A Database of Elliptic Curves—First Report

William A. Stein[1] and Mark Watkins[2]

[1] Harvard University
was@math.harvard.edu
http://modular.fas.harvard.edu
[2] The Pennsylvania State Univerisity
watkins@math.psu.edu
http://www.math.psu.edu/watkins

1 Introduction

In the late 1980s, Brumer and McGuinness [2] undertook the construction of a database of elliptic curves whose discriminant Δ was both prime and satisfied $|\Delta| \leq 10^8$. While the restriction to primality was nice for many reasons, there are still many curves of interest lacking this property. As ten years have passed since the original experiment, we decided to undertake an extension of it, simultaneously extending the range for the type of curves they considered, and also including curves with composite discriminant. Our database can be crudely described as being the curves with $|\Delta| \leq 10^{12}$ which either have conductor smaller than 10^8 or have prime conductor less than 10^{10}—but there are a few caveats concerning issues like quadratic twists and isogenous curves. For each curve in our database, we have undertaken to compute various invariants (as did Brumer and McGuinness), such as the Birch–Swinnerton-Dyer L-ratio, generators, and the modular degree. We did not compute the latter two of these for every curve. The database currently contains about 44 million curves; the end goal is find as many curves with conductor less than 10^8 as possible, and we comment on this direction of growth of the database below. Of these 44 million curves, we have started a first stage of processing (computation of analytic rank data), with point searching to be carried out in a later second stage of computation.

Our general frame of mind is that computation of many of the invariants is rather trivial, for instance, the discriminant, conductor, and even the isogeny structure. We do not even save these data, expecting them to be recomputable quite easily in real time. For instance, for each isogeny class, we store only one representative (the one of minimal Faltings height), as we view the construction of isogenous curves as a "fast" process. It is only information like analytic ranks, modular degrees (both

of which use computation of the Frobenius traces a_p), and coordinates of generators that we save; saving the a_p would take too much storage space. It might be seen that our database could be used a "seed" for other more specialised databases, as we can quickly calculate the less time-consuming information and append it to the saved data.

2 Generating the Curves.

While Brumer and McGuinness fixed the a_1, a_2, a_3 invariants of the elliptic curve (12 total possibilities) and then searched for a_4 and a_6 which made $|\Delta|$ small, we instead decided to break the c_4 and c_6 invariants into congruence classes, and then find small solutions to $c_4^3 - c_6^2 = 1728\Delta$. We write $c_4^\star$ for the least nonnegative residue of c_4 modulo 576, and $c_6^\star$ for the least nonnegative residue of c_6 modulo 1728. The work of Connell [3] gives necessary and sufficient conditions on c_4 and c_6 for an elliptic curve with such invariants to exist. We first need that $c_6 \equiv 3 \pmod 4$ (when it follows that c_4 is odd), or $2^4 \mid c_4$ and $c_6 \equiv 0, 8 \pmod{32}$, and secondly we require a local condition at the prime 3, namely that $c_6 \not\equiv \pm 9 \pmod{27}$. Using this information and the fact that $1728 \mid (c_4^3 - c_6^2)$, this leads to 288 possible $(c_4^\star, c_6^\star)$ pairs.

For each fixed such $(c_4^\star, c_6^\star)$ pair, we can simply loop over c_4 and c_6, finding the curves with $|\Delta| \leq 10^{12}$. Of course, it is only under the ABC-conjecture that we would have an upper bound on c_4 to ensure that we would have found all such curves, and even then the bound would be too large. Our method was to take $c_4 \leq 1.44 \cdot 10^{12}$ in this first step; in any case, curves with larger c_4 are most likely found more easily using the method of Elkies [5].

2.1 Minimal Twists

In the sequel, we shall write E_d for the quadratic twist of E by d. For each (c_4, c_6) pair (again with $c_4 \leq 1.44 \cdot 10^{12}$) which satisfies the $|\Delta| \leq 10^{12}$ condition, we then determine whether this curve is minimal—not only in the traditional sense for its minimal discriminant, but also whether it is has the minimal discriminant in its family of quadratic twists. For $p \geq 5$, this is rather easy to determine; unless $p^6 \mid \Delta$ and $p \mid c_4$, the curve is minimal for quadratic twists (the only difference between this and the standard notion of minimality is that the exponent here is 6 instead of 12). If both the above conditions hold, then we throw the curve out, as $E_{\tilde{p}}$, where $\tilde{p} = \left(\frac{-1}{p}\right) p$, is a curve with lesser discriminant (which will be

found by our search procedure). Given that the curve is minimal at a prime divisor $p \geq 5$ of Δ, the local conductor at p is p^2 if $p \mid c_4$ and p^1 otherwise.

The case with $p = 3$ is a bit harder. By Connell's conditions, we see that if $3^9 \mid (c_4^3 - c_6^2)$ while $3 \mid c_6$ but 3^5 does not exactly divide c_6, then E_{-3} is a curve with invariants $(c_4/9, -c_6/27)$ which has the discriminant reduced by 3^6. This is the only prohibition against the curve being the minimal twist at 3. If $3 \mid\mid c_4$, the curve has good reduction (at 3), while if c_4 is not divisible by 3, the curve has either good or multiplicative reduction. In both cases, the local conductor can be computed readily, it being 3^0 for good reduction and 3^1 for multiplicative. To compute the conductor in the remaining cases, let $\tilde{c}_4$ be the the least nonnegative residue of $(c_4/9)$ modulo 3, and $\tilde{c}_6$ be the the least nonnegative residue of $(c_6/27)$ modulo 9. Table 1 then gives us the exponent of the local conductor. Here $e = 5$ if $3^4 \mid c_4$ and $e = 4$ if $3^3 \mid\mid c_4$ (note that we must have $3^5 \mid\mid c_6$ in this case for the curve to be twist-minimal).

Table 1. Local Conductors at 3

$\tilde{c}_4 \backslash \tilde{c}_6$	0	1	2	3	4	5	6	7	8
0	e	3	3	5	2	2	5	3	3
1	2	3	4	3	4	4	3	4	3
2	2	3	2	3	3	3	3	2	3

For $p = 2$, the minimality test and conductor computation is much more complicated. We include the prime at infinity (twisting by -1) in the test for $p = 2$. By Connell's conditions, if $2^6 \mid c_4$ and $2^8 \mid c_6$, we see that E_2 is a curve with invariants $(c_4/4, c_6/8)$, and has a lesser discriminant. Also if $2^6 \mid c_4$ and $2^6 \mid\mid c_6$, then one of the twists $E_{\pm 2}$ (the sign depending on whether $c_6/8$ is 8 mod 32) has lesser discriminant. And finally if we have $2^4 \mid\mid c_4$ and $2^6 \mid\mid c_6$ and $2^{18} \mid (c_4^3 - c_6^2)$, then one of $E_{\pm 1}$ (depending on whether $c_6/64$ is 3 mod 4) is nonminimal (in the standard sense) at 2, and hence can be ignored. If none of these events happens, then the curve is twist-minimal at $p = 2$ and the infinite prime. We next describe how to compute the local conductor at $p = 2$ in terms of congruence conditions. If c_4 is odd, then the local conductor is 2^0 or 2^1, depending on whether 2 divides Δ. If c_4 is even, then it is divisible by 16. In this case, if c_6 is 8 mod 32, there is good reduction at 2, and again the local conductor is 2^0. So we are left to consider the cases of additive reduction where $2^4 \mid c_4$ and $2^5 \mid c_6$. Let $\tilde{c}_4$ be the the least nonnegative residue of

$(c_4/16)$ modulo 8, and $\tilde{c}_6$ be the the least nonnegative residue of $(c_6/32)$ modulo 8. Table 2 then gives the exponent of the local conductor at 2. In this, the dashed entries simply do not occur. For the entries marked by e, let $\tilde{c}_4$ be the the least nonnegative residue of $(c_4/16)$ modulo 16, and $\tilde{c}_6$ be the the least nonnegative residue of $(c_6/32)$ modulo 16. We then use the further Table 3. All the conductor computations are exercises with Tate's algorithm [9].

Table 2. Local Conductors at 2

$\tilde{c}_4 \backslash \tilde{c}_6$	0	1	2	3	4	5	6	7
1,5	6	4	e	3	6	4	e	3
2,6	8	3	6	4	7	3	6	4
3,7	5	2	7	2	5	2	7	4
4	6	2	-	4	3	2	-	4
0	6	2	-	4	2	2	-	4

Table 3. More of the Same

$\tilde{c}_4 \backslash \tilde{c}_6$	2	6	10	14
1	4	5	5	3
5	3	2	4	4
9	5	3	4	5
13	4	4	3	2

A curve which has minimal discriminant at $p = 2$ will be of minimal conductor at $p = 2$ unless $2^4 \mid\mid N$ or $2^6 \mid\mid N$; we can throw out the curve in the first case, since E_{-1} will be found in the search process (and it has lesser conductor). But in the latter case, we cannot immediately discard the curve, as E_2 will have conductor smaller by a factor of 2, but the discriminant rises by a factor of 64. So only if $|\Delta| \leq 10^{12}/64$ do we discard the curve; in the alternative case we replace the curve by E_2, so that we have the twist of minimal conductor. Finally, if we have $2^5 \mid\mid N$ (possibly after the above twisting by 2), or $2^7 \mid N$, we make the arbitrary decision to discard the curve if $c_6 < 0$, as we will also find E_{-1} in the search, which will have the same conductor and discriminant.

Using the above method, we can rid ourselves of all curves which are not minimal twists, and simultaneously compute the conductor. If $N > 10^{10}$, we simply ignore the curve; if $N > 10^8$ (and $N \leq 10^{10}$), we

check whether N is a strong pseudoprime for 2, 13, 23, and 1662803, this being sufficient to prove primality [6]. At this point, we have a list of curves which meet our size conditions on the discriminant, and which have the minimal conductor in a family of quadratic twists, and minimal discriminant at primes other than $p = 2$.

2.2 Isogenous Curves

The next step will be to get rid of isogenous curves. The process of finding all curves isogenous to a given one is described in [4]. This is a fairly fast process, as most curves will have no nontrivial isogenies. Amongst the isogenous curves, we then take the curve of largest fundamental volume, that is, minimal Faltings height (which is unique by [8], as our representative. Note that this curve might not have the minimal discriminant in the isogeny class. Our final set of curves is then: the set of elliptic curves E such that E has minimal height in its isogeny class, and has some isogenous curve F for which we have $c_4 \leq 1.44 \cdot 10^{12}$ and either $N \leq 10^{10}$ with $|\Delta|$ prime, or $N \leq 10^8$ with $|\Delta| \leq 10^{12}$ for either the curve F or F_2.

2.3 Future Extension of the Database

As stated above, we would desire to have all minimal twists which have conductor less than 10^8. There are three ways of enlarging the database. The first is extending the range on c_4 by using the algorithm of [5]. The second is to incorporate the data from the exhaustive methods of Cremona. The third is to find families in which we expect the conductor to be substantially less than the discriminant; for instance, curves with a rational point of order 5 often have some prime to the 5th power dividing the discriminant. In the same vein, curves with (say) a 5-isogeny are parametrised from $X_0(5)$, and in such a parametrised family we again expect a large difference between the conductor and discriminant. We could also extend the discriminant limit to (say) 10^{13} for certain $(c_4^\star, c_6^\star)$ pairs, especially those for which we know ahead of time that we will save significant powers of 2 and 3 in the conductor compared to the discriminant.

3 Data Computed for Each Curve

One object of interest for an elliptic curve is its algebraic rank. This is hard to compute; indeed, there is no known algorithm to do this, only ones which work conditionally. By the process given in [4], we can try

to determine the **analytic rank** of the curve, which is the degree of vanishing of its L-series at the central point. Of course, as there is no way to determine if a computed number is exactly zero, we can only give a good guess as to the analytic rank. The conjecture of Birch and Swinnerton-Dyer asserts that the algebraic rank and the analytic rank are equal, and that the first nonzero derivative of the L-function at the central point has arithmetic significance. For each curve in the database, we computed the suspected analytic rank and first nonzero derivative for both the curve itself and some of its quadratic twists.

Each curve in our database is the curve of minimal Faltings height in its isogeny class. A conjecture of Stevens [8] asserts that this curve should be the **optimal** curve for parametrisations from $X_1(N)$, in the sense that the parametrisations to the isogenous curves factor through the parametrisation to the strong curve (the existence of a modular parametrisation from $X_1(N)$ was proved in [1] following the methods initiated by Wiles [11]). It is sometimes the case that the optimal curve for parametrisations from $X_0(N)$ differs from the curve we find; in [10], a process is given to find the $X_0(N)$-optimal curve, assuming a technical condition, namely that the Manin constant of the optimal curve is 1 (this is similar to the Stevens conjecture). As many of the Frobenius traces were already computed for the analytic rank computation, these can be re-used at this stage. In a section below, we discuss the data obtained.

In the aforementioned paper [10], a process is given to compute the modular degree of an elliptic curve, again assuming that the Manin constant is 1. Compared to the computation of the analytic rank, which requires about the first $\sqrt{N}$ of the Frobenius traces, this method requires on the order of N of these (actually $\tilde{N}$, the symmetric-square conductor; see below). Thus for $N \geq 300000$ or so, it becomes rather time-consuming to compute the modular degree. We therefore compromised, computing the modular degree only if the symmetric-square conductor of the elliptic curve was sufficiently small (if we write $N = \prod_p p^{f_p}$ as a product of local conductors, then the symmetric-square conductor is simply $\tilde{N} = \prod_p p^{\lceil f_p/2 \rceil}$, except possibly when $f_2 = 8$, when the local symmetric-square conductor at 2 might be either 2^3 or 2^4; see [10] for details). We also computed the modular degree in some other interesting cases, for instance, when the rank is large.

4 Differing Optimal Curves

Here we discuss the question of differing optimal curves for parametrisations from $X_0(N)$ and $X_1(N)$. Note that we do not compute the actual optimal curve for the latter, relying instead on the Stevens conjecture, and compute the optimal curve for $X_0(N)$ only under the assumption that the Manin constant is 1. But the results are still interesting.

There appear to be three major cases when the optimal curves differ by a 2-isogeny. One of these, the so-called Setzer-Neumann curves, was considered in [7]. These curves are parametrised by $c_4 = u^2 + 48$ and $c_6 = -u\left(u^2 + 72\right)$, with the discriminant $u^2 + 64$ being a prime and u being taken to be congruent to 1 mod 4. The second family corresponds to taking $c_4 = 16\left(u^2 + 3\right)$ and $c_6 = -32u\left(2u^2 + 9\right)$ with u again being 1 mod 4 and $p = u^2 + 4$ being prime. Here the conductor is $4p$ and the discriminant is $16p$; the differing optimal curves property appears to be preserved upon twisting by -1.

The third family we found is obtained by taking $c_4 = p\left(p + 16\right) + 16$ and $c_6 = \left(p + 8\right)\left(p^2 + 16p - 8\right)$ of discriminant $p\left(p + 16\right)$ with both p and $p + 16$ primes congruent to 3 mod 4. A similar thing occurs if p and $p + 16$ are more generally powers of primes, but at least one of the two must be a power of a prime which is congruent to 3 mod 4 (i.e. $p = 11$ or $p = 2401$ works, but $p = 625$ does not). If p is congruent to 1 mod 4, then the sign of c_6 must be switched. Finally, p can be taken to be negative, for instance $p = -5$. Note that $p = 9$ leads to 15A, in which the optimal curves differ by a 4-isogeny; also, 17A might be thrown into consideration here with $p = 1$, which also has the optimal curves differing by a 4-isogeny.

With these considerations, there are but a couple of outstanding cases of optimal curves differing by a 2-isogeny (though proofs of this classification are lacking), those being the isogeny classes 24A/48A, 40A/80A, 32A/64A, and 128B/128D, though this last case can be seen as the $p = 8$ case of the second family. Ignoring the 5-isogeny example of 11A as being spurious, this leaves just the occasions of the optimal curves differing by a 3-isogeny. Here, all known examples are parametrised by

$$c_4 = (n + 3)\left(n^3 + 9n^2 + 27n + 3\right)$$

and

$$c_6 = -\left(n^6 + 18n^5 + 135n^4 + 504n^3 + 891n^2 + 486n - 27\right)$$

with the discriminant being $n\left(n^2 + 9n + 27\right)$. The n's for which the optimal curves differ are (experimentally) precisely those for which $n^2+9n+27$

is a prime power and n has no prime factors congruent to 1 mod 6; else the optimal curves are the same. We have no theoretical justification of this observation.

5 Data Obtained

This may seem strange for a comprehensive database project, but we do not dwell on large-scale phemonemon; indeed, the Brumer–McGuinness work is probably already sufficient in this manner, at least for prime conductor. As noted there, telling the difference between a small power of 10^8 (or whatever the upper limit of consideration may be) and a large power of its logarithm is rather hopeless—extending their data by a factor of 5/4 on the logarithmic scale does not help matters much. We mention that there are 11386955 isogeny classes of curves with prime conductor less than 10^{10} in our database (this should grow slightly when curves with $c_4 \geq 1.44 \cdot 10^{12}$ are added). Of these curves with prime conductor, of the ones we have processed, we have that 62.5% of the curves with even functional equation possess rank 0, compared to about 60% for Brumer–McGuinness. It is conjectured that asymptotically this percentage should be 100%. Similarly, 92.5% of the curves with odd functional equation have rank 1, slightly more than the previous results; there is no real reason to think that our numbers will change drastically upon extending the rank computation to all the prime conductor curves we have. The least conductor for a rank 5 curve we have found is 48012824 for $[0, 1, 0, -625, 6099]$, and for rank 6 we have $[0, 0, 1, -277, 4566]$ of conductor 7647224363. These respectively fall short to the best-known (to the authors) examples of $[0, 0, 1, -79, 342]$ of conductor 19047851 and $[0, 0, 1, -7077, 235516]$ of conductor 5258110041.

Instead of concentrating on large-scale behavior, we see our database as more of a tool to be used by other mathematicians. For instance, Neil Dummigan queried us concerning examples of strong Weil curves with rank 2 and a rational point of order 5 for which the conductor is not divisible by 5, and we were able to provide him with the example $[0, 1, 1, -840, 39800]$ of conductor 13881 (and modular degree 52000), among other examples which were beyond the range of Cremona's tables (which include $[1, 1, 1, -2365, 43251]$ of conductor 5302). Though we would likely be better able to answer the question after extending our database with parametrisations from $X_0(5)$, the efficacy of our database was evinced. As another example, the second author has conjectured in [10] that 2^r divides the modular degree for any curve (where r is the rank),

and perhaps higher powers of 2 should divide the modular degree when the conductor is composite, due to factorisation through Atkin–Lehner involutions. For many large-rank curves in the Brumer–McGuinness database, we verified this. With our extension to curves of composite conductor, we are able to give more evidence for this conjecture. Also, the third 2-isogeny family in the previous section was discovered after looking at our data, as was the parametrisation of the 3-isogeny family, and finally our analytic rank data concerning quadratic twists could be of use.

6 Acknowledgements

The authors would like to thank Neil Dummigan for the question mentioned in Section V, Noam Elkies for many useful observations, especially concerning Section IV, and Blair Kelly III and Wayne Whitney for providing computing power.

References

1. C. Breuil, B. Conrad, F. Diamond, and R. Taylor, *On the modularity of elliptic curves over* $\mathbf{Q}$*: Wild 3-adic exercises.* J. Amer. Math. Soc. **14** (2001), 843–939.
2. 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.
3. I. Connell, Lecture Notes from class at McGill University, 1991.
4. J. Cremona, *Algorithms for modular elliptic curves.* Cambridge University Press, Cambridge, 1992. Second edition 1997.
5. N. Elkies, *Rational points near curves and small nonzero* $|x^3 - y^2|$ *via lattice reduction.* In *Algorithmic number theory* (Leiden 2000), 33–63, Lecture Notes in Comput. Sci., 1838, Springer, Berlin, 2000.
6. G. Jaeschke, *On strong pseudoprimes to several bases.* Math. Comp. **61** (1993), no. 204, 915–926.
7. 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.
8. G. Stevens, *Stickelberger elements and modular parametrizations of elliptic curves.* Invent. Math. **98** (1989), no. 1, 75–106.
9. 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.
10. M. Watkins, *Computing the modular degree of an elliptic curve,* preprint, 2001.
11. A. Wiles, *Modular elliptic curves and Fermat's last theorem.* Ann. of Math. (2) **141** (1995), no. 3, 443–551.