761 8808 9080 9388 12280 12313 12545 13373 13516 13897 19164 22204 23241 25036 25653 41205 41480 42665 43429 44121 44285 44508 45660 48828 50584 52989 64037 74585 75324 76921 81885 85036 96220
43^2	88 152 440 2044 4268 5852 6376 7880 8908 9880 14252 15681 17864 20085 20353 28492 29477 45368 55948 56172 57409 60177 68136 79916 84524 85580 86853 96216
67^2	17 57 869 1612 1628 3260 6380 6385 7469 8328 11017 13772 14152 14268 14552 15901 22513 24605 24664 27992 29676 33541 33789 36344 36588 38028 40280 43041 49884 62353 67260
163^2	30720

TABLE 5. Primitive d with $\operatorname*{ord}_{s=k} L(\psi_d^{2k-1}, s) = 2$ for some $k \geq 3$

27a	5th: −13091 4040 18044	49a	5th: 437 19317
32a	5th: 1704	121a	5th: −183 7th: 27365
36a	5th: −856 −2104 −31592 −88580 22909	256a	5th: −79 −21252
36a	7th: −95 2488	256b	5th: −511 89320

In [46, §8] Rodriguez Villegas and Zagier mention the possibility of a formula of Waldspurger type for the twists of the Hecke Grössencharacters, but it does not seem that an exact formula has ever appeared. Using the techniques developed by Basmaji and Frey [27], we are able to compute the weight 5/2 lift for (say) the symmetric cube of Hecke Grössencharacter attached to 49a. However, as we are currently unable to write it as a twisted ternary theta series as in [48], it does not seem to aid our computations. In the non-CM case, it has been noted by R. Schulze-Pillot that a special case of a result of Ramakrishnan and Shahidi [44] re-interprets the symmetric cube L-function as a weight 3 spinor L-function associated to a degree 2 Siegel modular form; again we are currently unable to use this in our computations.

7. Acknowledgements

The author was partially supported by Engineering and Physical Sciences Research Council (EPSRC) grant GR/T00658/01 (United Kingdom). He thanks I. A. Burhanuddin, N. D. Elkies, H. A. Helfgott, R. C. Vaughan, and A. Venkatesh for useful comments, and N. P. Jones for the reference [22], and G. Tenenbaum for the reference [49].

References

[1] B. J. Birch, *How the number of points of an elliptic curve over a fixed prime field varies.* J. London Math. Soc. **43** (1968), 57–60.
[2] B. J. Birch, H. P. F. Swinnerton-Dyer, *Notes on elliptic curves. I. II.* J. reine angew. Math. **212** (1963), 7–25, **218** (1965), 79–108.
[3] W. Bosma, J. Cannon, C. Playoust, *The Magma algebra system. I. The user language.* In *Computational algebra and number theory* Proceedings of the 1st MAGMA Conference held at Queen Mary and Westfield College, London, August 23–27, 1993. Edited by J. Cannon and D. Holt, Elsevier Science B.V., Amsterdam (1997), 235–265. Cross-referenced as J. Symbolic Comput. **24** (1997), no. 3-4, 235–265. Online at `magma.maths.usyd.edu.au`
[4] 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** (2001), no. 4, 843–939.
[5] N. G. de Bruijn, *On the number of integers $\leq x$ whose prime factors divide n.* Illinois J. Math. **6** (1962), 137–141.
[6] A. Brumer, *The average rank of elliptic curves. I.* Invent. Math. **109** (1992), no. 3, 445–472.
[7] A. Brumer, O. McGuinness, *The behavior of the Mordell-Weil group of elliptic curves.* Bull. Amer. Math. Soc. (N.S.) **23** (1990), no. 2, 375–382.
[8] J. Buhler, C. Schoen, J. Top, *Cycles, L-functions and triple products of elliptic curves.* J. Reine Angew. Math. **492** (1997), 93–133.
[9] S. Burris, K. Yeats, *Admissible Dirichlet series.* Preprint, 2005, available online at `arxiv.org/math/0507487`
[10] H. Cohen, *A Course in Computational Algebraic Number Theory.* Grad. Texts in Math. **138**, Springer-Verlag, 1993.
[11] I. Connell, *The Elliptic Curve Handbook.* Lecture Notes from class at McGill University, 1991. See `www.math.mcgill.ca/connell/public/ECH1`
[12] B. Conrad, F. Diamond, R. Taylor, *Modularity of certain potentially Barsotti-Tate Galois representations.* J. Amer. Math. Soc. **12** (1999), no. 2, 521–567.
[13] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, N. C. Snaith, *Integral moments of L-functions.* Proc. London Math. Soc. (3) **91** (2005), no. 1, 33–104. See `arxiv.org/math/0206018`
[14] J. B. Conrey, J. P. Keating, M. O. Rubinstein, N. C. Snaith, *On the frequency of vanishing of quadratic twists of modular L-functions.* Number theory for the millennium, I (Urbana, IL, 2000), A K Peters, Natick, MA (2002), 301–315. See `arxiv.org/math/0012043`
[15] ———, *Random Matrix Theory and the Fourier Coefficients of Half-Integral Weight Forms*, Experimental Math. **15** (2006), no. 1, 67–82. Online at `arxiv.org/math/0412083`

[16] J. B. Conrey, A. Pokharel, M. O. Rubinstein, and 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*, edited by J. B. Conrey, D. W. Farmer, F. Mezzadri, and N. C. Snaith, London Mathematical Society Lecture Note Series **341**, Cambridge University Press (2007), 215–232.

[17] J. E. Cremona, `mwrank` (sotfware), `www.warwick.ac.uk/~masgaj/ftp/progs`

[18] J. E. Cremona, Elliptic Curve Data, at `www.warwick.ac.uk/~masgaj/ftp/data/`

[19] C. Delaunay, M. Watkins, *The powers of logarithm for quadratic twists*, in *Ranks of Elliptic Curves and Random Matrix Theory*, edited by J. B. Conrey, D. W. Farmer, F. Mezzadri, and N. C. Snaith, London Mathematical Society Lecture Note Series **341**, Cambridge University Press (2007), 189–193,

[20] F. Diamond, *On deformation rings and Hecke rings.* Ann. of Math. (2) **144** (1996), no. 1, 137–166.

[21] T. Dokchiter, V. Dokchitser, *Root numbers of elliptic curves in residue characteristic 2.* Preprint, 2006, available online at `arxiv.org/math/0612054`

[22] W. Duke, *Elliptic curves with no exceptional primes.* C. R. Acad. Sci. Paris Sér. I Math. **325** (1997), no. 8, 813–818.

[23] W. Duke, E. Kowalski, *A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations.* Invent. Math. **139** (2000), no. 1, 1–39.

[24] N. Dummigan, P. Martin, M. Watkins, *Euler factors and local root numbers for symmetric powers of elliptic curves.* Preprint, 2006.

[25] T. A. Fisher, *Finding rational points on elliptic curves using 6-descent and 12-descent.* Preprint, 2007. Available from `www.dpmms.cam.ac.uk/~taf1000/papers/sixandtwelve.html`

[26] É. Fouvry, M. Nair, G. Tenenbaum, *L'ensemble exceptionnel dans la conjecture de Szpiro.* (French) [The exceptional set in Szpiro's conjecture]. Bull. Soc. Math. France **120** (1992), no. 4, 485–506.

[27] G. Frey (ed.), *On Artin's conjecture for odd 2-dimensional representations.* Lecture Notes in Mathematics, **1585**, Springer-Verlag, Berlin, 1994.

[28] D. Goldfeld, *Conjectures on elliptic curves over quadratic fields.* In *Number theory, Carbondale 1979* (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), M. B. Nathanson, ed., Lect. Notes in Math. **751**, Springer-Verlag, Berlin (1979), 108–118,

[29] R. Greenberg, *On the Birch and Swinnerton-Dyer conjecture.* Invent. Math. **72** (1983), no. 2, 241–265.

[30] H. A. Helfgott, *On the behaviour of root numbers in families of elliptic curves.* Preprint (2004), available online at `arxiv.org/math/0408141`

[31] M. Hindry, *Why is it difficult to compute the Mordell-Weil group?* Preprint, available from `www.math.jussieu.fr/~hindry/MW-size.pdf`

[32] N. M. Katz, *Moments, monodromy, and perversity: a Diophantine perspective.* Annals of Mathematics Studies, **159**. Princeton University Press, Princeton, NJ, 2005. Online from `www.math.princeton.edu/~nmk/mmp`

[33] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society Colloquium Publications, vol. 45, American Mathematical Society, Providence, RI, 1999.

[34] J. P. Keating, N. C. Snaith, *Random matrix theory and* $\zeta(1/2+it)$, and *Random matrix theory and L-functions at* $s=1/2$. Comm. Math. Phys. **214** (2000), no. 1, 57–89, 91–110.

[35] S. Kobayashi, *The local root number of elliptic curves with wild ramification.* Math. Ann. **323** (2002), no. 3, 609–623.

[36] S. Lang, *Conjectured Diophantine estimates on elliptic curves.* In *Arithmetic and geometry, Vol. I. Arithmetic.* Papers dedicated to I. R. Shafarevich on the occasion of his sixtieth birthday. Edited by M. Artin and J. Tate. Progress in Mathematics **35** (1983), Birkhäuser Boston, Boston, MA, 155–171.

[37] C. Liu, L. Xu, *The vanishing order of certain Hecke L-functions of imaginary quadratic fields.* J. Number Theory **108** (2004), no. 1, 76–89. Preliminary version online at `arxiv.org/math/0210313`

[38] V. A. Marčenko, L. A. Pastur, *Distribution of eigenvalues in certain sets of random matrices.* (Russian). Mat. Sb. (N.S.) **72**, no. 114, (1967), 507–536. English translation appears in: V. A. Marchenko, L. A. Pastur, *Distribution of eigenvalues for some sets of random matrices,* Math. USSR-Sb. **1** (1967), 457–483.

[39] P. Martin, M. Watkins, *Symmetric powers of elliptic curve L-functions.* In *Algorithmic Number Theory*, Proceedings of the 7th International Symposium, ANTS-VII, Berlin, Germany, July 2006, edited by F. Hess, S. Pauli, and M. Pohst, Springer Lecture Notes in Computer Science, **4076** (2006), 377–392. See `arxiv.org/math/0604095`

[40] M. L. Mehta, *Random matrices.* Third edition. Pure and Applied Mathematics (Amsterdam), 142. Elsevier/Academic Press, Amsterdam, 2004.

[41] P. Michel, *Rang moyen de familles de courbes elliptiques et lois de Sato-Tate.* (French) [Average rank of families of elliptic curves and Sato-Tate laws]. Monatsh. Math. **120** (1995), no. 2, 127–136.

[42] S. J. Miller, *One- and two-level densities for rational families of elliptic curves: evidence for the underlying group symmetries.* Compos. Math. **140** (2004), no. 4, 952–992.

[43] A. Pacetti. G. Tornaría, *Computing central values of twisted L-series: the case of composite levels,* to appear in Exp. Math., `arxiv.org/math/0607008`

[44] D. Ramakrishnan, F. Shahidi, *Siegel modular forms of genus 2 attached to elliptic curves.* Math. Res. Lett. **14** (2007), no. 2, 315–332.

[45] M. Rapoport, N. Schappacher, P. Schneider (ed.), *Beilinson's conjectures on special values of L-functions.* Perspectives in Mathematics, **4**, Academic Press, Inc., Boston, MA (1988).

[46] F. Rodriguez Villegas, D. Zagier, *Square roots of central values of Hecke L-series.* In *Advances in number theory*, Proceedings of the Third Conference of the Canadian Number Theory Association held at Queen's University, Kingston, Ontario, August 18–24, 1991. Edited by F. Q. Gouvêa and N. Yui, Oxford Sci. Publ., Oxford Univ. Press, New York (1993), 81–99.

[47] D. E. Rohrlich, *The nonvanishing of certain Hecke L-functions at the center of the critical strip,* Duke Math. J. **47** (1980), no. 1, 223–232, and *On the L-functions of canonical Hecke characters of imaginary quadratic fields. I. II.,* Duke Math. J. **47** (1980), no. 3, 547–557, **49** (1982), no. 4, 937–942 (with H. L. Montgomery).

[48] H. Rosson, G. Tornaría, *Central values of quadratic twists for a modular form of weight 4.* In *Ranks of Elliptic Curves and Random Matrix Theory*, edited by J. B. Conrey, D. W. Farmer, F. Mezzadri, and N. C. Snaith, London Mathematical Society Lecture Note Series **341**, Cambridge University Press (2007), 315–321.

[49] W. Schwarz, *Einige Anwendungen Tauberscher Sätze in der Zahlentheorie.* (German) [Some number theoretic applications of Tauberian theorems]. J. Reine Angew. Math. **219** (1965), 157–179.

[50] C. L. Siegel, *A mean value theorem in the geometry of numbers.* Ann. of Math. (2) **46** (1945), 340–347.

[51] J. H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, New York, 1992, Corrected reprint of the 1986 original.

[52] S. K. M. Stamminger, *Explicit 8-Descent on Elliptic Curves.* Doctoral Thesis, International University Bremen, under M. Stoll, 2005. See `www.iu-bremen.de/research/dissertations` or possibly directly from `www.iu-bremen.de/phd/files/1135329284.pdf`

[53] W. A. Stein, M. Watkins, *A Database of Elliptic Curves—First Report.* In *Algorithmic number theory*, ANTS-V (Sydney 2002), edited by C. Fieker and D. R. Kohel, Springer Lecture Notes in Computer Science, **2369** (2002), 267–275.

[54] J. T. Tate, *Algebraic cycles and poles of zeta functions.* In *Arithmetical Algebraic Geometry.* Proceedings of a Conference at Purdue Univ. 1963. Edited by O. F. G. Schilling. Harper & Row, New York (1965), 93–110.

[55] R. Taylor, A. Wiles, *Ring-theoretic properties of certain Hecke algebras.* Ann. of Math. (2) **141** (1995), no. 3, 553–572.

[56] G. Tenenbaum, *La méthode du col en théorie analytique des nombres.* (French) [The method of steepest descent in analytic number theory]. Séminaire de Théorie des Nombres, Paris 1986–87 (C. Goldstein, ed.), Progr. Math., **75**, Birkhäuser Boston, Boston, MA (1988), 411–441.

[57] J.-L. Waldspurger, *Sur les coefficients de Fourier des formes modulaires de poids demi-entier.* (French) [On the Fourier coefficients of modular forms of half-integral weight]. J. Math. Pures Appl. (9) **60** (1981), no. 4, 375–484.

[58] D. Whitehouse, *Root numbers of elliptic curves over 2-adic fields.* Preprint, 2004.

[59] E. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions.* Ann. of Math. (2) **62** (1955), 546–564.

[60] A. Wiles, *Modular elliptic curves and Fermat's last theorem.* Ann. of Math. (2) **141** (1995), no. 3, 443–551.

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

[62] M. P. Young, *Low-lying zeros of families of elliptic curves.* J. Amer. Math. Soc. **19** (2006), no. 1, 205–250 (electronic).