Computing the modular degree of an elliptic curve

Mark Watkins, Department of Mathematics, McAllister Building, The Pennsylvania State University, University Park, PA, 16802 (watkins@math.psu.edu)

Abstract

We review previous methods of computing the modular degree of an elliptic curve, and present a new method (conditional in some cases), which is based upon the computation of a special value of the symmetric square L-function of the elliptic curve. Our method is sufficiently fast to allow large-scale experiments to be done. The data thus obtained on the arithmetic character of the modular degree show two interesting phenomena. First, in analogy with class number in the number field case, there seems to be a Cohen–Lenstra heuristic for the probability that an odd prime divides the modular degree. Secondly, the experiments indicate that 2^r should always divide the modular degree, where r is the Mordell–Weil rank of the elliptic curve. We also discuss the size distribution of the modular degree, or more exactly of the special L-value which we compute, again relating it to the number field case.

1. Introduction

Let E be an elliptic curve over the rationals. We can assume that E is in the form $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ and that this is a minimal Weierstrass equation for E; we will refer to such a curve as $[a_1, a_2, a_3, a_4, a_6]$. By the work of Wiles and others ([Wiles 1995, Breuil et al. 2001]), it is known that there is a surjective morphism (called a modular parametrisation) $\phi : X_0(N) \to E$, where $X_0(N)$ is the (compactification of the) standard curve classifying cyclic N-isogenies and N is the conductor of E. The curve $X_0(N)$ can also be viewed as the upper half-plane modulo the action of the group $\Gamma_0(N) = \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} : N|c,\ ad - bc = 1\right\}$, with appropriate cusps added. Since both E and $X_0(N)$ can be realised as Riemann surfaces, this modular parametrisation has a topological degree. We call this the **modular degree** of E, and denote it by $\deg\phi$. Equivalently, this degree is also the usual notion of degree from algebraic geometry, namely the index of the pullback of the function field of E in the function field of $X_0(N)$. It shall be our goal to compute the modular degree for a large set of elliptic curves, and study its size and arithmetic properties.

There are a few problems when talking about the running time of an algorithm to compute the modular degree. The first is that good upper bounds on the modular degree are only known under the assumption of the ABC-conjecture. To avoid problems with the time needed simply to output the result and questions of precision needed in our calculations, we tacitly assume that there is a polynomial bound (in N) on the modular degree, as would follow from the ABC-conjecture. As such, our time estimates are heuristic. Secondly, some of the algorithms we present below require the computation of the pth trace of Frobenius of E for various primes p. There is an algorithm in [Schoof 1985] which does this in time $O\big((\log p)^8\big)$. However, for our range of interest, the asymptotically inferior baby steps/giant steps method of Mestre (see [Cohen 1993]), which takes $O(p^{1/4})$ time, is faster. This is the more practical running time estimate, and the one which we report below; we also leave out powers of $\log N$ in our time bounds.

There are a number of algorithms known for computing the modular degree. The first to appear in the literature seems to be [Zagier 1985], whose method (explicit only for prime N) involves triangulating the fundamental domain for $\Gamma_0(N)$, and then traversing this, noting how different edges glue together. The proper choice of fundamental domain is very effective in the case of prime N, giving an algorithm which, using the fast Fourier transform, runs in $O(N^{5/4})$ time and $O(N)$ space. However, in the general case when N is composite, the algorithm becomes markedly more complex and the running time appears then to be no better than N^2. For comparison with the other algorithms, we note that Zagier's method computes neither the $X_0(N)$-optimal curve nor the Manin constant, but given one of the two the other can be computed (see below for the definitions of these). Another method is given in [Mestre 1986], involving the "method of graphs" which utilises supersingular j-invariants. Again this is described explicitly only for prime N, but here the relative gains from a generalisation to composite N are sufficient to make such possibly worthwhile. This algorithm takes about N^2 time, and computes both the Manin constant and the $X_0(N)$-optimal curve. Via the use of sparse matrix techniques, the memory requirements can be kept to size about N. In the early 1990s, Elkies (personal communication) used this method to compute the modular degree of the rank 4 curve $[0, 1, 1, -72, 210]$. Related to this is the method of [Birch 1991], which uses ternary quadratic forms.

However, it only works for the -1 quotient, i.e. those curves whose L-function satisfies an even functional equation. This is a special case of the method of Brandt matrices developed in [Eichler 1973], and generalised in [Pizer 1976]. Finally there are methods using modular symbols, one of which is explained in [Cremona 1995], it being described as a variant of Zagier's method. But one can alternatively give such methods a more algebraic formulation; for instance, [Frey and Müller 1999] expresses the modular degree in terms of an intersection pairing, which can then be computed using the techniques of [Merel 1993]. A similar approach appears in [Merel 1995]. And [Kohel and Stein 2000] expresses the modular degree as the square of the order of a cokernel of a natural restriction map involving modular symbols (as such, it is computable given the modular symbols, and works for all quotients of the Jacobian, not just the elliptic ones). The computation of modular symbols naïvely takes around N^3 time, due to matrix operations on matrices of size N by N, but sparse matrix techniques might reduce this (and the memory requirements). Admittedly, these methods using modular symbols do much more than just compute the modular degree (and the Manin constant and $X_0(N)$-optimal curve) — for instance, they enumerate all the elliptic curves of a given conductor.

Our method is to compute a special value of a certain L-function, which is related to the modular degree via a formula that comes from a Rankin-Selberg convolution. Indeed, as in [Flach 1993] (reformulating a result of [Shimura 1976]), we have that

$$\frac{L(\mathrm{Sym}^2 E, 2)}{\pi i \Omega} = \frac{\deg \phi}{N c^2} \prod_{p^2 | N} U_p(2), \qquad (\star)$$

where $L(\mathrm{Sym}^2 E, s)$ is the (motivic) symmetric-square L-function, $\Omega = \int_{E(\mathbf{C})} \omega \wedge \bar{\omega}$ is the complex volume (which is $2/i$ times the volume of the fundamental parallelogram; see below for the definition of the Néron differential ω), c is the Manin constant, and the product over bad primes can be described explicitly (see section 2). The L-value here is at the edge of the critical strip, and there is a strong link with Dirichlet's class number formula. The quantity Ω plays the role of the regulator — one major difference is that Ω can be computed extremely rapidly to high precision, via the arithmetic-geometric mean of Gauss. If the analogy to the class number formula holds, then $\deg \phi$ corresponds to the class number, and in section 4 we shall comment on the group that is associated to the modular parametrisation. The product over bad primes comes from two sources. The first is the possibility of our elliptic curve not being minimal in its family of quadratic twists — this corresponds to a non-fundamental discriminant in Dirichlet's case. The other effect of bad primes is more subtle. In section 2, we define the symmetric-square L-function in full generality; when N is squarefree, it is quite straightforward, but square divisors of N cause enough problems for there to be two notions of the symmetric-square (analytic and motivic) and in this case, $U_p(2)$ measures the difference between the two. Finally, to expound further on the links to algebraic number theory, we mention that similar to the theory of genera for number fields (which involves the 2-divisibility of the class number), here we have a theory of Atkin-Lehner involutions through which the modular parametrisation map often factors, correspondingly affecting the 2-divisibility of the modular degree. As we shall see in section 4, there appears to be another influence on the 2-divisibility of the modular degree, namely the rank of the elliptic curve. Also in section 4, we shall give some experimental evidence that a Cohen–Lenstra heuristic (see [Cohen and Lenstra 1984]) holds for the divisibility of the modular degree by odd primes — and also some evidence that such a heuristic does not hold.

But how does the above formula ($\star$) help us compute $\deg \phi$? Using the work of [Shimura 1975] and [Gelbart and Jacquet 1978], we know that $L(\mathrm{Sym}^2 E, s)$ (the motivic version if N is not squarefree) has an analytic continuation to an entire function, and $\Lambda(\mathrm{Sym}^2 E, s) = (\tilde{N}^2/4\pi^3)^{s/2}\Gamma(s)\Gamma(s/2)L(\mathrm{Sym}^2 E, s)$ satisfies the functional equation $\Lambda(\mathrm{Sym}^2 E, s) = \Lambda(\mathrm{Sym}^2 E, 3 - s)$. Here $\tilde{N}$ is the symmetric-square conductor (fully defined in section 2), which always divides N and is equal to it if the conductor is squarefree. This functional equation is almost all that is needed to compute $L(\mathrm{Sym}^2 E, 2)$ fast. The appendix of [Cohen 2000] gives a method (whose roots date back to Hecke, but are generalised in a form suitable for us by [Lavrik 1967]) for computing any (reasonable) special L-value to a precision of D bits using only knowledge of the functional equation and the first $O(D^g\sqrt{C})$ terms of the Dirichlet series, where C is the conductor of the functional equation, and g is the number of Γ-factors in the functional equation. How much precision do we need for $L(\mathrm{Sym}^2 E, 2)$? Assuming the ABC-conjecture, we need only compute a constant times the number of digits of N, so that the D^g term is a power of $\log N$. The conductor here is $\tilde{N}^2$, so the method requires computation of about $\tilde{N}$ series coefficients. The series coefficients follow immediately upon calculation of the traces of Frobenius, and thus, using the baby steps/giant steps algorithm, our time estimate is $\tilde{N}^{5/4}$. This is a smaller exponent than any of the methods mentioned above (save Zagier's for prime N), and it works for any elliptic curve. The main downside of our algorithm is that we need to know the Manin constant.

In order to obtain $\deg \phi$ from the above equation ($\star$), we must also have good algorithms for computing the objects other than $L(\mathrm{Sym}^2 E, 2)$. The conductor N can be obtained about as fast as the discriminant can be factored using the algorithm in [Tate 1975]. In section 2 we describe the bad Euler factors $U_p(s)$, and these follow immediately (from divisibility and congruence conditions) once the conductor is known. The complex volume can be computed to high precision extremely fast (quadratic convergence) using the arithmetic-geometric mean, a process essentially known to Gauss (see [Cohen 1993]). Hence the above method computes $\frac{1}{c^2}\deg \phi$ in time no worse than $N^{5/4}$ (times some power of $\log N$), with the dominant amount of time being the computation of the coefficients of the L-series from the traces of Frobenius. This is fast enough to be used in some large-scale experiments. The differential ω in the definition of the complex volume is the heart of the problem with the Manin constant. The canonical Néron differential ω on $E = [a_1, a_2, a_3, a_4, a_6]$ is defined to be $dx/(2y + a_1 x + a_3)$. Under the modular parametrisation map ϕ, this pulls back to a differential on $X_0(N)$. Letting $f(z) = \sum_n l_n e^{2\pi i n z}$ be the weight 2 level N newform associated to E (so that l_p is the pth trace of Frobenius of E), we know that $f(z)\,dz$ is also a differential on $X_0(N)$, which by the multiplicity-one theorems of [Atkin and Lehner 1970] differs from ω by a constant. The Manin constant c is defined (up to sign, taken positive) by $\phi^\star(\omega) = 2\pi i c f(z)\,dz$. It is conjectured in [Manin 1972] that $c = 1$ for the so-called optimal (or strong) curve in an isogeny class. The work of [Katz and Mazur 1985] implies that c is an integer; this is treated (without reference) as a well-known fact on pg. 310 of [Gross and Zagier 1986], and is mentioned in [Frey 1987] as being an observation of Oesterlé. The most general upper bounds for c are due to [Edixhoven 1991] — he has indicated that he has sharper results in unpublished work. His paper appears to be the first to write down the 1-paragraph derivation (after Katz–Mazur) of the integrality of c, and in his thesis, Edixhoven indicates that the correct attribution for this might belong to Gabber (unpublished). Most relevant to our experiments is [Abbes and Ullmo 1996], which shows that (in particular) when N is odd and squarefree, we have $c = 1$ for the optimal curve. If another conjecture (of Stevens, regarding which curve is optimal for parametrisations from $X_1(N)$) is assumed, we can quickly determine which isogenous curve is $X_0(N)$-optimal (see section 3). If the curve we are given is not optimal, it is easy to determine the relative factor between its modular degree and that of the optimal curve (this applies to all the algorithms). Thus assuming both the Manin and Stevens conjectures, we are able to compute the modular degree of any elliptic curve using our method (the assumption of the ABC-conjecture is only needed for estimates on the running time).

While herein we consider the value of $L(\mathrm{Sym}^2 E, s)$ at the edge of the critical strip (which is the point $s = 2$), some work has been done for $s = 3$, particularly with respect to values of elliptic trilogarithms and their relation to conjectures of Beilinson and Bloch–Kato. Notable is section 10 of the recent [Zagier and Gangl 2000], while [Mestre and Schappacher 1991] has many computations, and indicates that Henniart has probably anticipated much of the calculations in our section 2; however, the "table numérique" (section 3.3) of this latter work unfortunately seems replete with errors — for instance the curve $[0, 0, 0, -15, -50]$ is asserted to have conductor 900, while its conductor is actually 3600. Furthermore, the Euler factor at 2 is often incorrect, possibly due to the incompleteness (see below) of the classification of [Coates and Schmidt 1987]. It should also be noted that similar work to ours has already been done for the symmetric cube L-function. Buhler, Schoen, and Top [Buhler et al. 1997] investigate the experimental validity of a Birch–Swinnerton-Dyer type formula which relates the central value $L(\mathrm{Sym}^3 E, 2)$ to the Griffiths group. As the critical value is shifted to the center, the behaviour is very much different, and hence the results are not all that comparable. We also make a practical note on the implementation of the computation of the special L-values. We need to compute what might be called "incomplete K-Bessel functions," in analogy with the incomplete Γ-functions which come up when (say) computing the analytic rank of an elliptic curve. There are some sophisticated ways of doing this, but we found that the fastest was simply to compute the relevant functions (and sufficiently many derivatives) once-and-for-all on a mesh of values, and then use local power series to interpolate. In fact, the derivatives of the functions in question satisfy recurrence relations, making the task even simpler. We also used the memory-efficient algorithm of [Buhler and Gross 1985] for computing multiplicative sums, but with the memory sizes of today's computers this might be unneeded.

2. Symmetric-Square L-functions and minimal twists

Let $L(E,s)=\prod_p(1-\alpha_p/p^s)^{-1}(1-\beta_p/p^s)^{-1}$ be the standard L-function for E. Here, for p not dividing N, we have $\beta_p=\bar{\alpha}_p$ and $\alpha_p+\beta_p=l_p$, where l_p is the pth trace of Frobenius of E. For $p\|N$, we have $\beta_p=0$ and $\alpha_p^2=1$, while $\beta_p=\alpha_p=0$ when $p^2|N$. The analytic symmetric-square L-function is now defined as

$$L^A(\mathrm{Sym}^2 E,s)=\prod_p L_p^A(\mathrm{Sym}^2 E,s)=\prod_p(1-\alpha_p^2/p^s)^{-1}(1-\alpha_p\beta_p/p^s)^{-1}(1-\beta_p^2/p^s)^{-1}.$$

This is the "imprimitive" $D(E,s)$ in equation (1.11) of [Coates and Schmidt 1987]; it is not stable under quadratic twists, though twisting by a fundamental discriminant D does not affect the Euler factors of primes not dividing D. In the derivation of the formula ($\star$), this is the more useful symmetric-square L-function due to the fact that it is a convolution of $L(E,s)$ with itself, and hence can be analysed via the Rankin-Selberg method of unfolding as in [Shimura 1976], from which we get the formula

$$\frac{L^A(\mathrm{Sym}^2 E,2)}{\pi i\Omega}=\frac{\deg\phi}{Nc^2}.$$

However, for the functional equation to hold, we must adjust $L^A(\mathrm{Sym}^2 E,s)$ by appropriate Euler factors when $p^2|N$. This is described automorphically in [Gelbart and Jacquet 1978] and via techniques of Iwasawa theory in Coates–Schmidt. We give an explicit formulation involving nothing more than divisibility and congruence conditions, largely following the exposition of Coates–Schmidt, and correcting a couple of errors therein. We define the Euler product $U(s)$ via $L^M(\mathrm{Sym}^2 E,s)=L^A(\mathrm{Sym}^2 E,s)\cdot U(s)$ where $\Lambda^M(\mathrm{Sym}^2 E,s)=(\tilde{N}^2/4\pi^3)^{s/2}\Gamma(s)\Gamma(s/2)L^M(\mathrm{Sym}^2 E,s)$ satisfies the functional equation given by $\Lambda^M(\mathrm{Sym}^2 E,s)=\Lambda^M(\mathrm{Sym}^2 E,3-s)$. This motivic L-function is stable under quadratic twists; Theorem 2.4 of Coates–Schmidt makes explicit that it satisfies the functional equation (they denote it by script-D). Denote by $U_p(s)$ the local factor of $U(s)$ at a prime p. Below we shall that this is identically 1 unless $p^2|N$, in which case its description is more complicated. So if N is squarefree, that is, E is semistable, $U(s)$ itself is identically 1. There is also the aspect of the symmetric-square conductor $\tilde{N}$ in the functional equation, which is also easy in the semistable case, where $\tilde{N}=N$.

2.1. Quadratic twists and minimality

We define the notion of a p-minimal quadratic twist E_p of an elliptic curve E (in minimal Weierstrass form). We let E^n be the nth quadratic twist of E, and for odd primes p write $\tilde{p}=\left(\frac{-1}{p}\right)$. For each odd prime p, we let E_p be either E or $E^{\tilde{p}}$, choosing the one which has smaller local conductor, with ties being broken by smaller local discriminant. For $p=2$, we let E_2 be one of E, E^{-4}, and $E^{\pm 8}$, again choosing the one with smallest local conductor then discriminant, and (arbitrarily) taking the curve with $c_6\geq 0$ if twisting by -1 results in curves with the same local conductor and discriminant. For $p\geq 5$, we have that $E_p=E^{\tilde{p}}$ if and only if $p^2|c_4$ and $p^3|c_6$. A more complicated criterion can be written down for $p=2$ and $p=3$ (see [Stein and Watkins 2002]). In particular, if p^2 does not divide N_E, then E is p-twist-minimal, and twisting by fundamental discriminants ensures that we do not affect minimality at other primes. By iteratively minimising a curve locally prime-by-prime, we end up with a global minimal twist. Since the symmetric-square L-function of an elliptic curve is isogeny-invariant, the form of the functional equation must end up the same no matter which isogenous curve we use. As such, the fact that p-twist-minimality is not necessarily isogeny-invariant for $p=2$ or $p=3$ is not overly important. There are reasons to make the primary sorting by discriminant instead of conductor (this possibly affects only $p=2$), but here we regard conductor as more important.

Let F be the global minimal twist of E, letting N_F and N_E be their respective conductors. We compare the modular degrees of E and F, using the above formula, proceeding prime-by-prime. We have $L_p^A(\mathrm{Sym}^2 F,s)=L_p^A(\mathrm{Sym}^2 E_p,s)$ since the Euler factor is stable under twists by fundamental discriminants coprime to p. So if $E_p=E$ then $L_p^A(\mathrm{Sym}^2 E,s)=L_p^A(\mathrm{Sym}^2 F,s)$. For primes with $E_p\neq E$ we have that $p^2|N_E$, and thus $L_p^A(\mathrm{Sym}^2 E,s)\equiv 1$. We write

$$\deg\phi_E=\deg\phi_F\cdot\frac{c_E^2}{c_F^2}\cdot\prod_p V_p,$$

so that $V_p = 1$ when $E_p = E$ and $V_p = \frac{\Omega_{E_p}}{\Omega_E} \cdot \frac{N_E}{N_{E_p}} \cdot L_p^A(\mathrm{Sym}^2 E_p, 2)^{-1}$ when $E_p \neq E$. Every term in V_p is easily computed, and thus it is quite straightforward to determine the modular degree of a curve once that of its minimal twist has been found (if we assume the Manin constants are the same). When $p \neq 2$ and $E_p \neq E$, we can describe V_p more directly. Firstly, if E_p has good reduction at p, then we compute that $L_p^A(\mathrm{Sym}^2 E_p, s)^{-1} = \bigl(1 - b_p/p^s + pb_p/p^{2s} - p^3/p^{3s}\bigr)$ where $b_p = l_p^2 - p$ and l_p is the pth trace of Frobenius of E_p. Evaluating this at $s = 2$, we get $\frac{1}{p^3}(p-1)(p+1-l_p)(p+1+l_p)$. We have that $N_E/N_{E_p} = p^2$ and $\Omega_E/\Omega_{E_p} = 1/p$. Thus $V_p = (p-1)(p+1-l_p)(p+1+l_p)$ (this appears already in [Zagier 1985]). Secondly, if E_p has multiplicative reduction at p, we have $L_p^A(\mathrm{Sym}^2 E_p, s)^{-1} = (1 - 1/p^s)$. Again $\Omega_E/\Omega_{E_p} = 1/p$, but here $N_E/N_{E_p} = p$. So $V_p = (p-1)(p+1)$ in this case. Finally, if E_p has additive reduction at p, then the twisting does not change the L-function or the conductor, but does increase the volume by a factor of p, thus decreasing the modular degree by $V_p = p$.

2.2. Calculating $U(s)$ for a global minimal twist

So we have reduced the problem to computing the modular degree of a global minimal twist, which we continue to call F. We define local conductors δ_p by $N_F = \prod_p p^{\delta_p}$, and write the symmetric-square conductor as a product of local conductors as $\tilde{N} = \prod_p p^{\tilde{\delta}_p}$. If F has good reduction at p, then Case 1 of Coates–Schmidt on pg. 107 implies that $L_p^M(\mathrm{Sym}^2 F, s) = L_p^A(\mathrm{Sym}^2 F, s)$ (and so $U_p(s) \equiv 1$) while $\tilde{\delta}_p = 0$. If F has multiplicative reduction at p, then Lemma 1.2 of Coates–Schmidt implies that $L_p^M(\mathrm{Sym}^2 F, s) = L_p^A(\mathrm{Sym}^2 F, s)$ again, and their comments below Lemma 2.12 on page 119 show that $\tilde{\delta}_p = 1$. This leaves the most difficult case where F has additive reduction at p. Note that $L_p^A(\mathrm{Sym}^2 F, s) \equiv 1$ in this case, so that $L_p^M(\mathrm{Sym}^2 F, s) = U_p(s)$. We write F as $y^2 = x^3 - 27c_4 x - 54c_6$; the fact that this model is not minimal at 2 and 3 will not matter. Because F has additive reduction at p, we have $p|c_4$ and $p|c_6$. From Lemma 1.4 of Coates–Schmidt, there are three possibilities for $U_p(s)$, it being $(1 \pm p/p^s)^{-1}$ or identically 1.

We first consider $p \geq 5$, where the argument following Lemma 2.12 of Coates–Schmidt tells us that $\tilde{\delta}_p = 1$. Letting F_3 be the set of coordinates of the 3-torsion points of F, Lemma 1.4 of Coates–Schmidt tells us that $U_p(s) = (1 - p/p^s)^{-1}$ if $\mathbf{Q}_p(F_3)/\mathbf{Q}_p$ is an abelian extension, and $U_p(s) = (1 + p/p^s)^{-1}$ if it is not. Let $G = \mathrm{Gal}\bigl(\mathbf{Q}_p(F_3)/\mathbf{Q}_p\bigr)$, and Φ_p be the inertia group of this extension, recalling that G/Φ_p is cyclic. There are three possibilities for Φ_p; it is cyclic of order 3, 4, or 6 (see page 108 of Coates–Schmidt). We also have $G \subseteq GL_2(\mathbf{F}_3)$, due to the fact that the 3-torsion is isomorphic to $\mathbf{Z}/3\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$. We let $\mathbf{Q}_p(F_3^x)$ be the extension of $\mathbf{Q}_p$ by just the x-coordinates of the 3-torsion. Modding G out by conjugation (if any), we get the Galois group of this extension, which we denote by H, so that $H \subseteq PGL_2(\mathbf{F}_3)$. We let C_n be the cyclic group of order n and D_{2n} is the dihedral group of order $2n$. If $\Phi_p \cong C_3$, the requirement that G/Φ_p be cyclic implies that G is C_3, C_6, or D_6. By the conjugation action the first two lead to $H \cong C_3$, and the third to $H \cong D_6$. If $\Phi_p \cong C_4$, then G is one of C_4, C_8, D_8, or the quarternion group of order 8, denoted Q_8. The last two imply that $H \cong C_2 \times C_2$, while if G is C_8 then $H \cong C_4$, and if G is C_4 then $H \cong C_2$. When $\Phi_p \cong C_6$, we get that either G is C_6 and $H \cong C_3$ as before, or $G \cong D_{12}$ and $H \cong D_6$. So the question of the abelian nature of G can be answered by determining H — we see that G is abelian iff H is cyclic. This turns out only to depend on the congruence class of p mod 12 and various p-divisibilities of c_4 and c_6. Let $p^\alpha \| c_4$ and $p^\beta \| c_6$. Because $p \geq 5$ and F is twist-minimal, we cannot have both $\alpha \geq 2$ and $\beta \geq 3$, and so it follows that either $\alpha \geq \beta \geq 1$ or $\alpha = 1$ and $\beta \geq 2$. We have the following theorem and corollary.

Theorem. *Assume that $p \geq 5$ is prime and $p^\alpha \| c_4$ and $p^\beta \| c_6$. If $\alpha \geq \beta \geq 1$, then G is abelian iff $p \equiv 1$ (3). If $\alpha = 1$ and $\beta \geq 2$, then G is abelian iff $p \equiv 1$ (4).*

Corollary. *Assume F is twist-minimal with additive reduction at a prime $p \geq 5$. The minus sign always occurs in $U_p(s)$ when $p \equiv 1$ (12), and the plus sign always occurs when $p \equiv 11$ (12). When $p \equiv 5$ (12) the minus sign occurs iff $p^2 | c_6$ and $p \| c_4$, and when $p \equiv 7$ (12) the plus sign occurs exactly when $p^2 | c_6$ and $p \| c_4$.*

The Corollary follows immediately from the Theorem and Lemma 1.4 of Coates–Schmidt. For a curve of the form $y^2 = x^3 + ax + b$, the x-coordinates of the 3-torsion points are the roots of the polynomial $3x^4 + 6ax^2 + 12bx - a^2$. We divide out by powers of 3 to get that the field $\mathbf{Q}_p(F_3^x)$ is defined by the roots of the polynomial $f(x) = x^4 - 6c_4 x^2 - 8c_6 x - 3c_4^2$. We now compute $H = \mathrm{Gal}\bigl(\mathbf{Q}_p(F_3^x)/\mathbf{Q}_p\bigr)$ in the various cases. We write $c_4 = p^\alpha u_4$ and $c_6 = p^\beta u_6$. First suppose that $\alpha \geq \beta \geq 1$. Here $f(x)$ factors as $(x - \xi)g(x) =$

$(x-\xi)(x^3+\xi x^2+Ax+B)$ with $\xi = -\frac{3u_4^2}{8u_6}p^{2\alpha-\beta} + O(p^{2\alpha-\beta+1})$, $A = \xi^2 - 6u_4p^\alpha = -6u_4p^\alpha + O(p^{\alpha+1})$, and $B = 3u_4^2p^{2\alpha}/\xi = -8u_6p^\beta + O(p^{\beta+1})$. We have $\mathrm{disc}(g) = \xi^2A^2 - 4\xi^3B - 4A^3 + 18\xi AB - 27B^2 = -27B^2 + O(p^{3\alpha})$, which is a square in $\mathbf{Q}_p$ iff -3 is a square, that is, iff $p \equiv 1$ (3). When $\mathrm{disc}(g)$ is a square, we have $H \cong C_3$, while $H \cong D_6$ if not. Using the H-G-correspondence then gives us the first statement of the theorem.

Next suppose that $\alpha = 1$ and $\beta \geq 2$. Here f has no roots modulo p^2, and thus none in $\mathbf{Q}_p$. We try to factor $f(x)$ as $(x^2+Ax+B)(x^2-Ax+C)$, getting the 3 equations $B+C-A^2 = -6u_4p$, $A(B-C) = 8u_6p^\beta$, and $BC = -3u_4^2p^2$. We write $\tilde{B} = B/pu_4$ and $\tilde{C} = C/pu_4$, so that we have the mod-p-congruences $\tilde{B} + \tilde{C} \equiv -6$ and $\tilde{B}\tilde{C} \equiv -3$. These imply that $\sqrt{3} \in \mathbf{Q}_p$, so that there is no solution (and hence $f(x)$ is irreducible) when $p \equiv \pm 5$ (12) — we return to this possibility below. When $p \equiv \pm 1$ (12), we substitute the first equation into the square of the second to eliminate A, and then eliminate C by using the third. This gives us that $\tilde{B}$ is a root of the sextic polynomial $(y^2+6y-3)(y^2+3)^2 - \frac{64u_6^2p^{2\beta}}{p^3u_4^3}y^3$. Since $\beta \geq 2$, the last term is 0 mod p. We note that the polynomial y^2+6y-3 has distinct roots mod p, so by Hensel's Lemma there is some $\mathbf{Q}_p$-root of this sextic, and from it we get a factorization $B = pu_4\omega_+ + O(p^2)$, $C = pu_4\omega_- + O(p^2)$, and $A = \frac{2u_6}{\sqrt{3}u_4}p^{\beta-1} + O(p^\beta)$, where $\omega_\pm = -3 \pm 2\sqrt{3}$. Now we have that $\mathbf{Q}_p(F_3^x) = \mathbf{Q}_p\bigl(\sqrt{A^2-4B}, \sqrt{A^2-4C}\bigr)$, and compute that $\frac{A^2-4B}{A^2-4C} = \frac{\omega_+}{\omega_-} + O(p) = \bigl(-7+4\sqrt{3}\bigr) + O(p)$, and $-7+4\sqrt{3}$ is a square exactly when $p \equiv 1$ (12). Thus $H \cong C_2$ when p is 1 mod 12, and $H \cong C_2 \times C_2$ when p is 11 mod 12, so by using the H-G-correspondence, we get half of the second statement of the theorem. We now analyse the cases $p \equiv \pm 5$ (12) for which $f(x)$ is irreducible in $\mathbf{Q}_p$. The discriminant Δ of f is $-2^{12}3^3(u_4^3p^3 - u_6^2p^{2\beta})^2$, and since none of the above possibilities for H contains A_4, the resolvent cubic must factor. When $p \equiv 7$ (12), the discriminant is a $\mathbf{Q}_p$-square, so that $H \cong C_2 \times C_2$ and G is nonabelian. When $p \equiv 5$ (12), the discriminant is not a square. However, the above factorization of $f(x)$ into quadratics works in the discriminant field $\mathbf{Q}_p(\sqrt{\Delta}) = \mathbf{Q}_p(\sqrt{-3}) = \mathbf{Q}_p(\sqrt{3})$. Thus $H \cong C_4$, and $G \cong C_8$ is abelian. This shows the Theorem.

We next discuss $p = 2$. Here the minimal twist will have neither 16 nor 64 exactly dividing its conductor (this follows from the table on pg. 121 of Coates–Schmidt, or more simply from an analysis of Tate's algorithm), so that δ_2 is neither 4 nor 6. If δ_2 is odd, the Coates–Schmidt table tells us that $U_2(s) \equiv 1$ and $\tilde{\delta}_2 = (1+\delta_2)/2$. If $\delta_2 = 2$, again there is not much problem; the table says that $U_2(s) = (1+2/2^s)^{-1}$ and $\tilde{\delta}_2 = 1$. The case of $\delta_2 = 8$ is the most difficult. The appendix of Coates–Schmidt makes two errors, leading to the classification being incomplete. The first error they make is on the fifth line of the $r = 2$ case on pg. 151: the quoted work of Atkin and Li requires the underlying form to have 16 dividing the level, and if the level of the absolute minimal quadratic twist of the form f associated to F (no longer necessarily rational, i.e., the twisted form can have a nontrivial Nebentypus character) is exactly divisible by 2^3, this does not apply. An explicit example is the curve 768H (given by $[0,1,0,1,-3]$) for which the absolute minimal twist is of level 24. Another error is on pg. 153 in the analysis of the case where the inertia subgroup is Q_8, where they state that "G is obviously a 2-Sylow group of $GL_2(\mathbf{F}_3)$, hence dihedral of order 16." This should be semidihedral of order 16. This causes them to miss the possibility that the absolute minimal twist can have 2^7 exactly dividing its level. An example is 256B (given by $[0,0,0,-2,0]$) where the absolute minimal twist is of level 128. So with $\delta_2 = 8$ this gives three different types of behaviour for the absolute minimal twist: it can have 2^3, 2^6, or 2^7 exactly divide its level. We can write $U_2(s) = (1+w/2^s)^{-1}$. The first case corresponds to $w = -2$, the second case to $w = 2$, and the third case to $w = 0$. The local symmetric-square conductor $\tilde{\delta}_2$ is respectively 3, 3, or 4. Both of these statements follow from a corrected Coates–Schmidt table. Finally, we re-interpret this in terms of congruences for c_4 and c_6.

Theorem. *If $2^8 | N_F$ and F is twist-minimal, then $2^5 \| c_4$ and $2^8 | c_6$. If $2^9 | c_6$ then $U_2(s) \equiv 1$ so that $\tilde{\delta}_2 = 4$. If $2^8 \| c_6$, then $\tilde{\delta}_2 = 3$, and if $c_4 \equiv 32$ (128) we have that $U_2(s) = (1+2/2^s)^{-1}$, while if $c_4 \equiv 96$ (128) we have that $U_2(s) = (1-2/2^s)^{-1}$.*

The first statement follows from an exercise using Tate's algorithm. By Lemma 1.4 of Coates–Schmidt, to find $U_2(s)$ it suffices to determine whether the inertia group Φ_2 of the extension $\mathbf{Q}_2(F_3)/\mathbf{Q}_2$ is cyclic and whether the Galois group G of this extension is abelian (the statements concerning $\tilde{\delta}_2$ follow as above, using the corrected Coates–Schmidt table). The corrected table tells us that when $2^8 | N$, we have that Φ_2 is either C_4 or Q_8. We first show that $\Phi_2 \cong Q_8$ iff $H \cong D_8$. At the top of pg. 153, Coates–Schmidt (corrected) shows that if $\Phi_2 \cong Q_8$, then $G \cong SD_{16}$, the semidihedral group of order 16, and consideration of the conjugation action then implies that $H \cong D_8$. Conversely, conjugation tells us that if $H \cong D_8$, then

$G \cong SD_{16}$. Since G/Φ_2 is cyclic but SD_{16}/C_4 is not, we must have $\Phi_2 = Q_8$ here. As before, everything follows upon determination of H.

We write $c_6 = 2^8 u_6$ and $c_4 = 2^5 u_4$, so that u_4 is odd, but u_6 need not be. We remove some powers of 2 from the 3-torsion polynomial, transforming it to $x^4 - 2^2 3 u_4 x^2 - 2^5 u_6 x - 2^2 3 u_4^2$. This has no $\mathbf{Q}_2$-roots, and we try to factor it as $(x^2 - Ax + B)(x^2 + Ax + C)$. Writing $\tilde{B} = B/2u_4$ and $\tilde{C} = C/2u_4$, as in the $p \geq 5$ case we find that $\tilde{B}$ satisfies a sextic relation, which we write here as $(y^2 + 6y - 3)(y^2 + 3)^2 = 2^7 \frac{u_6^2}{u_4^3} y^3$. For 2^7 to divide the left side we must have y be 3 mod 4, and then the left side is 384 mod 512. So if u_6 is even or u_4 is 3 mod 4, there are no $\mathbf{Q}_2$ solutions to this sextic, implying that $f(x)$ is irreducible over $\mathbf{Q}_2$ — we return to these cases below. When u_6 is odd and u_4 is 1 mod 4, we substitute $y = 3 + 4z$ into the above sextic, getting a new sextic relation $g(z) = (2z^2 + 6z + 3)(4z^2 + 6z + 3)^2 - \frac{u_6^2}{u_4^3}(4z + 3)^3 = 0$. We compute that $2 \| g'(\alpha)$ for all $\alpha \in \mathbf{Z}_2$, and note that u_6^2/u_4^3 is congruent to u_4 modulo 8. By taking $z = 0$ if u_4 is 1 mod 8 and $z = 1$ if u_4 is 5 mod 8 we get a mod 8 root of g. By Hensel's Lemma, this then lifts to a $\mathbf{Q}_2$ root of g, and thus a $\mathbf{Q}_2$ solution to the y-sextic. This gives us a factorization of $f(x)$ into quadratics. Since $\tilde{B}\tilde{C} = -3$, we find that $\tilde{B}$ and $\tilde{C}$ are congruent modulo 4, but not modulo 8. Thus $2^2 \| A$, and we get that $A^2 - 4B$ and $A^2 - 4C$ are also congruent mod 4 but not mod 8. So $\mathbf{Q}_2(F_3^x) = \mathbf{Q}_2\big(\sqrt{A^2 - 4B}, \sqrt{A^2 - 4C}\big)$ has Galois group $C_2 \times C_2$, implying that Φ_2 is cyclic, G is nonabelian, and $U_2(s) = (1 + 2/2^s)^{-1}$.

We now return to cases where $f(x)$ is irreducible over $\mathbf{Q}_2$. As with $p \geq 5$, the resolvent cubic must have a root in $\mathbf{Q}_2$, while the discriminant Δ is $-2^{18} 3^3 (u_4^3 - 2u_6^2)^2$, so that the discriminant field $\mathbf{Q}_2(\sqrt{\Delta})$ is $\mathbf{Q}_2(\omega)$ where $\omega^2 + \omega + 1 = 0$. We first consider the case where u_6 is odd and u_4 is 3 mod 4, and look at the $g(z)$-sextic. Similar to before, we have that $2 \| g'(\alpha)$ for all $\alpha \in \mathbf{Z}_2[\omega]$. When u_4 is 7 mod 8 we find that $3 + 3\omega$ is a mod 8 root of g, while if u_4 is 3 mod 8 we get that $1 + 7\omega$ is one. This root of g then lifts to $\mathbf{Q}_2(\omega)$, which gives us a factorization of $f(x)$ over $\mathbf{Q}_2(\sqrt{\Delta})$. Thus $H \cong C_4$, Φ_2 is cyclic, $G \cong C_8$ is abelian, and $U_2(s) = (1 - 2/2^s)^{-1}$ in this case. For the case where u_6 is even, we show that there is no solution in $\mathbf{Z}_2[\omega]$ to the previous y-sextic. Writing $y = a + b\omega$, we see that the left side of the sextic relation is not divisible by 2^8 unless a is odd and $2 \| b$. But in this case, we get that $2^2 \| (y^2 + 6y - 3)$, so that the left side has even 2-valuation while that of the the right side is odd. Thus there are no solutions to the y-sextic in $\mathbf{Q}_2(\sqrt{\Delta})$, implying that $f(x)$ is irreducible in this field. So $H \cong D_8$, $\Phi_2 \cong Q_8$ is not cyclic, and $U_2(s) \equiv 1$. This proves the Theorem.

We lastly consider the case where F has additive reduction at $p = 3$. The table on page 121 of Coates–Schmidt tells us that if $\delta_3 = 3$ or $\delta_3 = 5$, then $U_3(s) \equiv 1$ and $\tilde{\delta}_3 = (1 + \delta_3)/2$. Furthermore, in the case $\delta_3 = 2$, the same table says that we must have $U_3(s) = (1 + 3/3^s)^{-1}$ and $\tilde{\delta}_3 = 1$. It is only in the case $\delta_3 = 4$ that there is ambiguity, though here we have always have $\tilde{\delta}_3 = 2$ and $U_3(s) = (1 \pm 3/3^s)^{-1}$.

Theorem. *If $3^4 \| N_F$ with F twist-minimal, then either $c_4 \equiv 9\ (27)$ with $3^3 \| c_6$ but $c_6/27 \not\equiv \pm 1\ (9)$, or else $3^3 \| c_4$ and $3^5 \| c_6$. If $c_4 \equiv 9\ (27)$, then $U_3(s) = (1 + 3/3^s)^{-1}$ if $c_6 \equiv \pm 54\ (243)$ and $U_3(s) = (1 - 3/3^s)^{-1}$ if $c_6 \equiv 108\ (243)$. If $3^5 \| c_6$, then $U_3(s) = (1 - 3/3^s)^{-1}$ if $c_4 \equiv 27\ (81)$ and $U_3(s) = (1 + 3/3^s)^{-1}$ if $c_4 \equiv 54\ (81)$.*

Again the first statement follows from an exercise using Tate's algorithm. For the second part, we compute whether $G = \mathrm{Gal}\big(\mathbf{Q}_3(F_4)/\mathbf{Q}_3\big)$ is abelian, which by Lemma 1.4 of Coates–Schmidt will tell us which sign occurs in $U_3(s)$. We write H for the quotient of G by the conjugation operation, so that H is the Galois group of $\mathbf{Q}_3(F_4^x)/\mathbf{Q}_3$, the extension by just the x-coordinates of the points of exact order 4, noting that $H \subseteq PGL_2(\mathbf{Z}/4\mathbf{Z})$.

We first consider the case $c_4 \equiv 9\ (27)$ and $3^3 \| c_6$ with $c_6/27 \not\equiv \pm 1\ (9)$, writing $u_4 = c_4/9$ and $u_6 = c_6/27$. The 2-torsion polynomial $x^3 - 3^5 u_4 x - 3^6 2 u_6$ has no $\mathbf{Q}_3$-roots (thus is irreducible) and its discriminant is $2^2 3^{15} (u_4^3 - u_6^2)$. When u_6^2 is 4 mod 9, this is nonsquare, and so $\mathrm{Gal}\big(\mathbf{Q}_3(F_2^x)/\mathbf{Q}_3\big) \cong D_6$. This is a quotient group of G which is hence also nonabelian, so that $U_3(s) = (1 + 3/3^s)^{-1}$. When u_6^2 is 7 mod 9, the discriminant is square, implying that $\mathbf{Q}_3(F_2^x)$ is a normal cubic subfield of $\mathbf{Q}_3(F_4^x)$, which gives us a normal index 3 subgroup in H by the Galois correspondence. Ramification theory implies that the wild inertia group of order 3 is a normal subgroup of the Galois group G, and its quotient upon conjugation becomes an order 3 normal subgroup in H. The only subgroups of $PGL_2(\mathbf{Z}/4\mathbf{Z})$ that have normal subgroups of both index 3 and order 3 are C_3 and C_6. (In actuality, an arduous computation shows the exact-4-torsion polynomial is always irreducible in $\mathbf{Q}_3$ in this case, so that $H \cong C_6$.) Considering the action of conjugation, these H-possibilities imply that G is one of C_3, C_6, or $C_6 \times C_2$, in each case abelian. Thus $U_3(s) = (1 - 3/3^s)^{-1}$ in this case.

We finally turn to the case where $3^3\|c_4$ and $3^5\|c_6$, writing $u_4 = c_4/3^3$ and $u_6 = c_6/3^5$. The 2-torsion polynomial $x^3 - 3^6u_4x - 3^8 2u_6$ is irreducible and has discriminant $2^2 3^{18}(u_4^3 - 3u_6^2)$. This is nonsquare when u_4 is 2 mod 3, which as above implies that G is nonabelian, so that $U_3(s) = (1 + 3/3^s)^{-1}$. This discriminant is square when u_4 is 1 mod 3, so $\mathbf{Q}_3(F_2^x)$ is again a normal cubic subfield, and it follows as above that G is abelian and $U_3(s) = (1 - 3/3^s)^{-1}$. This proves the theorem and completes the description of the extra Euler factors and symmetric-square conductor in the functional equation.

As an example of all the above, take $E = [0, 0, 0, -8892, 731025]$, where $N = 2^2 \cdot 3^2 \cdot 19^2 \cdot 37 \cdot 1697$. Twisting by -3 gives $F = [0, 0, 0, -988, -27075]$ which has good reduction at $p = 3$. Since $l_3 = 0$ for this latter curve, the modular degree of E is 32 times that of F (assuming each Manin constant is 1). We have that $\delta_2 = 2$ so that $U_2(s) = (1 + 2/2^s)^{-1}$ and $\tilde{\delta}_2 = 1$. We compute (using F) that $c_4 = 47424$ and $c_6 = 23392800$, so that $19^2|c_6$ but $19\|c_4$. Hence $U_{19}(s) = (1 + 19/19^s)^{-1}$ and $\tilde{N} = 2 \cdot 19 \cdot 37 \cdot 1697$, which is much less than N.

3. Optimal Curves

Let ϕ be a modular parametrisation from $X_0(N)$ to E. We say that ϕ (and also the parametrised curve) is **optimal** if every modular parametrisation (from $X_0(N)$) to an isogenous curve of E factors through ϕ. By algebraic considerations, there is a unique such curve in any isogeny class (see [Birch and Swinnerton-Dyer 1975], where the concept is called strong). Similarly, if we consider parametrisations from $X_1(N)$, there is again the notion of optimality. Alternatively, we can view the parametrisations as coming from the relevant Jacobians, and then optimality simply means that the kernel is connected. Taking the canonical Néron differential $\omega = dx/(2y + a_1x + a_3)$, we define the complex volume $\Omega = \int_{E(\mathbf{C})} \omega \wedge \bar{\omega}$ (which is $2/i$ times the volume of the fundamental parallelogram). In terms of the Parshin–Faltings height H, we have that $H = \sqrt{2\pi/i\Omega}$. In [Stevens 1989], we find the following conjectures: in any isogeny class, the curve with largest $|\Omega|$, that is, minimal height, is optimal for $X_1(N)$ (Conjecture II, pg. 77), and has Manin constant (from $X_1(N)$) equal to 1 (Conjecture I, pg. 76). Indeed, this latter conjecture implies that the $X_1(N)$-Manin constant for *any* curve is 1 (see the comments on page 85). This is not true for $X_0(N)$, as $[0, 1, 1, 0, 0]$ has a $X_0(11)$-Manin constant of 5. However, if the optimal curves for $X_0(N)$ and $X_1(N)$ are the same (as they frequently are — only 95 counterexamples exist for $N \le 10000$), and the $X_0(N)$-Manin constant for the strong curve is its conjectural value of 1, then all the isogenous curves have $X_0(N)$-Manin constant of 1 also (this follows in the same manner as the argument on page 85 of [Stevens 1989]). Moreover, by assuming this Stevens conjecture we can ameliorate the seemingly difficult process of determining the $X_0(N)$-optimal curve. Note that the process of [Cremona 1992, Section 3.8] allows us to list all the isogenous curves for a given curve, and computing Ω for each takes little time, so under our assumption of the Stevens conjectures, computing the optimal curve for $X_1(N)$ is easy.

We next show how to pass from the $X_1(N)$-optimal curve to the $X_0(N)$-optimal curve. We first define the full period lattice Λ_f^G of a congruence group $G \subseteq \Gamma_0(N)$ for a weight 2 newform f of level N. This is defined as the image of the homomorphism $I_f : G \to \mathbf{C}$ given by $I_f(\gamma) = 2\pi i \int_\infty^{\gamma(\infty)} f(z)\, dz$. Under our assumptions it follows that Λ_f^G is a discrete rank 2 subgroup of $\mathbf{C}$, and if we let $E_f^G = \mathbf{C}/\Lambda_f^G$, then E_f^G is the G-optimal curve (see [Birch and Swinnerton-Dyer 1975]). We next define the invariant period lattice of E. For simplicity of exposition, assume that the discriminant of E is positive (see Algorithm 7.4.7 of [Cohen 1993] for the other case). We write E in the form $y^2 = g(x) = 4x^3 + b_2x^2 + 2b_4x + b_6$, and let $e_1 > e_2 > e_3$ be the (necessarily real) roots of $g(x)$. Put $\omega_1 = \pi/\mathrm{agm}\big(\sqrt{e_1 - e_3}, \sqrt{e_1 - e_2}\big)$ and $\omega_2 = i\pi/\mathrm{agm}\big(\sqrt{e_1 - e_3}, \sqrt{e_2 - e_3}\big)$, where agm is the arithmetic-geometric mean. Then the invariant period lattice of E is that which is generated (over $\mathbf{Z}$) by ω_1 and ω_2. For an optimal curve, the Manin constant can be shown to be the lattice index of the invariant period lattice in the full period lattice (see [Birch and Swinnerton-Dyer 1975]). Note that the full period lattice depends on the group but not which isogenous curve is chosen (being a function only of the newform), while the invariant period lattice depends on the choice of isogenous curve but not the group. There appears to be no standard terminology in the literature for this distinction between the lattices.

Lemma. *Let f be a weight 2 newform of level N. Let L_0 and L_1 be the full period lattices of $\Gamma_0(N)$ and $\Gamma_1(N)$ for f respectively, and M a lattice with $L_1 \subseteq M \subseteq L_0$. Then we have a surjective homomorphism $h : \big(\mathbf{Z}/N\mathbf{Z}\big)^* \to L_0/M$.*

Corollary. *Let f be a weight 2 newform of level N. Suppose that M is a lattice with $L_1 \subseteq M \subset L_0$. Then there is some prime p which divides $\phi(N)$ and some subgroup P of $(\mathbf{Z}/N\mathbf{Z})^\star$ of order p such that $h(d) \neq \mathrm{id}$ for any d for which $d^{\lambda(N)/p}$ generates P.*

We note that the surjective homomorphism $I_f : \Gamma_0(N) \to L_0$ restricts to a surjective homomorphism $\bar{I}_f : \Gamma_1(N) \to L_1$, and so induces a surjective homomorphism $\left(\mathbf{Z}/N\mathbf{Z}\right)^* \cong \Gamma_0(N)/\Gamma_1(N) \to L_0/L_1$. Now if M is any lattice with $L_1 \subseteq M \subseteq L_0$, we obtain an induced surjective homomorphism $h : \left(\mathbf{Z}/N\mathbf{Z}\right)^* \to L_0/M$. Explicitly, $h(D) = I_f(\gamma) \pmod{M}$ where $\gamma \in \Gamma_0(N)$ is any matrix with D as its lower-right entry. Since all the groups involved are finite and abelian, the corollary (where $\lambda(N)$ is the Carmichael function, that is the exponent of the group $\mathbf{Z}/N\mathbf{Z}$) follows directly from the classification of finite abelian groups.

In our case, we can limit the choices for p by consideration of the p-isogenies of E. So now our algorithm is as follows: given an elliptic curve E, find all the isogenous curves using [Cremona 1992], and specifically the one of minimal height, which we denote by $\tilde{E}$. By our assumption of the Stevens conjecture, the full period lattice of $\Gamma_1(N)$ for f associated to E is the invariant period lattice of $\tilde{E}$, which we call M. For each plausible p-subgroup of $(\mathbf{Z}/N\mathbf{Z})^\star$ (or better, a basis for them), we find some d such that $d^{\phi(N)/p}$ generates the subgroup, and see if $h(d) = \mathrm{id}$. If it is, we continue, while if not, we enlarge M and iterate. At the end of the process, we have L, the full period lattice of $\Gamma_0(N)$. If there is a curve in the isogeny class with L as its invariant lattice, then this curve is the desired $X_0(N)$-optimal curve, as else the Manin constant of the optimal curve would be nonintegral. There is no proof that L must be the invariant lattice for some curve in the isogeny class, but if it is not, then the $X_0(N)$-Manin constant would not be 1, contrary to conjectural behaviour. Computing $h(d)$ is expedited by a method of [Cremona 1997], and can be done in $(d\sqrt{N})^{5/4}$ time in practice. Standard conjectures of analytic number theory imply that we need not take d very large, so this amount is very reasonable compared to the other parts of the modular degree algorithm. Alternatively, in [Stein and Watkins 2002], the authors conjecture what they believe to be a complete classification of curves with differing optimal curves from $X_1(N)$ and $X_0(N)$; they find 3 families (one being the Setzer–Neumann curves considered below) where the optimal curves (conjecturally) differ by a 2-isogeny, and a family where they differ by a 3-isogeny, to go with the 4-isogeny examples 15A and 17A and the 5-isogeny example 11A.

As an example, we consider $E = [0, 1, 1, -3343, 73293]$ of conductor 8027. This curve is of minimal height in its isogeny class, having real volume ≈ 0.422966, while the 3-isogenous curve $F = [0, 1, 1, -3243, 77986]$ has a volume smaller by a factor of 3. Now $\phi(8027)$ equals 7656 which is divisible by 3. Using $d = 2$, we have $2^{7656/3} \equiv 2699 \not\equiv 1 \pmod{8027}$, and find that $I_f\left(\begin{pmatrix} 4024 & 1 \\ 8027 & 2 \end{pmatrix}\right) \approx -3.591969$, which is $-4/3$ times the real period ω_1 of E (≈ 2.6947). Hence $h(d) \neq \mathrm{id}$, and we quickly conclude (subject to our belief of the Stevens conjecture) that F is the $X_0(N)$-optimal curve.

4. Experimental Results

There are four main data sets of isogeny classes of elliptic curves with which we did experiments. The first set is simply the 38042 classes with conductor less than 10000, a list which has been compiled by Cremona, using his modular symbol technique. We call this set S_1. The others are (almost) subsets of the large set of data found in [Brumer and McGuinness 1990], who made a list of 310716 curves for which $|\Delta|$ is prime and less than 10^8. However, the curve $[0, 0, 1, -10000, 384900]$ inexplicably appears twice in their data, and a pair of isogenous curves are computed by their method for $N = 11, 17, 19, 37$. Hence there are only 310711 isogeny classes. Our set S_2 is related to the 860 Setzer–Neumann curves (see [Setzer 1975], [Neumann 1971], and [Neumann 1973]) with prime discriminant $p \leq 10^8$ of the form $p = u^2 + 64$, but we choose a different representative in the isogeny class than Brumer–McGuinness does. Other than the four above examples, these are the only curves with prime (absolute value of) discriminant for which there is more than just the one curve in the isogeny class, there being two isogenous curves in this case. A direct computation shows that the curve with prime discriminant $p = u^2 + 64$ is the one of minimal height, while the work of [Mestre and Oesterlé 1989] (following directly from the appendix of [Mazur 1977]) implies that the isogenous curve with discriminant $-p^2$ is the $X_0(p)$-optimal curve. We denote by S_2 this set of 860 optimal curves. The third set of curves we consider is all the non-Setzer–Neumann curves in the Brumer–McGuinness list with $|\Delta| \leq 10^7$, additionally excluding the four above curves possessing nontrivial isogenies. This set (called S_3) has 52878 curves. Finally, the fourth set (S_4) is the 804 curves in the Brumer–McGuinness list which have rank 4. We have also computed the modular degree for the five rank 5 examples of Brumer–McGuinness, and about 50 other rank 5 curves from the data of Tom Womack (available from `http://www.maths.nott.ac.uk/personal/jec/ftp/data/womack.data`).

The set S_1, while being the most comprehensive, is perhaps the worst for data analysis, as it contains quadratic twists and other non-semistable phenomena such as the motivic/analytic symmetric-square difference. However, it does provide a good testing ground for an implementation of the algorithm. On the other hand, the set S_2 has some very nice properties, especially that Ω follows a simple trend. The set S_3 is sufficiently large to produce data on a larger scale. The fact that there is only one prime dividing the conductor also helps to make this a useful set for analysis. Finally, the set S_4 was taken simply to accrue more data for the rank conjecture (see below). For much of S_1, Cremona has rigourously determined the $X_0(N)$-optimal curve and corresponding Manin constant (and even the modular degree); in particular the verification is complete for $N < 8000$, and will be continued for all of S_1. By the work of [Abbes and Ullmo 1996] and the aforementioned [Mestre and Oesterlé 1989], we know the optimal curve and Manin constant for the other three sets. So, except for a few cases in S_1, we can be assured that we are actually computing the modular degree of the optimal curve for each of the curves considered. In all known cases, the Manin constant of the optimal curve is indeed 1.

4.1. Size distribution of $\deg\phi$ and $L(\mathrm{Sym}^2 E, 2)$

First we consider the size distribution of $\deg\phi$. This is largely controlled by Ω, with $L(\mathrm{Sym}^2 E, 2)$ playing a lesser role (similar to the number field case with the regulator and L-function). For curves of prime conductor, the ABC-conjecture predicts that $\Omega \approx N^{-1/6}$, while we can show that $L(\mathrm{Sym}^2 E, 2) \ll (\log N)^3$. In particular, average behavior is not so relevant as are the extreme cases vis-a-vis the ABC-conjecture. Instead of looking at the distribution of the modular degree, we look at how $L(\mathrm{Sym}^2 E, 2)$ is distributed. The set S_3 is the largest, and we look at it first. One thing to which we can compare this is $L(1, \chi)$ where χ is a quadratic character, so that this is the value at the edge of the critical strip of the L-function of a quadratic field. We chose to consider only imaginary quadratic fields since the L-values are slightly easier to compute in this case. We can also restrict the (absolute value of the) discriminant to be prime in order to correlate better with the data from S_3 which we have for $L(\mathrm{Sym}^2 E, 2)$. The distribution of $L(1, \chi)$-values for the prime discriminants up to 10^7 is displayed in Figure 1. Therein we also display the distribution of $L(1, \chi)$ for all negative fundamental discriminants up to 10^6, and those of the motivic and analytic $L(\mathrm{Sym}^2 E, 2)$ for the 20726 minimal quadratic twists in the set S_1. We use the logarithm of the L-value as it seems to be the more natural measure, due to the Euler product representation of the L-function, implying its positivity at the edge of the critical strip. In fact, by the appendix of [Hoffstein and Lockhart 1995], we have the equivalent of "no Siegel zeros" for the symmetric-square L-function. In the figure, the horizontal axis is divided into 100 parts, and the vertical axis indicates what proportion of the data falls into the intervals implied by the division, with each data set being line-connected in order to ease the viewing. There seems to be much more difference between the sets S_1 and S_3 than there is between the corresponding sets of fundamental discriminants. If we restrict S_1 to semistable curves, this does not change matters much. Many authors have determined the distribution function for $L(1, \chi)$ when averaging over all negative fundamental discriminants. The first appears to be [Chowla and Erdős 1951] who used methods of additive functions. A similar technique appears in [Elliott 1980], while [Barban 1964] used the large sieve to evaluate the moments of $L(1, \chi)$, from which the distribution is recoverable. The author has obtained a similar result for $L(\mathrm{Sym}^2 E, 2)$. The main tool is the large sieve for GL_n of [Duke and Kowalski 2000]. When averaging over all weight 2 modular forms (and not just the rational ones as here) the resulting distribution has already been found in [Royer 2001] (see also the Maaß form case in [Luo 1999]), again using the Duke–Kowalski large sieve. While of theoretical value, these results do not imply much beyond that the distributions should be different.

We next turn to the sets S_2 and S_4. Now all the curves in S_2 have rank zero, while the ones in S_4 have rank four. This could have an effect. It turns out to have more of an impact on S_4, and below we try to explain why. Figure 2 shows the $L(\mathrm{Sym}^2 E, 2)$ distribution for S_2, S_3, and S_4. The graph for S_2 takes a much different shape from that of S_3, being more peaked around the mean. It is unclear why this should be so, perhaps simply because the Ω-N correspondence is better behaved. Meanwhile, S_4 is violently different, suggesting that we might want to sort the data by rank. Figure 3 considers the four subsets of S_3 with fixed ranks 0-3, and also includes set S_4. As can be seen, the rank 0-3 data is largely similar, while rank 4 is rather different. My guess is that we do not have enough data; when the rank is comparatively large vis-a-vis the conductor, we might expect most of the small Frobenius traces l_p to be as negative as possible, or near $-2\sqrt{p}$. Upon squaring and subtracting p to get the symmetric-square coefficient, we are left with $3p$, oriented strongly in the positive direction. Hence we could expect the $L(\mathrm{Sym}^2 E, 2)$ value to be larger than expected. But as the conductor gets large with respect to a fixed rank, there is no reason to expect such behaviour. So while the data show some indication of a rank effect on the size of the L-value, it might be misleading. Finally, the extreme values of $L(\mathrm{Sym}^2 E, 2)$ in the data samples are ≈ 12.68 for the rank 0 curve $[0, 0, 1, -7, -89]$ of conductor 3380723, and ≈ 0.4176 for the rank 1 curve $[0, 0, 1, -58, -118]$ of conductor 6497461. As evinced by the figures, there is much more clustering around the lower value.

4.2. Arithmetic properties of $\deg\phi$

Next we turn to the arithmetic properties of $\deg\phi$. The first conjecture suggested by the data is:

Conjecture. *If an elliptic curve has rank r, then 2^r divides its modular degree.*

This conjecture holds for all curves checked in the experiment, in particular for the rank 4 and rank 5 curves, while the rank 0 curves often had odd modular degree. For some high-rank composite-conductor examples, even more powers of 2 divide the modular degree, as might be suggested from an analysis of Atkin-Lehner involutions. Using a refined Eisenstein Theorem due to Emerton, in joint work with M. Baker and W. Stein we have shown that about half of the Setzer–Neumann curves have the property that the $X_0(p)$-optimal curve in the isogeny class has odd modular degree. Writing the prime discriminant of the minimal height curve as $p = u^2 + 64$ with u congruent to 3 mod 4, the modular degree of the optimal curve is odd if and only if u is 3 mod 8 — unfortunately, we cannot prove that there are infinitely primes of this form. In this case of Setzer–Neumann isogeny classes, the data also suggest that if u is 7 mod 8, then 2 exactly divides the modular degree, while if u is 7 mod 16, then 4 divides the modular degree (with there being no apparent pattern in 2-divisibility beyond this). However, this does not seem to follow from Emerton's refined Eisenstein Theorem, and there is even some reason to think it might fail occasionally. The data also suggest the following:

Conjecture. *If an elliptic curve of prime conductor p has odd modular degree then either $p = 17$, p is of the form $u^2 + 64$, or p is 3 mod 8.*

There are 29436 curves in S_3 with p not 3 mod 8, of which 8240 have rank 0, yet all have even modular degree. On the other hand, there are 7322 rank 0 curves in S_3 with conductor congruent to 3 mod 8, of which 4589 have odd modular degree.

In general, for a prime p, we could ask how often p divides the modular degree. For odd primes and the set S_3, the answer appears to be given closely by the Cohen–Lenstra heuristics [Cohen and Lenstra 1985], namely p divides the modular degree $\bigl(1 - \prod_k (1 - 1/p^k)\bigr) = \bigl(1/p + 1/p^2 - 1/p^5 + \cdots\bigr)$ of the time, a quantity which we shall call $CL(p)$. Of course, distinguishing between $CL(p)$ and a function of the form $\frac{p^2}{p^2+1}\frac{1}{p-1}$ is not necessarily easy. Also, there seems to be a statistically significant difference between the expected and computed percentage for $p = 3$ (see Table 1). We exclude $p = 2$ from the table, as the Cohen–Lenstra heuristics are silent here, and we have already discussed the 2-divisibility above. As for the non-zero congruence classes modulo p, the modular degree appears to be equi-distributed amongst them. Yet data from S_1 gives a vastly different percentage of curves with (e.g.) 3 dividing the modular degree (even when 3 does not divide the level), and Cremona's data also has non-equidistribution in non-zero congruence classes. Below, we consider why Cremona's nonprime conductor curves might cause such a difference, but first we restrict ourselves to just the prime conductor curves.

p	3	5	7	11	13	17	19	23	29	31	37
number	23771	12770	8607	5233	4305	3263	2822	2459	1802	1788	1415
data%	44.84	24.15	16.28	9.90	8.14	6.17	5.34	4.65	3.41	3.38	2.68
$100 \cdot CL(p)$	43.99	23.97	16.32	9.92	8.28	6.23	5.54	4.54	3.57	3.33	2.78

Table 1: Percentages of curves from S_3 with a given odd prime dividing the modular degree, compared to the percentage predicted by the Cohen–Lenstra heuristic.

If the Cohen–Lenstra heuristic is correct, one would expect there to be a relevant finite group whose order determines the data; indeed, the heuristic comes from assuming that groups appear randomly, except weighted by the number of their automorphisms. The desired groups come from lattices. We follow the exposition of [Zagier 1985]. Let E be an elliptic curve of conductor N, and f the weight 2 newform of level N associated to it. Let S be the space of integral weight 2 forms of level N with integral coefficients, so that $S = S_2\bigl(\Gamma_0(N)\bigr) \cap \mathbf{Z}\bigl[[q]\bigr]$. Let $L = [f]^\perp \cap S$, where $[f]^\perp$ is the span of eigenforms other than f, or equivalently, the orthogonal complement of f with respect to the Petersson inner product. Let e be the exponent of the finite group $G = S/(\mathbf{Z}f + L)$. If N is prime, we have that $e = \deg\phi$, which Zagier attributes to Ribet. If we expect G to be a random group in the sense of Cohen–Lenstra, and note that a prime divides the exponent of a group if and only if it divides the order of the group, we do indeed recover the heuristic that p should divide the modular degree about $CL(p)$ of the time. However, there might be reasons to expect that G is not totally random, especially when N is not a prime. In general (note that Zagier's claim is transposed), we have $\deg\phi | e$ and also $e | N^i \deg\phi$ for some i. One of the referees informs us that Ribet has recently shown that if $p \| N$, then the powers of p dividing e and $\deg\phi$ are equal (this was a conjecture of Agashe and Stein).

Data from [Jacobson 1998] on real quadratic fields might give some indication on how fast we could expect convergence to the Cohen–Lenstra number (assuming that it is correct), but things still seem unclear, particularly for $p = 3$. We discuss some reasons why the Cohen–Lenstra heuristic might fail for S_1. With this set, we find that the modular degree is a multiple of 3 about 78% of the time, far distant from the 44% of Cohen–Lenstra. There are a number of things which could be affecting this. The first is that when 3 divides the level, the above heuristic argument fails, and there is even doubt simply when there are square divisors of the level. The existence of a 3-isogeny is a third factor possibly affecting the heuristic. After removing the curves for which 3 divides the level, we still have 10768/14816 or almost 73% whose modular degree is divisible by 3. Further pruning the non-semistable curves gives 4125/6555 or 63%. Adding the no-3-isogeny restriction gives 3668/6004 or 61%. So even with these three considerations there still seems to be another factor affecting divisibility of the modular degree. Suppose that there are (say) 15 isogeny classes for a given conductor, and there is a 3-congruence involving a linear combination of 13 of the associated newforms. This would have an influence on the 3-divisibility of the modular degree of all 13 curves involved. In other words, the p-divisibility probabilities for the different isogeny classes at a given level are often not independent. This appears to have little effect on the prime level data, possibly because of the small expected number of curves with the same conductor. We could try to ameliorate this problem by using just the 1023 squarefree and coprime-to-3 levels for which there is only one isogeny class. This gives 510/1023 (just under 50%), which is somewhat outside the error bounds of the Cohen–Lenstra number. So it appears that the number of isogeny classes at a given level has a strong effect on the heuristic. Unfortunately, there are few good heuristics for the number of such isogeny classes, the best results being the upper bounds of [Brumer and Silverman 1996], recently improved by [Wong 2001]. And even with all the adjustments to the heuristic, it is still unclear why the prime conductor results match closely with the Cohen–Lenstra prediction while the Cremona data has more variance. Finally, the elimination of non-level-unique curves does not fix the slight $p = 3$ anomaly in our data.

5. Acknowledgements

The author would like to thank Henri Cohen and Don Zagier for help with the special L-value algorithm, Noam Elkies for information about Mestre's algorithm and for noticing that 2^5 divided the modular degree of the first few rank 5 examples computed, Frank Calegari for pointing out that a congruence-class description of $U_p(s)$ should be possible, William Stein for help with computing the level of absolute minimal quadratic twists, John Cremona for verification of modular degrees, Joe Buhler for references concerning symmetric-cube L-functions, and the anonymous referees for much good advice. Much of this work was done during the Algorithmic Number Theory semester at the Mathematical Sciences Research Institute in Fall 2000.

References

[Abbes and Ullmo 1996] A. Abbes, E. Ullmo, *À propos de la conjecture de Manin pour les courbes elliptiques modulaires* (French). Compositio Math. **103** (1996), no. 3, 269–286.

[Atkin and Lehner 1970] A. Atkin, J. Lehner, *Hecke operators on* $\Gamma_0(m)$. Math. Ann. **185** (1970), 134–160.

[Barban 1964] M. Barban, *The 'Large Sieve' Method and its applications in the Theory of Numbers* (Russian). Uspehki Mat. Nauk **21** (1966), no. 1 (127), 51–102; Russian Math Surveys **21** (1966), no. 1, 49–103.

[Birch 1991] B. Birch, *Hecke action on classes of ternary quadratic forms.* In *Computational number theory (Debrecen, 1989),* edited by A. Pethö, M. Pohst, H. Williams, and H. Zimmer, 191–212, de Gruyter, Berlin, 1991.

[Birch and Swinnerton-Dyer 1975] B, Birch, H. Swinnerton-Dyer, *Elliptic curves and modular functions.* In *Modular functions of one variable IV*, edited by B. Birch and W. Kuyk, 2–32, Lecture Notes in Math., Vol. 476, Springer, Berlin, 1975.

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

[Brumer and McGuinness 1990] A. Brumer, O. McGuinness, *The behavior of the Mordell-Weil group of elliptic curves.* Bull. Amer. Math. Soc. (N.S.) **23** (1990), no. 2, 375–382.

[Brumer and Silverman 1996] A. Brumer, J. Silverman, *The number of elliptic curves over* **Q** *with conductor* N. Manuscripta Math. **91** (1996), no. 1, 95–102.

[Buhler and Gross 1985] J. Buhler, B. Gross, *Arithmetic on elliptic curves with complex multiplication II.* Invent. Math. **79** (1985), no. 1, 11–29.

[Buhler et al. 1997] J. Buhler, C. Schoen, J. Top, *Cycles, L-functions and triple products of elliptic curves.* J. Reine Angew. Math. **492** (1997), 93–133.

[Chowla and Erdős 1951] S. Chowla, P. Erdös, *A theorem on the distribution of values of L-series.* Journ. Indian Math. Soc. **15A** (1951), 11–18.

[Coates and Schmidt 1987] J. Coates, C.-G. Schmidt, *Iwasawa theory for the symmetric square of an elliptic curve.* J. Reine Angew. Math. **375/376** (1987), 104–156.

[Cohen 1993] H. Cohen, *A course in computational algebraic number theory.* Graduate Texts in Mathematics, 138. Springer-Verlag, New York, 1993.

[Cohen 2000] H. Cohen, *Advanced topics in computational number theory.* Graduate Texts in Mathematics, 193. Springer-Verlag, New York, 2000.

[Cohen and Lenstra 1984] H. Cohen, H. Lenstra, *Heuristics on class groups of number fields.* In *Number theory, Noordwijkerhout 1983*, edited by H. Jager, 33–62, Lecture Notes in Math., 1068, Springer, Berlin, 1984.

[Cremona 1992] J. Cremona, *Algorithms for modular elliptic curves.* Cambridge University Press, Cambridge, 1992. Second edition 1997.

[Cremona 1995] J. Cremona, *Computing the degree of the modular parametrization of a modular elliptic curve.* Math. Comp. **64** (1995), no. 211, 1235–1250.

[Cremona 1997] J. Cremona, *Computing periods of cusp forms and modular elliptic curves.* Experiment. Math. **6** (1997), no. 2, 97–107.

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

[Edixhoven 1991] B. Edixhoven, *On the Manin constants of modular elliptic curves.* In *Arithmetic algebraic geometry (Texel, 1989),* edited by G. van der Geer, F. Oort, and J. Steenbrink, 25–39, Progr. Math., 89, Birkhäuser Boston, Boston, MA, 1991.

[Eichler 1973] M. Eichler, *The basis problem for modular forms and the traces of the Hecke operators.* In *Modular functions of one variable I*, edited by W. Kuyk, 75–151, Lecture Notes in Math., Vol. 320, Springer, Berlin, 1973.

[Elliott 1980] P. Elliott, *Probabilistic number theory II. Central limit theorems.* Grundlehren der Mathematischen Wissenschaften, 240. Springer-Verlag, Berlin-New York, 1980.

[Flach 1993] M. Flach, *On the degree of modular parametrizations.* In *Séminaire de Théorie des Nombres, Paris 1991–92,* edited by S. David, 23–36, Progr. Math., 116, Birkhäuser Boston, Boston, MA, 1993.

[Frey 1987] G. Frey, *Elliptic curves and solutions of* $A - B = C$. In *Séminaire de Théorie des Nombres, Paris 1985–86,* edited by C. Goldstein, 39–51, Progr. Math., 71, Birkhäuser Boston, Boston, MA, 1987.

[Frey and Müller 1999] G. Frey, M. Müller, *Arithmetic of modular curves and applications.* In *Algorithmic algebra and number theory (Heidelberg, 1997),* edited by B. Matzat, G.-M. Greuel, and G. Hiss, 11–48, Springer, Berlin, 1999.

[Gelbart and Jacquet 1978] S. Gelbart, H. Jacquet, *A relation between automorphic representations of GL(2) and GL(3).* Ann. Sci. École Norm. Sup. (4) **11** (1978), no. 4, 471–542.

[Gross and Zagier 1986] B. Gross, D. Zagier, *Heegner points and derivatives of L-series.* Invent. Math. **84** (1986), no. 2, 225–320.

[Hoffstein and Lockhart 1994] J. Hoffstein, P. Lockhart, *Coefficients of Maass forms and the Siegel zero.* Appendix by D. Goldfeld, J. Hoffstein, and D. Lieman. Ann. of. Math. (2) **140** (1994), no. 1, 161–181.

[Jacobson 1998] M. Jacobson, *Experimental results on class groups of real quadratic fields.* In *Algorithmic number theory (Portland, OR, 1998),* edited by J. Buhler, 463–474, Lecture Notes in Comput. Sci., 1423, Springer, Berlin, 1998.

[Katz and Mazur 1985] N. Katz, B. Mazur, *Arithmetic moduli of elliptic curves.* Annals of Mathematics Studies, 108. Princeton University Press, Princeton, NJ, 1985.

[Kohel and Stein 2000] D. Kohel, W. Stein, *Component groups of quotients of $J_0(N)$.* In *Algorithmic number theory (Leiden, 2000),* edited by W. Bosma, 405–412, Lecture Notes in Comput. Sci., 1838, Springer, Berlin, 2000.

[Lavrik 1967] A. Lavrik, *Functional equations of the Dirichlet functions* (Russian). Izv. Akad. Nauk SSSR Ser. Mat. **31** (1967), 431–442.

[Luo 1999] W. Luo, *Values of symmetric square L-functions at 1.* J. Reine Angew. Math. **506** (1999), 215–235.

[Manin 1972] J. Manin, *Parabolic points and zeta functions of modular curves* (Russian). Izv. Akad. Nauk SSSR Ser. Mat. **36** (1972), 19–66.

[Mazur 1977] B. Mazur, *Modular curves and the Eisenstein ideal.* Inst. Hautes Études Sci. Publ. Math. **47** (1977), 33–186 (1978).

[Merel 1993] L. Merel, *Intersections sur des courbes modularies* (French). Manuscripta Math. **80** (1993), no. 3, 283–289.

[Merel 1995] L. Merel, *Homologie des courbes modulaires affines et paramétrisations modulaires* (French). In *Elliptic Curves, modular forms, and Fermat's last theorem (Hong Kong, 1993),* edited by J. Coates and S.-T. Yau, 110–130, Series in Number Theory, I, International Press, Cambridge, MA, 1995.

[Mestre 1986] J.-F. Mestre, *La méthode des graphes. Exemples et applications* (French). In *Proceedings of the international conference on class numbers and fundamental units of algebraic number fields (Katata, 1986),* (no editor listed), 217–242, Nagoya Univ., Nagoya, 1986.

[Mestre and Oesterlé 1989] J.-F. Mestre, J. Oesterlé, *Courbes de Weil semi-stables de discriminant une puissance m-ième* (French). J. Reine Angew. Math. **400** (1989), 173–184.

[Mestre and Schappacher 1991] J.-F. Mestre, N. Schappacher, *Séries de Kronecker et fonctions L des puissances symétriques de courbes elliptiques sur* $\mathbf{Q}$ (French). In *Arithmetic algebraic geometry (Texel, 1989),* edited by G. van der Geer, F. Oort, and J. Steenbrink, 209–245, Progr. Math., 89, Birkhäuser Boston, Boston, MA, 1991.

[Neumann 1971] O. Neumann, *Elliptische Kurven mit vorgeschriebenem Reduktionsverhalten. I.* Math. Nachr. **49** (1971), 107–123.

[Neumann 1973] O. Neumann, *Elliptische Kurven mit vorgeschriebenem Reduktionsverhalten. II.* Math. Nachr. **56** (1973), 269–280.

[Pizer 1976] A. Pizer, *On the arithmetic of quaternion algebras II.* J. Math. Soc. Japan **28** (1976), no. 4, 676–688.

[Royer 2001] E. Royer, *Statistiques de la variable aléatoire* $L(\mathrm{Sym}^2 f, 1)$ (French). Mathematische Annalen **321**, no. 3, 667–687.

[Schoof 1985] R. Schoof, *Elliptic curves over finite fields and the computation of square roots mod p.* Math. Comp. **44** (1985), no. 170, 483–494.

[Setzer 1975] B. Setzer, *Elliptic curves of prime conductor.* J. London Math. Soc. (2) **10** (1975), 367–378.

[Shimura 1975] G. Shimura, *On the holomorphy of certain Dirichlet series.* Proc. London Math. Soc. (3) **31** (1975), no. 1, 79–98.

[Shimura 1976] G. Shimura, *The special values of the zeta functions associated with cusp forms.* Comm. Pure Appl. Math **29** (1976), no. 6, 783–804.

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

[Stevens 1989] G. Stevens, *Stickelberger elements and modular parametrizations of elliptic curves.* Invent. Math. **98** (1989), no. 1, 75–106.

[Tate 1975] J. Tate, *Algorithm for determining the type of a singular fiber in an elliptic pencil.* In *Modular functions of one variable IV*, edited by B. Birch and W. Kuyk, 33–52, Lecture Notes in Math., Vol. 476, Springer, Berlin, 1975.

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

[Wong 2001] S. Wong, *Exponents of Class Groups and Elliptic Curves.* J. Number Theory **89** (2001), no. 1, 114–120.

[Zagier 1985] D. Zagier, *Modular parametrizations of elliptic curves.* Canad. Math. Bull. **28** (1985), no. 3, 372–384.

[Zagier and Gangl 2000] D. Zagier, H. Gangl, *Classical and elliptic polylogarithms and special values of L-series.* In *The arithmetic and geometry of algebraic cycles (Banff, AB, 1998),* edited by B. Gordon, J. Lewis, S. Müller-Stach, S. Saito, and N. Yui, 561–615, NATO Sci. Ser. C Math. Phys. Sci., 548, Kluwer Acad. Publ., Dordrecht, 2000.

Figure 1: Symmetric Square critical values compared to those from Imaginary Quadratic Fields

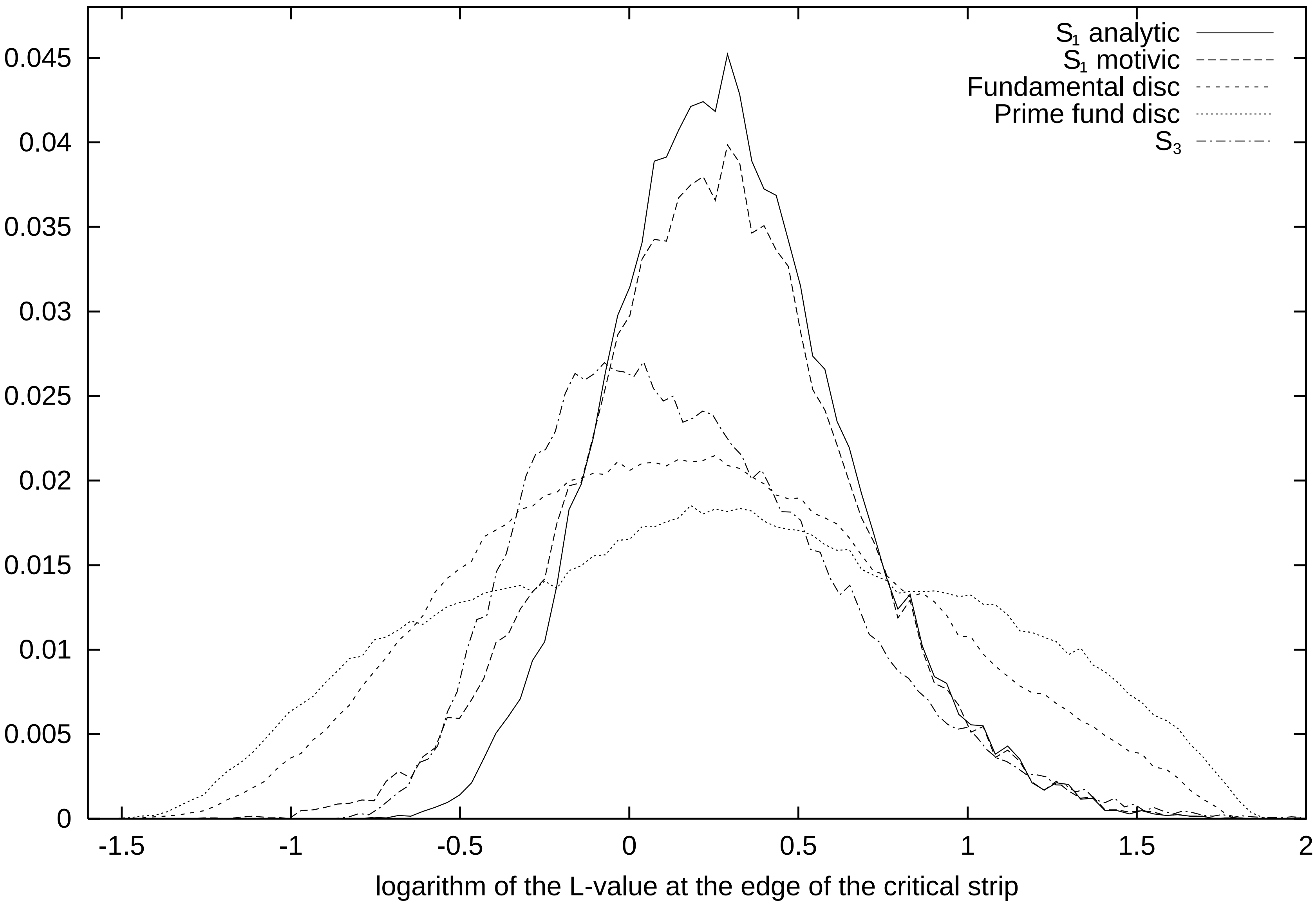

Figure 2: Special L-value distributions for sets S_2, S_3, and S_4

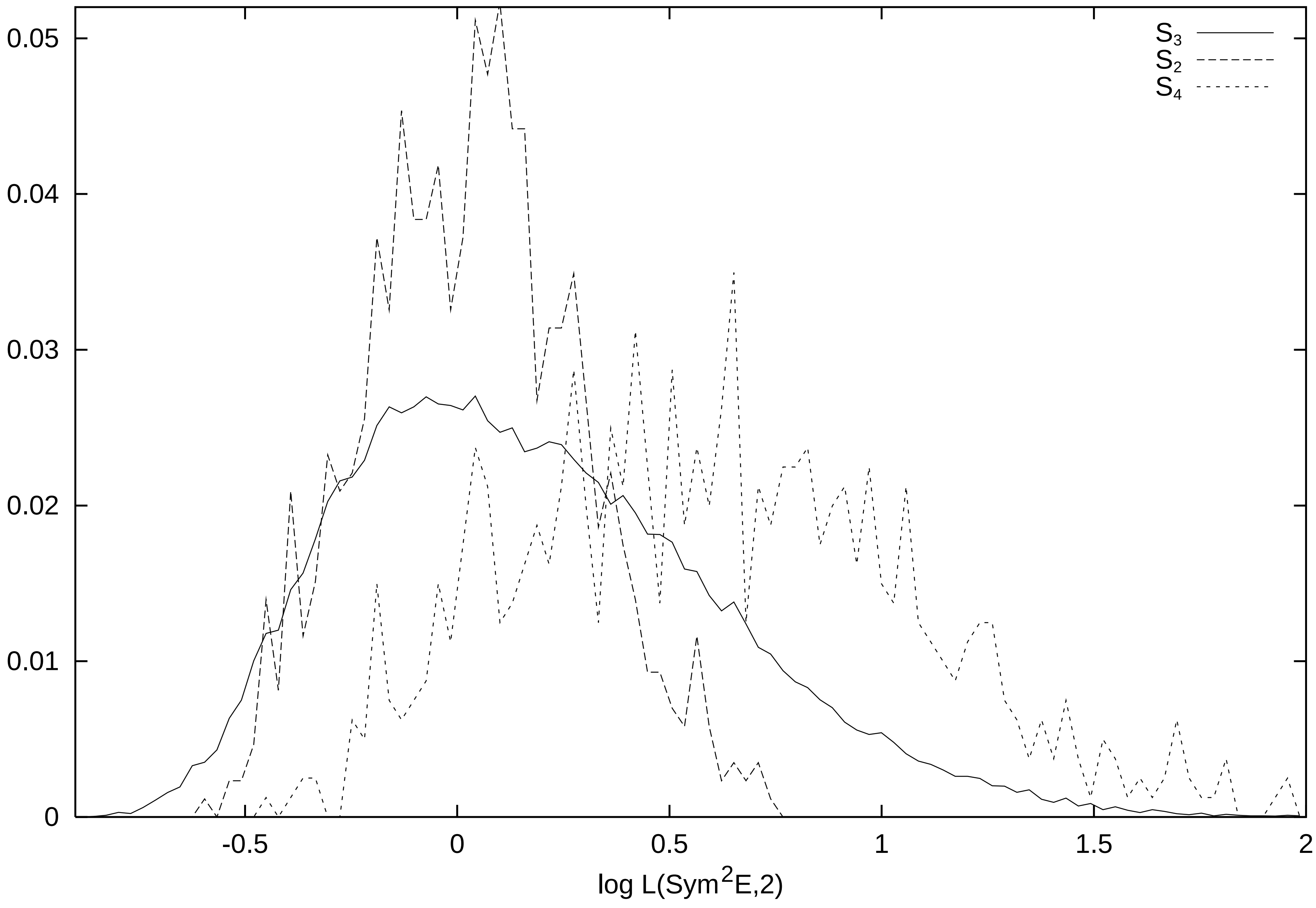

Figure 3: L-value distribution demarcated by rank