# Some Heuristics about Elliptic Curves

Mark Watkins

CONTENTS



We give some heuristics for counting elliptic curves with certain properties. In particular, we rederive 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 rederive 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 for enumeration 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. On 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.

2000 AMS Subject Classification: Primary 14H52, 14G10

Keywords: Elliptic curves, asymptotic count, vanishing $L$-function


1058-6458/2008 $0.50 per page

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 rederive the Brumer–McGuinness heuristic [Brumer and McGuinness 90] 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 $\mathbb{Q}$ whose minimal (integral) discriminant has absolute value less than $X$ is asymptotically given by (splitting into positive and negative discriminant)*

$$A_\pm(X) \sim \frac{\alpha_\pm}{\zeta(10)} X^{5/6},$$

*where*

$$\alpha_\pm = \frac{\sqrt{3}}{10} \int_{\pm 1}^{\infty} \frac{dx}{\sqrt{x^3 \mp 1}}.$$

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

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

since $\alpha_+ = \frac{1}{3}B\left(\frac{1}{2},\frac{1}{6}\right)$ and $\alpha_- = B\left(\frac{1}{2},\frac{1}{3}\right)$.

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

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

with

$$a_1, a_3 \in \{0,1\} \quad \text{and} \quad |a_2| \leq 1.$$

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 $\left|64a_4^3 + 432a_6^2\right| \leq X$, where we note 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 $\iint_U du_4\,du_6$ for the region $U$ given by $\left|64u_4^3 + 432u_6^2\right| < X$. After splitting into two parts based on the sign of the discriminant and performing the $u_4$-integration, we get

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

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

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

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

$$\frac{3}{5}\int_1^\infty \frac{dx}{\sqrt{x^3-1}} = \frac{1}{5}B\left(\frac{1}{2},\frac{1}{6}\right)$$

and

$$\frac{3}{5}\int_{-1}^\infty \frac{dx}{\sqrt{x^3+1}} = \frac{3}{5}B\left(\frac{1}{2},\frac{1}{3}\right).$$

This counts all models of curves; if we eliminate nonminimal models, for which we have $p^4 \mid c_4$ and $p^{12} \mid \Delta$ for some prime $p$, we expect to accrue an extra factor[2] of

[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 integrate again 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 nonminimality at $p$ occurs when $c_4/p^4$ and $c_6/p^6$ satisfy the Connell congruences [Connell 91] we mention below. A referee points out that Brumer [Brumer 92, Lemma 4.3] obtains this same $1/\zeta(10)$ factor via sieving.

$1/\zeta(10)$. From this we get the conjecture of Brumer and McGuinness presented above.

## 3. COUNTING CURVES OF EVEN PARITY WHOSE CENTRAL $L$-VALUE VANISHES

Due to work of Wiles [Wiles 95, Taylor and Wiles 95] and others [Diamond 96, Conrad et al. 99, Breuil et al. 01], we know that elliptic curves over $\mathbb{Q}$ 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. (See [Silverman 92, Sections 15-16] for the definitions of the conductor $N$ and $L$-function $L(E,s)$ of an elliptic curve $E$.)

We now try to count elliptic curves $E$ with even parity for which $L(E,1) = 0$. Throughout this section, $E$ will denote 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 will 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 that 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 $E/\mathbb{Q}$ 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] A referee reports that a preprint of [Helfgott 04] discusses this question for parameterized families.

### 3.1 Random Matrix Theory

Originally developed in mathematical statistics by Wishart [Wishart 28] in the 1920s and then in mathematical physics (especially the spectra of highly excited nuclei) by Wigner [Wigner 55], Dyson, Mehta, and others (particularly [Marčenko and Pastur 67]), random matrix theory [Mehta 04] 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 $\zeta$-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 a delicate one (see particularly [Conrey et al. 05, Section 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 square-free $d$. The work (most significantly a monodromy computation) of Katz and Sarnak [Katz and Sarnak 99] 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 modeled 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 $\mathrm{SO}(2M)$ 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 on frequency of small $L$-values and discretization (similar to the below) then gives that the number of $d$ with $E_d$ of even parity, $|d| < D$, 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 [Delaunay and Watkins 07]) depending on the splitting behavior of the cubic polynomial $x^3 + Ax + B$, while $c_E$ has yet to be determined explicitly. Various data have been given by Rubinstein [Conrey et al. 06] to lend credence to this guess. We note that the exponent $3/4$ can already be suspected from

work of Waldspurger [Waldspurger 81], which relates[4] to $E$ a modular form of weight $\frac{3}{2}$ whose $d$th coefficient $c(d)$ is such that $c(d)^2$ is proportional to $L(E_d, 1)$.

In particular, the Ramanujan conjecture predicts that the $d$th coefficient is bounded by $d^{1/4+\epsilon}$, and so, assuming a reasonable distribution, the probability that it is zero is about one 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 for 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 $\mathrm{SO}(2M)$ 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 [Keating and Snaith 00], 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 [Keating and Snaith 00] and then using Mellin inversion, as $t \to 0$ we have (see [Conrey et al. 02, (21)]) that (here $\frac{3}{8}$ is $\binom{-1/2}{2}$)

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

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, since 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 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 on 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

$$\begin{aligned} \alpha_A(E) &= \prod_p F(p) \\ &= \prod_p \left(1 - \frac{1}{p}\right)^{3/8} \left(\frac{p}{p+1}\right) \\ &\quad \times \left(\frac{1}{p} + \frac{L_p(1/p)^{-1/2}}{2} + \frac{L_p(-1/p)^{-1/2}}{2}\right), \end{aligned} \tag{3–2}$$

where $L_p(X) = (1 - a_p X + pX^2)^{-1}$ when $p \nmid \Delta$ and $L_p(X) = (1 - a_p X)^{-1}$ otherwise; see [Conrey et al. 02, (10)] evaluated at $k = -\frac{1}{2}$, though that equation is wrong at primes that divide the discriminant; see [Conrey et al. 07, (20)], where $Q$ should be taken to be 1. Note that the Sato–Tate conjecture [Tate 65] implies that $a_p^2$ is $p$ on average, and this implies that the above Euler product converges.[6]

### 3.2 Discretization of the $L$-Value Distribution

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 $\Omega_{\mathrm{re}}(E)$ 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 $L(E,1) = 0$. (For precise definitions of the Tamagawa numbers, torsion group, periods, and Shafarevich–Tate group, see [Silverman 92], though below we give a brief description of some of these.)

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 [Birch and Swinnerton-Dyer 63, Birch and Swinnerton-Dyer 65], which asserts that

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

Our discretization[7] will be that

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

[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 [Katz 05, Section 12.8] has some results in the function-field case related to (say) generalized or usual Weierstrass families of elliptic curves. A referee also notes that there is work of Miller [Miller 04] and Young [Young 06] in this direction.

[6] A referee points out that Birch [Birch 68] has already shown (using the Selberg trace formula) that the Sato–Tate distribution holds with respect to all curves over a fixed field $\mathbb{F}_p$, while Michel [Michel 95] has the best results in the case of one-parameter families.

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

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

Using (3–1), 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

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

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 [Stein and Watkins 02] uses the work of Connell [Connell 91] 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 nonminimal 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*

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

*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 that the torsion is trivial; below we will give these heuristics 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.

[8]Note that our methods do not readily generalize to positive rank, since there is no apparent way to model the heights of points (and thus the regulator). A referee points out that Lang [Lang 83] gives some bounds, and perhaps suggests a distribution, but this seems insufficient for our purposes.

- 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 Birch–Swinnerton-Dyer (BSD) conjecture.
- There is independence among 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

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

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

$$e_1-e_2=\Delta^{1/6}\lambda \quad\text{and}\quad e_2-e_3=\Delta^{1/6}\mu,$$

so that we have $\mu\lambda(\lambda+\mu)=\frac{1}{4}$, while

$$e_1=\frac{\Delta^{1/6}}{3}(\mu+2\lambda),\quad e_2=\frac{\Delta^{1/6}}{3}(\mu-\lambda),$$

and

$$e_3=-\frac{\Delta^{1/6}}{3}(2\mu+\lambda).$$

Thus we get

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

and

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

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

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

for some constant $\tilde{c} > 0$, where

$$\lambda = \frac{\sqrt{\mu^4+\mu}-\mu^2}{2\mu}.$$

Thus the variables are nicely separated, and since the $\mu$-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 parameterizations 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 $\Delta$, 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 $p$ much smaller than $X^{1/3}$. We estimate the probability that $p^k \mid \Delta$ 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 $\mathbb{F}_p$, and has $(p-1)$ nonsingular $\mathbb{F}_p$-solutions that 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 $p^k \mid \Delta$ 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 $p^4 \mid c_4$ and $p^6 \mid c_6$ will not be given by their minimal model; indeed, we want to exclude these curves, and so we 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 on conditioning on 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 $p^k \parallel D$ for some $k > 0$? In terms of Kodaira symbols, this is the case of $\mathrm{I}_k$. 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. If we assume these conditions and work 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 $p^k \mid D$. 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 $p^k \mid D$, which gives $1/p^k$ for the probability that $p^k \parallel D$. In symbols, we have that (for $p \ge 5$ and $k \ge 1$)

$$\operatorname{Prob}\left[p^k \parallel (c_4^3 - c_6^2) \;\middle|\; p \nmid c_4, p \nmid c_6\right] = 1/p^k.$$

Including the conditional probability for minimal models, we get

$$P_m(p,k) = \frac{1}{p^k}\left(1-\frac{1}{p^{10}}\right)^{-1}\left(1-\frac{1}{p}\right)^2,$$

for $p \ge 5$ and $k \ge 1$. Note that summing this over $k \ge 1$ 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 $p \ge 5$ and $p^k \parallel D$ for some $k > 0$? We shall temporarily ignore the factor $\kappa_p = (1-1/p^{10})^{-1}$ from nonminimal models and include it at the end. We must have that $p \mid c_4$ and $p \mid c_6$, and thus get that $k \ge 2$. For $k = 2, 3, 4$, which correspond to Kodaira symbols II, III, and IV respectively, the computation is not too bad: we get that $p^2 \parallel D$ exactly when $p \mid c_4$ and $p \parallel c_6$, so that the probability is

$$\frac{1}{p}\cdot\frac{1-1/p}{p} = \frac{1-1/p}{p^2};$$

for $p^3 \parallel D$ we need $p \parallel c_4$ and $p^2 \mid c_6$ and thus get

$$\frac{1-1/p}{p}\cdot\frac{1}{p^2} = \frac{1-1/p}{p^3}$$

for the probability; and for $p^4 \parallel D$ we need $p^2 \mid c_4$ and $p^2 \parallel c_6$, and so get

$$\frac{1}{p^2} \cdot \frac{1-1/p}{p^2} = \frac{1-1/p}{p^4}$$

for the probability. Note that the case $k = 5$ cannot occur. Thus we have (for $p \geq 5$) the formula

$$P_a(p,k) = \frac{1}{p^k}\left(1 - \frac{1}{p^{10}}\right)^{-1}\left(1 - \frac{1}{p}\right)$$

for $k = 2, 3, 4$.

More complications occur for $k \geq 6$, where now we split into two cases depending on whether additive reduction persists on taking the quadratic twist by $p$. This occurs when $p^3 \mid c_4$ and $p^4 \mid c_6$, and we denote by $P_a^n(p,k)$ the probability that $p^k \parallel D$ in this subcase. Just as above, we get that

$$P_a^n(p,k) = \frac{1}{p^{k-1}}\left(1 - \frac{1}{p^{10}}\right)^{-1}\left(1 - \frac{1}{p}\right)$$

for $k = 8, 9, 10$. These are respectively the cases of Kodaira symbols IV$^\star$, III$^\star$, and II$^\star$. 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 $p^6 \mid D$ with either $p^2 \parallel c_4$ or $p^3 \parallel c_6$. First we consider curves for which $p^7 \mid D$, and these have multiplicative reduction at $p$ upon twisting. In particular, these curves have $p^2 \parallel c_4$ and $p^3 \parallel c_6$, and the probability of this is

$$\frac{1-1/p}{p^2} \cdot \frac{1-1/p}{p^3}.$$

Consider $k \geq 7$. We then take $c_4/p^2$ and $c_6/p^3$ both modulo $p^{k-6}$, and get that $p^{k-6} \parallel (D/p^6)$ with probability $1/p^{k-6}$ in analogy with the above. So we get that

$$P_a^t(p,k) = \left(1 - \frac{1}{p^{10}}\right)^{-1} \frac{(1-1/p)^2}{p^{k-1}}$$

for $k \geq 7$. This corresponds to the case of $\mathrm{I}^\star_{k-6}$.

Finally, for $p^6 \parallel D$ (which is the case $\mathrm{I}^\star_0$) we get a probability of

$$\frac{1}{p^2} \cdot \frac{1}{p^3}$$

that $p^2 \mid c_4$ and $p^3 \mid c_6$, and since there are $p$ points mod $p$ on the auxiliary curve

$$\left(\frac{c_4}{p^2}\right)^3 \equiv \left(\frac{c_6}{p^3}\right)^2 \pmod{p},$$

we get a conditional probability of $(p^2 - p)/p^2$ that $p^6 \parallel D$. So we get that

$$P_a^t(p,6) = \frac{1}{p^5}\left(1 - \frac{1}{p^{10}}\right)^{-1}\left(1 - \frac{1}{p}\right).$$

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

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

and

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

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

$$\sum_{k=0}^{\infty} \left[P_m(p,k) + P_a^n(p,k) + P_a^t(p,k) + P_0(p,k)\right] = 1,$$

since 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$^\star$, III$^\star$, and II$^\star$ give $\kappa_p(1/p^7 - 1/p^{10})$; the cases of $\mathrm{I}^\star_k$ summed for $k \geq 1$ give $\kappa_p(1 - 1/p)/p^6$; the case of $\mathrm{I}^\star_0$ 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 $|\Delta| \leq X$ act like those in a big box with $|c_4| \leq X^{1/3}$ and $|c_6| \leq X^{1/2}$, 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 $p \mid D$ with $p \geq 5$, and compute the local Tamagawa number $t(p)$; this can be done as in [Cohen 93, Algorithm 7.5.1] (with a corrected line 2 of step 3 in early printings).

When $E$ has multiplicative reduction at $p$ and $p^k \parallel D$, then $t(p) = k$ if $-c_6$ is square mod $p$, and otherwise, $t(p) = 1, 2$, depending on whether $k$ is odd or even. So the average of $\sqrt{t(p)}$ for this case is

$$\epsilon_m(k) = \frac{1}{2}\left(1 + \sqrt{k}\right) \quad \text{or} \quad \frac{1}{2}\left(\sqrt{2} + \sqrt{k}\right)$$

for $k$ odd or even respectively.

When $E$ has potentially multiplicative reduction at $p$ with $p^k \parallel D$, 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 $\Delta/p^k$ 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 $p^6 \parallel D$, we have that $t(p) = 1, 2, 4$, corresponding to whether the cubic

$$x^3 - \frac{27c_4}{p^2}x - \frac{54c_6}{p^3}$$

has $0, 1, 3$ roots modulo $p$. So the average of $\sqrt{t(p)}$ is

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

in this case.

For the remaining cases, when $p^2 \parallel D$ or $p^{10} \parallel D$ we have $t(p) = 1$, while when $p^3 \parallel D$ or $p^9 \parallel D$ we have $t(p) = 2$. Finally, when $p^4 \parallel D$ we have $t(p) = 3, 1$, depending on whether $-6c_6/p^2$ is square mod $p$, and similarly when $p^8 \parallel D$ 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

$$\epsilon_m(k) = \begin{cases} \frac{1}{2}(1+\sqrt{k}), & k \text{ odd}, \\ \frac{1}{2}(\sqrt{2}+\sqrt{k}), & k \text{ even}, \end{cases} \tag{3–3}$$

and

$$\epsilon_a^t(p,k) = \begin{cases} \frac{2}{3} + \frac{\sqrt{2}}{2} - \frac{1}{3p}, & k = 6, \\ \frac{1}{2}(\sqrt{2}+\sqrt{4}), & k \geq 7, \end{cases} \tag{3–4}$$

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

$$K(p) = \sum_{k=0}^{\infty} \big[\epsilon_m(k)P_m(p,k) + \epsilon_a^n(k)P_a^n(p,k) + \epsilon_a^t(p,k)P_a^t(p,k) + P_0(p,k)\big] \tag{3–5}$$

and assume that all the primes act independently to get that the expected global[9] Tamagawa number is

$$\beta(\sqrt{\tau}) = \prod_p K(p).$$

The convergence of this product follows from an analysis of the dominant $k = 0, 1, 2$ terms of (3–5), which gives a behavior of $1 + O(1/p^2)$. So we get that the Tamagawa product is a constant on average, 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 (3–2), we similarly assume that each prime acts independently.[10] We then compute the average value for each prime by calculating the distribution of $F(p)$ for all the curves modulo $p$ (including those with singular reduction, and again making the slight adjustment for nonminimal 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 [Tate 65], 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 the following.

**Heuristic 4.1.** *The number $B(X)$ of elliptic curves over $\mathbb{Q}$ 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 [Cremona 06] do not coincide with this heuristic; in fact, the growth seems almost linear in the conductor, for taking the curves with conductor in the range 40,000–130,000 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 $\frac{5}{6}$. An upper bound of $B(X) \ll_\epsilon X^{1+\epsilon}$ is explicated in [Duke and Kowalski 00, Section 3.1].

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

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

**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 square-free kernel of the discriminant, by way of explanation we first consider the behavior of $f(n) = n/\mathrm{sqfree}(n)$. The probability that $f(n) = 1$ is given by the probability that $n$ is square-free, 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 square-free and coprime to $p$. The probability that $p^{m+1} \parallel n$ is $(1 - 1/p)/p^{m+1}$, and given this, the conditional probability that $\big(n/p^{m+1}\big)$ is square-free is $(6/\pi^2) \cdot (1 - 1/p^2)^{-1}$. Extending this multiplicatively beyond prime powers, we get that

$$\begin{aligned}&\mathrm{Prob}\big[n/\mathrm{sqfree}(n) = q\big]\\ &\quad= \frac{6}{\pi^2} \prod_{p^m \| q} \frac{1/p^{(m+1)}}{(1 + 1/p)} = \frac{6}{\pi^2}\frac{1}{q}\prod_{p|q}\frac{1}{p+1}.\end{aligned}$$

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

### 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 $p \ge 5$, the probability that $p^v \parallel (D/N)$ is given by the following: the probability that $E$ has multiplicative reduction at $p$ and $p^{v+1} \parallel D$, that is, $P_m(p, v + 1)$; plus the probability that $E$ has additive reduction at $p$ and $p^{v+2} \parallel D$, 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 $p = 2, 3$)

$$\begin{aligned}&\mathrm{Prob}\big[D/N = q\big] \qquad (4\text{–}1)\\ &\quad= \prod_p \big[P_m(p, 1 + v) + P_a(p, 2 + v) + P_0(p, v)\big].\end{aligned}$$

We emphasize that this probability is with respect to curve-ordering by discriminant (as in the last section), and as previously, we have assumed that the primes act independently, that curves with $|\Delta| \le X$ 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

$$\begin{aligned}\sum_{E:N_E\le X} 1 = \sum_{q=1}^{\infty} \sum_{\substack{E:N\le X\\ D/N=q}} 1 &\approx \sum_{q=1}^{\infty} \sum_{E:D\le qX} \mathrm{Prob}\big[D/N = q\big]\\ &\sim \sum_{q=1}^{\infty} \alpha \cdot (qX)^{5/6} \cdot \mathrm{Prob}\big[D/N = q\big], \qquad (4\text{–}2)\end{aligned}$$

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

To show that the last sum in (4–2) does indeed converge, we get an upper bound for the probability in (4–1). We have that $P_m(p, v + 1) \le 1/p^{v+1}$ and $P_a(p, v + 2) \le 2/p^{v+1}$, which implies

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

We then estimate

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

and the last product is seen to be convergent on comparison to $\zeta\left(\frac{7}{6}\right)^3$. Thus we shown that the last sum in (4–2) converges, so that Heuristic 4.1 follows.

We note that Fouvry, Nair, and Tenenbaum [Fouvry et al. 92] have shown that the number of minimal models with $D \le X$ is at least $cX^{5/6}$, and that the number of curves with $D \le X$ with Szpiro ratio $\frac{\log D}{\log N} \ge \kappa$ 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 (4–1) and using the $\epsilon$ defined in (3–3) and (3–4), we find that this is given by

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

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

[11] The saddle-point method as indicated in [Tenenbaum 88] and [Burris and Yeats 05] might be applicable, but it appears to involve quite careful estimation to achieve an asymptotic rather than a log-asymptotic. It was pointed out to us by G. Tenenbaum that [Schwarz 65] improves on the result of [de Bruijn 62], though the result is not that explicit.

[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 that they may be ignored.

### 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

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

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 $\sum_p \frac{1}{p \log p} \ll 1$)

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

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

$$\begin{aligned}
&\frac{\hat{\alpha}_R \bar{\alpha}_A}{3456\,\zeta(10)} \\
&\quad \times \iint\limits_{\sqrt{X} \leq \frac{u_4^3 - u_6^2}{1728} \leq X} \Omega_{\mathrm{re}}(E) \\
&\qquad \times \left[\sum_{q < \Delta} \eta(q) (\log \Delta/q)^{3/8} \mathrm{Prob}\big[D/N = q\big]\right] du_4\, du_6.
\end{aligned}$$

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 have restricted to $\Delta > \sqrt{X}$). So the bracketed term becomes the desired

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

on 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 $0 < \Delta < qX$, $D/N = q$ and $L(E,1) = 0$ is given by

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

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 (after arguing similarly for curves with negative discriminant).

### 4.5 Relations to Other Work

It is proposed by Hindry [Hindry 05, Conjectures 5.4 and 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 that the product of the regulator and $\#Ш$ cannot be too small. We view this as unlikely; already in the rank-zero 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 discretization 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

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

and note that the torsion and Tamagawa contributions are small compared to the reciprocal of the real pe-

riod.[13] A generalization 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 a 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 predicts 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 parameterization

$$c_4 = (t + 16)(t + 64)T^2 \quad \text{and} \quad c_6 = (t - 8)(t + 64)^2T^3$$

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 square-free kernel of $pq(p + 64q)$ smaller than $X$ in 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 [Duke 97] 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 data sets of [Brumer and McGuinness 90] and [Stein and Watkins 02] 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 310,700 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 data sets are expected to be nearly exhaustive[15] among 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 nontrivial to get a good data set, 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 [Stein and Watkins 02], we divided the $(c_4, c_6)$ pairs into 288 congruence classes with

$$(\tilde{c}_4, \tilde{c}_6) = \big(c_4 \bmod 576,\ c_6 \bmod 1728\big).$$

[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 relies only on the size of the real period.

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

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 10,000 parameter selections

$$(c_4, c_6) = \big(576(1000 + i) + \tilde{c}_4, 1728(100000 + j) + \tilde{c}_6\big)$$

for $(i, j) \in [1..10] \times [1..1000]$, and then of these 2,880,000 curves, took the 89,913 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

$$(c_4, c_6) = \big(576(-883 + i) + \tilde{c}_4, 1728(100000 + j) + \tilde{c}_6\big)$$

for $(i, j) \in [1..10] \times [1..1000]$, took the subset of 89,749 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 data sets, and indeed was noted in [Brumer and McGuinness 90].[16] We do not average the results from positive and negative discriminants; the Brumer–McGuinness conjecture, 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 curves of prime positive discriminant it was 44.1% for Brumer–McGuinness and 41.7% for Stein–Watkins, but only 36.0% for our data set; for curves of negative discriminant and 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 role, and so we wish to understand how our choice has affected it. Indeed, as was pointed out to us by X.-F. Roblot, the variation of real period from enumerating in a large $(c_4, c_6)$-box is quite different from the result of enumerating by discriminant.

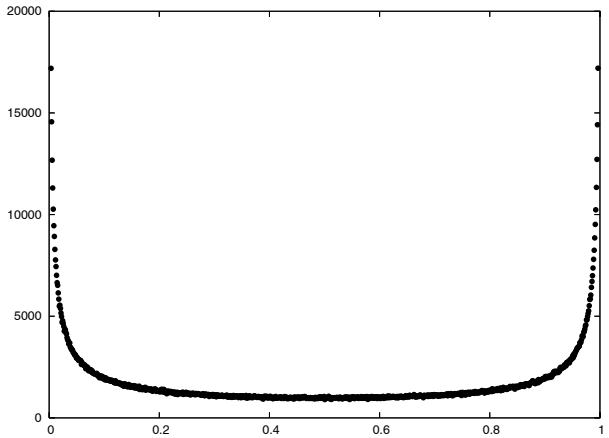


**FIGURE 1**. $\Delta > 0$: Curve distribution as a function of $t$.

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 data sets, 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]

$$\Omega_{\mathrm{re}}\Delta^{1/12} = \frac{2^{1/3}\pi(t - t^2)^{1/6}}{\mathrm{agm}(1, \sqrt{t})}.$$

The curves we considered all had $0.617 < t < 0.629$.

However, in analogy to our consideration of curves ordered by conductor, before counting curves with extra rank we should first simply count curves. Figure 1 indicates the distribution of the root quotient $t$ for the curves of prime (positive) discriminant and even parity from the Stein–Watkins database (more than two million curves

[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 maximized at $t \approx 0.0388505246188$ with a maximum just below 4.414499094.

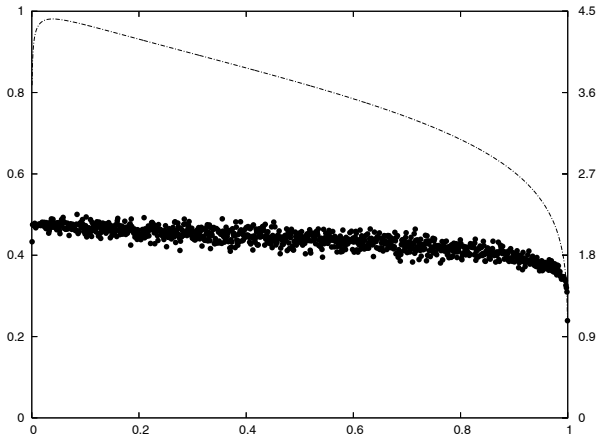


**FIGURE 2**. $\Delta > 0$: Positive-rank frequency as a function of the root quotient $t$, and $\Omega_{\mathrm{re}}\Delta^{1/12}$ as a function of $t$.

meet the criteria). The $x$-axis is divided up into bins of size 1/1000; there are more than one hundred times as many curves with $t < 0.001$ as 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 one thousand curves in some of our bins, we do not get such a nice graph. Note that the leftmost and especially the rightmost dots are much below their nearest neighbors and that 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% among all curves of even parity and prime discriminant in the Stein–Watkins database, and is 42.8% for the 12,324 such curves with $0.617 < t < 0.629$. The function plotted (labeled on the right axis) in Figure 2 is

$$\Omega_{\mathrm{re}}\Delta^{1/12} = \frac{2^{1/3}\pi(t - t^2)^{1/6}}{\mathrm{agm}(1, \sqrt{t})}$$

as a function of $t$, and note that this goes to zero 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.

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

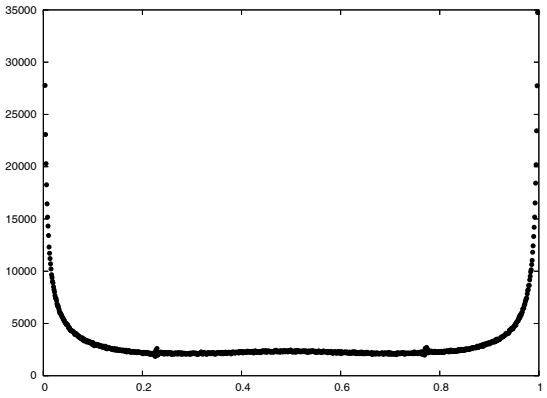


**FIGURE 3**. $\Delta < 0$: Distribution of curves as a function of $C$.

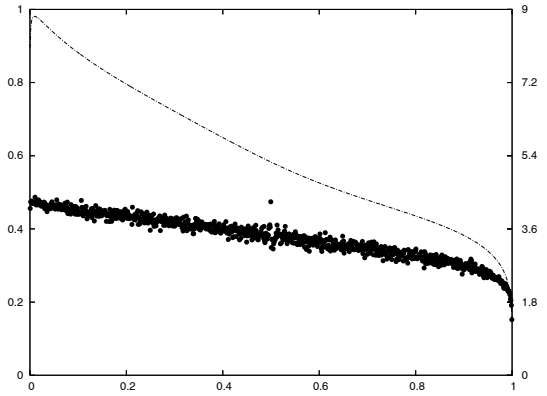


**FIGURE 4**. $\Delta < 0$: Positive-rank frequency as a function of $C$, and $\Omega_{\mathrm{re}}|\Delta|^{1/12}$ as a function of $C$.

$\tilde{r} = r + b_2/12$ and $c = \tilde{r}/Z$, we then have[18]

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

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

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 $C = \frac{1}{2}$. Indeed,

[18] This is maximized 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 $\pm 1$) tend to have $C$ in these ranges (for our discriminant range), and this causes instability in the counting function.

the hundred closest curves with $C < \frac{1}{2}$ all have positive rank; this breaks down when the barrier $\frac{1}{2}$ is crossed. 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 $C = \frac{1}{2}$ 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% among all curves of even parity and negative prime discriminant in the Stein–Watkins database, and is 37.0% for the 4695 such curves with $0.555 < C < 0.557$.

### 6.3 Other Considerations

The idea that the "probability" that a curve of even parity possesses positive rank should be proportional to $\sqrt{\Omega_{\mathrm{re}}}$ is perhaps overly simplistic; in particular, it is not borne out too precisely by the Stein–Watkins data set. We consider curves of positive prime discriminant with even parity; for those with $0.64 < \Omega_{\mathrm{re}} < 0.65$ we have 78,784 curves, of which 45.9% have positive rank, while of the 9872 with $0.32 < \Omega_{\mathrm{re}} < 0.325$, 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 $0.32 < \Omega_{\mathrm{re}} < 0.325$ that we have not considered. This, however, is in contrast 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 $L(E, 1) = 0$ than our estimate of $c\sqrt{\Omega_{\mathrm{re}}}$ would suggest. Indeed, it might be argued (perhaps due to arithmetic considerations, or perhaps explicit formulas for the zeros of $L$-functions) that curves with such small discriminant cannot realize their nominal expected frequency of positive rank.

Unfortunately, we cannot do much to quantify these musings, since the effect would likely be in a secondary term, making it difficult to detect experimentally. Note also that a relative depression of rank for curves of small discriminant 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 2,143,079 curves of (analytic) rank 2 of prime conductor less than $10^{10}$ in the Stein–Watkins database.[20] J. E. Cremona ran his `mwrank` program [Cremona 05] on all these curves, and it was successful in provably finding the Mordell–Weil group for 2,114,188 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 18,000 had 2-Selmer rank greater than 2.

We then used the `FourDescent` machinery of MAGMA, 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 [Stamminger 05] will complete (assuming the generalized Riemann hypothesis 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 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 normalizing $P$ 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 $\mathrm{SL}_2(\mathbb{Z})$. 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 $0 \leq x \leq \frac{1}{2}$.

With no other guidance, we might expect that the obtained distribution for the $z$ is given by[22] the Haar mea-

[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 minimal generator of height more than 600, so that other methods are likely to be needed in order to find it. Indeed, T. A. Fisher [Fisher 07] has recently used 6-descent and 12-descent to find the missing generators on these 35 curves.

[22] Siegel [Siegel 45] similarly uses Haar measure to put a natural measure on $n$-dimensional lattices of determinant 1.

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

**TABLE 1**. Horizontal distribution of rank-2 lattices with $y \ge 1$.

sure $(dx\,dy)/y^2$. 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 $y \ge 1$, while the experimental result is about 93.5%. Furthermore, we should expect that the proportion of curves with $y \ge Y$ should die off like $1/Y$ as $y \to \infty$; however, we get that 35.6% of the curves have $y \ge 2$, only 9.7% have $y \ge 4$, while 1.97% have $y \ge 8$ and 0.35% have $y \ge 16$.

The validity of the vertical distribution data might be arguable based on concerns regarding the discriminant cutoff of our data set, but the horizontal distribution is also skewed. If we consider only curves with $y \ge 1$, then we should get uniform distribution in the $x$-aspect; however, Table 1 shows that we do not have such uniformity.

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 ($\approx 1500$ per bin) is typically ten times the experimental amount; this increases to a factor of 100 for $y \approx 70$. Overall, our data seem to imply that the lattices are not as skinny and are less orthogonal than might be guessed.

### 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 [Martin and Watkins 06] for more details about this, but mention that due to conjectures of Deligne and more generally Bloch and Beĭlinson [Rapoport et al. 88], we expect that we should have a formula similar to that of Birch and Swinnerton-Dyer, stating that

$$L(\mathrm{Sym}^{2k-1}E, k)\frac{(2\pi N)^{\binom{k}{2}}}{\Omega_+^{\binom{k+1}{2}}\Omega_-^{\binom{k}{2}}}$$

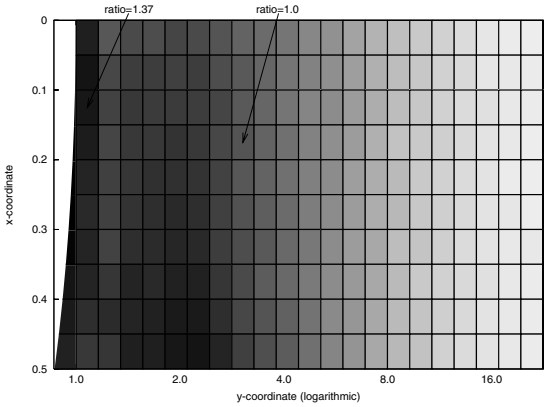


**FIGURE 5**. Density plot of ratio of experimental to conjectural counts.

should be rational with small denominator. Here, for $k$ odd, $\Omega_+$ is the real period and $\Omega_-$ the imaginary period, with this reversed for $k$ even. As noted in [Buhler et al. 97], 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 discretization as before gives that the probability that $L(\mathrm{Sym}^{2k-1}E, s)$ has even parity and that $L(\mathrm{Sym}^{2k-1}E, k) = 0$ is bounded from above by $c(\log\Delta)^{3/8} \cdot \sqrt{1/\Delta^{k^2/12}}$.Again following the analogy of the 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 $k \ge 5$ (that is, finitely many extra vanishings for the ninth 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 [Martin and Watkins 06] find a handful of curves for which the 9th, 11th, and even the 13th symmetric powers appear (up to 12 dig-

[23]Data in [Martin and Watkins 06, Table 6] are inexact, since [Cremona 06] missed curves with $6^3 \mid N$ and $90000 \le N \le 10^5$.

its of precision) to have a central zero of order 2. We find this surprising, and it casts some doubt about the validity of our methodology of modeling 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 $d$th quadratic twist, taking $d$ to be a fundamental discriminant. From an analogue of the Birch and Swinnerton-Dyer conjecture we expect to get a rational number with small denominator from the quotient[24]

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

We have that

$$\Omega_{\mathrm{im}}(E_d)^3 \Omega_{\mathrm{re}}(E_d) \approx \frac{\Omega_{\mathrm{im}}(E)^3}{d^{3/2}} \cdot \frac{\Omega_{\mathrm{re}}(E)}{d^{1/2}}$$

(with the periods reversed when $d < 0$), and so we expect the number of fundamental discriminants $|d| < D$ such that $L(\mathrm{Sym}^3 E_d, s)$ has even parity with $L(\mathrm{Sym}^3 E_d, 2) = 0$ 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 third symmetric power; generalizing 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 $L(\mathrm{Sym}^3 E_d, 2)$ or $L'(\mathrm{Sym}^3 E_d, 2)$ for all fundamental discriminants $d$ 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 nine 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 [Conrey et al. 06] does not compute $L(E_d, 1)$ directly, but rather uses the Waldspurger formula (as made explicit in works such as [Pacetti and Tornaría 08]) and then just computes the coefficients of a modular form of weight $\frac{3}{2}$ by enumerating lattice points in an ellipsoid.

[24]The contribution from the conductor actually comes from nonintegral 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.

6.6.2 CM Curves. Next we consider CM curves, for which we can compute significantly more data, but the modeling of vanishings is slightly different. Let $E$ be a rational elliptic curve with CM, and $\psi$ its Hecke character. We shall take $\psi$ to be "twist-minimal"; this is not the same as the "canonical" character of Rohrlich [Rohrlich 80, Rohrlich and Montgomery 80], but rather we just take $E$ to be a minimal (quadratic) twist. Indeed, we shall consider only 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 normalize 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 $\psi_d$ be the Hecke Grössencharakter $\psi$ 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 $\psi^{2k-1}$ to be the primitive underlying Grössencharakter.

We then expect $L(\psi^3, 2)(2\pi)/\Omega_{\mathrm{im}}(E)^3$ to be rational with small denominator. We can then use discretization 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

$$\Omega_{\mathrm{im}}(E_d)^3 \approx \frac{\Omega_{\mathrm{im}}(E)^3}{d^{3/2}},$$

we expect the number of discriminants $d$ with even parity and $L(\psi_d^3, 2) = 0$ to be 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 note that $\psi^3$ corresponds to a weight-4 modular form, so that there is a weight-$\frac{5}{2}$ Shimura lift of it whose coefficients are related to $L(\psi_d^3, 2)$ via the Waldspurger correspondence; on considering how often these coefficients should vanish, we obtain a similar heuristic.

For higher symmetric powers, we expect that

$$L(\psi^{2k-1}, k)\frac{(2\pi)^{k-1}}{\Omega_+(E)^{2k-1}}$$

is rational with small denominator, and thus 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

[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 seventh symmetric power of the Hecke character for $y^2 = x^3 + 127x$ yields a double central vanishing. See [Greenberg 83] for related information.

| | |
|---|---|
| 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$^\star$ 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$^\star$ −2047 −2235 −2280 −2407 −2443 −2563 −2824 −2831 −3127 −3135 −3523 −4119 −4179 −4191 −4323 137 141$^\star$ 229 233 281 345 469 473 492 497$^\star$ 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$^\star$ −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$^\star$ −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 |

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

| | 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**. Counts of double-order zeros for primitive twists.

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 that $K = \mathbb{Q}(i)$ when $d > 0$ is additionally primitive when $8 \parallel d$. Note also that 27a and 36a have the same symmetric cube $L$-function.

Table 3 lists our results for counts of central double zeros (to 32 digits) for the $L$-functions of the third, fifth, and seventh 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 04] has the latest results 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 67,260 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-à-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 seventh symmetric power for the 27,365th twist of 121a has a double zero, as does the third symmetric power, while the $L$-function of the twist itself has a triple zero. Also, the 22,909th twist of 36a has double zeros for its first, third, and fifth powers (note that 36a does not appear in Table 4, since the data are identical to those 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 third 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 data set is almost one hundred 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 modeling vanishings.

In [Rodriguez Villegas and Zagier 91, Section 8], the authors mention the possibility of a formula of Waldspurger type for the twists of the Hecke Grössencharakters, but it does not seem that an exact formula has ever appeared. Using the techniques devel-

[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 [Kobayashi 02, Whitehouse 04, Dokchiter and Dokchitser 06, Dummigan et al. 06] 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 $10^5$.

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

| | | | |
|---|---|---|---|
| 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 |

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

oped by Basmaji and Frey [Frey 94], we are able to compute the weight-$\frac{5}{2}$ lift for (say) the symmetric cube of Hecke Grössencharakter attached to 49a. However, since we are currently unable to write it as a twisted ternary theta series as in [Rosson and Tornaría 07], 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 [Ramakrishnan and Shahidi 07] reinterprets 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.

## ACKNOWLEDGMENTS

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 [Duke 97], and G. Tenenbaum for the reference [Schwarz 65].

## REFERENCES

[Birch 68] 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.

[Birch and Swinnerton-Dyer 63] B. J. Birch and H. P. F. Swinnerton-Dyer. "Notes on Elliptic Curves, I." *J. Reine Angew. Math.* 212 (1963), 7–25.

[Birch and Swinnerton-Dyer 65] B. J. Birch and H. P. F. Swinnerton-Dyer. "Notes on Elliptic Curves, II." *J. Reine Angew. Math.* 218 (1965), 79–108.

[Breuil et al. 01] C. Breuil, B. Conrad, F. Diamond, and R. Taylor. "On the Modularity of Elliptic Curves over $\mathbb{Q}$: Wild 3-adic Exercises." *J. Amer. Math. Soc.* 14:4 (2001), 843–939.

[de Bruijn 62] N. G. de Bruijn. "On the Number of Integers $\leq x$ Whose Prime Factors Divide $n$." *Illinois J. Math.* 6 (1962), 137–141.

[Brumer 92] A. Brumer. "The Average Rank of Elliptic Curves, I." *Invent. Math.* 109:3 (1992), 445–472.

[Brumer and McGuinness 90] A. Brumer and O. McGuinness. "The behavior of the Mordell–Weil Group of Elliptic Curves." *Bull. Amer. Math. Soc. (N.S.)* 23:2 (1990), 375–382.

[Buhler et al. 97] J. Buhler, C. Schoen, and J. Top. "Cycles, $L$-Functions and Triple Products of Elliptic Curves." *J. Reine Angew. Math.* 492 (1997), 93–133.

[Burris and Yeats 05] S. Burris and K. Yeats. "Admissible Dirichlet Series." Preprint available online (arxiv.org/math/0507487), 2005.

[Cohen 93] H. Cohen. *A Course in Computational Algebraic Number Theory*, Grad. Texts in Math., 138. New York: Springer-Verlag, 1993.

[Connell 91] I. Connell. *The Elliptic Curve Handbook*, Lecture Notes from a course taught at McGill University. Available online (www.math.mcgill.ca/connell/public/ECH1), 1991.

[Conrad et al. 99] B. Conrad, F. Diamond, and R. Taylor. "Modularity of Certain Potentially Barsotti–Tate Galois Representations." *J. Amer. Math. Soc.* 12:2 (1999), 521–567.

[Conrey et al. 02] J. B. Conrey, J. P. Keating, M. O. Rubinstein, and N. C. Snaith. "On the Frequency of Vanishing of Quadratic Twists of Modular $L$-Functions." In *Number Theory for the Millennium, I (Urbana, IL, 2000)*, pp. 301–315. Natick, MA: A K Peters, 2002.

[Conrey et al. 05] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith. "Integral Moments of $L$-Functions." *Proc. London Math. Soc. (3)* 91:1 (2005), 33–104.

[Conrey et al. 06] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith. "Random Matrix Theory and the Fourier Coefficients of Half-Integral Weight Forms." *Experimental Math.* 15:1 (2006), 67–82.

[Conrey et al. 07] 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, pp. 215–232, London Mathematical Society Lecture Note Series 341. Cambridge: Cambridge University Press, 2007.

[Cremona 05] J. E. Cremona. `mwrank` (software). Available online (www.warwick.ac.uk/~masgaj/ftp/progs), 2005.

[Cremona 06] J. E. Cremona. Elliptic Curve Data. Available online (www.warwick.ac.uk/~masgaj/ftp/data/), 2006.

[Delaunay and Watkins 07] C. Delaunay and 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, pp. 189–193, London Mathematical Society Lecture Note Series 341. Cambridge: Cambridge University Press, 2007.

[Diamond 96] F. Diamond. "On Deformation Rings and Hecke Rings." *Ann. of Math. (2)* 144:1 (1996), 137–166.

[Dokchiter and Dokchitser 06] T. Dokchiter and V. Dokchitser, "Root Numbers of Elliptic Curves in Residue Characteristic 2." Preprint available online (arxiv.org/math/0612054), 2006.

[Duke 97] W. Duke. "Elliptic curves with No Exceptional Primes." *C. R. Acad. Sci. Paris Sér. I Math.* 325:8 (1997) 813–818.

[Duke and Kowalski 00] W. Duke and E. Kowalski. "A Problem of Linnik for Elliptic Curves and Mean-Value Estimates for Automorphic Representations." *Invent. Math.* 139:1 (2000) 1–39.

[Dummigan et al. 06] N. Dummigan, P. Martin, and M. Watkins. "Euler Factors and Local Root Numbers for Symmetric Powers of Elliptic Curves." Preprint, 2006.

[Fisher 07] T. A. Fisher. "Finding Rational Points on Elliptic Curves Using 6-Descent and 12-Descent." Preprint available online (www.dpmms.cam.ac.uk/~taf1000/papers/sixandtwelve.html), 2007.

[Fouvry et al. 92] É. Fouvry, M. Nair, and G. Tenenbaum. "L'ensemble exceptionnel dans la conjecture de Szpiro." *Bull. Soc. Math. France* 120:4 (1992), 485–506.

[Frey 94] G. Frey, editor. *On Artin's Conjecture for Odd 2-Dimensional Representations*, Lecture Notes in Mathematics, 1585. Berlin: Springer-Verlag, 1994.

[Goldfeld 79] D. Goldfeld. "Conjectures on Elliptic Curves over Quadratic Fields." In *Number Theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979)*, edited by M. B. Nathanson, pp. 108–118, Lect. Notes in Math. 751. Berlin: Springer-Verlag, 1979.

[Greenberg 83] R. Greenberg. "On the Birch and Swinnerton-Dyer Conjecture." *Invent. Math.* 72:2 (1983), 241–265.

[Helfgott 04] H. A. Helfgott. "On the Behaviour of Root Numbers in Families of Elliptic Curves." Preprint available online (arxiv.org/math/0408141), 2004.

[Hindry 05] M. Hindry. "Why Is It Difficult to Compute the Mordell–Weil Group?" Preprint available online (www.math.jussieu.fr/~hindry/MW-size.pdf), 2005.

[Katz 05] N. M. Katz. *Moments, Monodromy, and Perversity: A Diophantine Perspective*, Annals of Mathematics Studies, 159. Princeton: Princeton University Press, 2005.

[Katz and Sarnak 99] N. M. Katz and P. Sarnak. *Random Matrices, Frobenius Eigenvalues, and Monodromy*, American Mathematical Society Colloquium Publications, 45. Providence: American Mathematical Society, 1999.

[Keating and Snaith 00] J. P. Keating and 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.* 214:1 (2000), 57–89 and 91–110.

[Kobayashi 02] S. Kobayashi. "The Local Root Number of Elliptic Curves with Wild Ramification." *Math. Ann.* 323:3 (2002), 609–623.

[Lang 83] 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, pp. 155–171, Progress in Mathematics 35. Boston: Birkhäuser, 1983.

[Liu and Xu 04] C. Liu and L. Xu. "The Vanishing Order of Certain Hecke $L$-Functions of Imaginary Quadratic Fields." *J. Number Theory* 108:1 (2004), 76–89.

[Marčenko and Pastur 67] V. A. Marčenko and L. A. Pastur. "Distribution of Eigenvalues in Certain Sets of Random Matrices." *Math. USSR-Sb.* 1 (1967), 457–483. (Russian original in *Mat. Sb. (N.S.)* 72:114 (1967), 507–536.)

[Martin and Watkins 06] P. Martin and 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, pp. 377–392, Springer Lecture Notes in Computer Science 4076. Berlin: Springer, 2006.

[Mehta 04] M. L. Mehta. *Random Matrices*, third edition, Pure and Applied Mathematics (Amsterdam) 142. Amsterdam: Elsevier/Academic Press, 2004.

[Michel 95] P. Michel. "Rang moyen de familles de courbes elliptiques et lois de Sato–Tate." *Monatsh. Math.* 120:2 (1995), 127–136.

[Miller 04] S. J. Miller. "One- and Two-Level Densities for Rational Families of Elliptic Curves: Evidence for the Underlying Group Symmetries." *Compos. Math.* 140:4 (2004), 952–992.

[Pacetti and Tornaría 08] A. Pacetti and G. Tornaría. "Computing Central Calues of Twisted $L$-Series: The Case of Composite Levels." To appear in *Exp. Math.*, 2008.

[Ramakrishnan and Shahidi 07] D. Ramakrishnan and F. Shahidi. "Siegel Modular Forms of Genus 2 Attached to Elliptic Curves." *Math. Res. Lett.* 14:2 (2007), 315–332.

[Rapoport et al. 88] M. Rapoport, N. Schappacher, and P. Schneider, editors. *Beilinson's Conjectures on Special Values of L-Functions*, Perspectives in Mathematics, 4. Boston: Academic Press, 1988.

[Rodriguez Villegas and Zagier 91] F. Rodriguez Villegas and 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, pp. 81–99, Oxford Sci. Publ. New York: Oxford Univ. Press, 1993.

[Rohrlich 80] D. E. Rohrlich. "The Nonvanishing of Certain Hecke $L$-Functions at the Center of the Critical Strip." *Duke Math. J.* 47:1 (1980), 223–232.

[Rohrlich and Montgomery 80] D. E. Rohrlich and H. L. Montgomery. "On the $L$-Functions of Canonical Hecke Characters of Imaginary Quadratic Fields, I, II." *Duke Math. J.* 47:3 (1980), 547–557, and *Duke Math. J.* 49:4 (1982), 937–942.

[Rosson and Tornaría 07] H. Rosson and 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, pp. 315–321, London Mathematical Society Lecture Note Series 341. Cambridge: Cambridge University Press, 2007.

[Schwarz 65] W. Schwarz. "Einige Anwendungen Tauberscher Sätze in der Zahlentheorie." *J. Reine Angew. Math.* 219 (1965), 157–179.

[Siegel 45] C. L. Siegel. "A Mean Value Theorem in the Geometry of Numbers." *Ann. of Math. (2)* 46 (1945), 340–347.

[Silverman 92] J. H. Silverman. *The Arithmetic of Elliptic Curves*. New York: Springer-Verlag, 1992.

[Stamminger 05] S. K. M. Stamminger. "Explicit 8-Descent on Elliptic Curves." PhD thesis, International University Bremen.

[Stein and Watkins 02] W. A. Stein and 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, pp. 267–275, Springer Lecture Notes in Computer Science, 2369. New York: Springer, 2002.

[Tate 65] 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, pp. 93–110. New York: Harper & Row, 1965.

[Taylor and Wiles 95] R. Taylor and Wiles. "Ring-Theoretic Properties of Certain Hecke Algebras." *Ann. of Math. (2)* 141:3 (1995), 553–572.

[Tenenbaum 88] G. Tenenbaum. "La méthode du col en théorie analytique des nombres." In *Séminaire de Théorie des Nombres, Paris 1986–87*, edited by C. Goldstein, pp. 411–441, Progr. Math. 75. Boston: Birkhäuser Boston, 1988.

[Waldspurger 81] J.-L. Waldspurger. "Sur les coefficients de Fourier des formes modulaires de poids demi-entier." *J. Math. Pures Appl. (9)* 60:4 (1981), 375–484.

[Whitehouse 04] D. Whitehouse. "Root Numbers of Elliptic Curves over 2-adic Fields." Preprint, 2004.

[Wigner 55] E. Wigner. "Characteristic Vectors of Bordered Matrices with Infinite Dimensions." *Ann. of Math. (2)* 62 (1955), 546–564.

[Wiles 95] A. Wiles. "Modular Elliptic Curves and Fermat's Last Theorem." *Ann. of Math. (2)* 141:3 (1995), 443–551.

[Wishart 28] J. Wishart. "The Generalized Product Moment Distribution in Samples from a Normal Multivariate Population." *Biometrika* 20A (1928), 32–52.

[Young 06] M. P. Young. "Low-Lying Zeros of Families of Elliptic Curves." *J. Amer. Math. Soc.* 19:1 (2006), 205–250.

Mark Watkins, School of Mathematics and Statistics F07, University of Sydney, NSW 2006, Australia
(watkins@maths.usyd.edu.au)

Received February 20, 2007; accepted September 26, 2007.