

4-5-2006

Some Remarks on Heegner Point Computations

Mark Watkins
University of Bristol

Follow this and additional works at: <https://surface.syr.edu/mat>

 Part of the [Mathematics Commons](#)

Recommended Citation

Watkins, Mark, "Some Remarks on Heegner Point Computations" (2006). *Mathematics Faculty Scholarship*. 106.
<https://surface.syr.edu/mat/106>

This Article is brought to you for free and open access by the Mathematics at SURFACE. It has been accepted for inclusion in Mathematics Faculty Scholarship by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

SOME REMARKS ON HEEGNER POINT COMPUTATIONS

by

Mark Watkins

Abstract. — We give an overview of the theory of Heegner points for elliptic curves, and then describe various new ideas that can be used in the computation of rational points on rank 1 elliptic curves. In particular, we discuss the idea of Cremona (following Silverman) regarding recovery a rational point via knowledge of its height, the idea of Delaunay regarding the use of Atkin-Lehner involutions in the selection of auxiliary parameters, and the idea of Elkies regarding descent and lattice reduction that can result in a large reduction in the needed amount of real-number precision used in the computation.

1. Introduction

We make some remarks concerning Heegner point computations. One of our goals shall be to give an algorithm (perhaps conditional on various conjectures) to find a non-torsion rational point on a given rank 1 elliptic curve. Much of this is taken from a section in Henri Cohen’s latest book [9], and owes a great debt to Christophe Delaunay. The ideas in the section about lattice reduction are largely due to Noam Elkies. We do not delve deeply into the theory of Heegner points, but simply give references where appropriate; the recent MSRI book “Heegner points and Rankin L -series” [11] contains many good articles which consider Heegner points and their generalisations from the standpoint of representation theory.

The author thanks the Institut Henri Poincaré for its hospitality and the Centre National de la Recherche Scientifique for financial support. The algorithms described here have been implemented in the Magma computer algebra system [4]; the author thanks the Magma computer algebra group at the University of Sydney for their hospitality and financial support. The author was also partially funded by an NSF VIGRE Postdoctoral Fellowship during part of this work.

2000 Mathematics Subject Classification. — 11G05, 11G40, 14G05.

Key words and phrases. — elliptic curves, Heegner points, descent, lattice reduction.

2. Definitions and Outline of Theory

Let τ be a quadratic surd in the upper-half-plane \mathbf{H} . Let $f_\tau = (A, B, C)$ be the associated integral primitive positive-definite binary quadratic form, so that $A\tau^2 + B\tau + C = 0$ with $A > 0$ and $\gcd(A, B, C) = 1$. The discriminant $\Delta(\tau)$ is $\Delta(f_\tau) = B^2 - 4AC$, which is negative. For simplicity we take $\Delta(\tau)$ to be fundamental, though much of the theory can be made to work when it is not. However, consideration of positive discriminant does not follow in the same manner.

Definition 2.1. — A Heegner point of level N and discriminant D is a quadratic surd in the upper-half-plane with $\Delta(\tau) = D = \Delta(N\tau)$. We let \mathcal{H}_N^D be the set of Heegner points of level N and discriminant D .

Proposition 2.2. — Let $\tau \in \mathbf{H}$ be a quadratic surd with discriminant D and $f_\tau = (A, B, C)$. Then $\tau \in \mathcal{H}_N^D$ iff $N|A$ and $\gcd(A/N, B, CN) = 1$.

Proof. — Note that $\tau = \frac{-B+\sqrt{D}}{2A}$ and $N\tau = \frac{-NB+N\sqrt{D}}{2A}$. For $\Delta(\tau) = \Delta(N\tau)$ we need $N\tau = \frac{-B'+\sqrt{D}}{2A'}$ and by equating imaginary and real parts we get $A = NA'$ and $B = B'$, so that $N|A$. Also note that $(A/N)(N\tau)^2 + B(N\tau) + (CN) = 0$, from which we get the rest of the lemma. \square

Note that \mathcal{H}_N^D will be empty unless $D = B^2 - 4N(A/N)C$ is a square modulo $4N$.

Lemma 2.3. — The set \mathcal{H}_N^D is closed under $\Gamma_0(N)$ -action.

Proof. — If $\gamma \in SL_2(\mathbf{Z}) \supseteq \Gamma_0(N)$ then $\Delta(\gamma(\tau)) = \Delta(\tau)$ since the discriminant is fixed. A computation shows that $\gamma \in \Gamma_0(N)$ and $\tau \in \mathcal{H}_N^D$ imply $\gamma(\tau) \in \mathcal{H}_N^D$. \square

Lemma 2.4. — The set \mathcal{H}_N^D is closed under the W_N -action that sends $\tau \rightarrow -1/N\tau$.

Proof. — Follows from the above proposition since $f_{(-1/N\tau)} = (CN, -B, A/N)$. \square

Definition 2.5. — We let $SS(D, N)$ be the set of square roots mod $2N$ of D mod $4N$.

Theorem 2.6. — The sets $\mathcal{H}_N^D/\Gamma_0(N)$ and $SS(D, N) \times Cl(\mathbf{Q}(\sqrt{D}))$ are in bijection.

Proof. — This can be shown by chasing definitions. Essentially, $[\tau] \in \mathcal{H}_N^D/\Gamma_0(N)$ gets mapped to $(B \bmod 2N) \times [\mathbf{Z} + \tau\mathbf{Z}]$ where $f_\tau = (A, B, C)$, and in the other direction, when we are given $\beta \times l \in SS(D, N) \times Cl(\mathbf{Q}(\sqrt{D}))$ we take $(A, B, C) \in l$ with $N|A$ and $B \equiv \beta \pmod{2N}$, and then $\tau = \frac{-B+\sqrt{D}}{2A}$. \square

From now on we let E be a global minimal model of a rational elliptic curve of conductor N , and take D to be a negative fundamental discriminant such that D is a square mod $4N$. We let \mathbf{H}^* be the union of \mathbf{H} with the rationals and $i\infty$. We let $\mathcal{P}(z)$ be the function that sends $z \in \mathbf{C}/\Lambda$ to the point $(\wp(z), \wp'(z))$ on E .

Theorem 2.7. — There is a surjective map $\hat{\phi} : X_0(N) \rightarrow E$ (the modular parametrization) where $X_0(N) = \mathbf{H}^*/\Gamma_0(N)$ and E can be viewed as \mathbf{C}/Λ for some lattice Λ . This map can be defined over the rationals.

Proof. — This is due to Wiles and others [29, 28, 12, 10, 6]. We let ϕ be the associated map from $\mathbf{H}^*/\Gamma_0(N)$ to \mathbf{C}/Λ . Explicitly, we have that $\tau \in \mathbf{H}^*$ gets mapped to the complex point $\phi(\tau) = 2\pi i \int_{i_\infty}^\tau \psi_E = \sum_n (a_n/n) e^{2\pi i n \tau}$, where ψ_E is the modular form of weight 2 and level N associated to E . The lattice Λ is generated by the real and imaginary periods,⁽¹⁾ which we denote by Ω_{re} and Ω_{im} . We assume that the Manin constant is 1, which is conjectured always to be the case for curves of positive rank (see [27] and [25]). \square

Theorem 2.8. — *Let $\tau = \beta \times l \in \mathcal{H}_N^D$. Then $\mathcal{P}(\phi(\tau))$ has its coordinates in the Hilbert class field of $\mathbf{Q}(\sqrt{D})$. Also we have*

1. $\overline{\phi(\beta \times l)} = \phi(-\beta \times l^{-1})$, in \mathbf{C}/Λ .
2. $\phi(W_N(\beta \times l)) = \phi(-\beta \times l n^{-1})$ in \mathbf{C}/Λ where $n = [N\mathbf{Z} + \frac{\beta + \sqrt{D}}{2}\mathbf{Z}]$,
3. $\mathcal{P}(\phi(\beta \times l))^{\text{Artin}(m)} = \mathcal{P}(\phi(\beta \times l m^{-1}))$ for all $m \in \text{Cl}(\mathbf{Q}(\sqrt{D}))$.

Proof. — This is the theorem of complex multiplication of Shimura [22, 21]. We outline the proof of the first statement, for which we work via the modular j -function. We have that $j(\tau)$ is in the Hilbert class field H (see [23, II, 4.3]) and similarly with $j(N\tau)$. Thus we get that $X_0(N)$ over H contains the moduli point corresponding to the isogeny between curves with these j -invariants. Since the modular parametrisation map $\hat{\phi}$ can be defined over the rationals, the image of the moduli point under $\hat{\phi}$ has its coordinates in the Hilbert class field. \square

Note that $\mathcal{P}(\overline{\phi(\tau)}) = \overline{\mathcal{P}(\phi(\tau))}$, so that there is no danger of confusing complex conjugation in \mathbf{C}/Λ with complex conjugation of the coordinates of the point on E . Using the third fact of Theorem 2.8, we can take the trace of $\mathcal{P}(\phi(\tau))$ and get a point that has coordinates in $\mathbf{Q}(\sqrt{D})$. Indeed, writing H for the Hilbert class field and K for the imaginary quadratic field $\mathbf{Q}(\sqrt{D})$ we get that

$$\begin{aligned} P = \text{Trace}_{H/K}(\mathcal{P}(\phi(\tau))) &= \sum_{\sigma \in \text{Gal}(H/K)} \mathcal{P}(\phi(\tau))^\sigma = \sum_{m \in \text{Cl}(K)} \mathcal{P}(\phi(\beta \times l))^{\text{Artin}(m)} \\ &= \sum_{m \in \text{Cl}(K)} \mathcal{P}(\phi(\beta \times l m^{-1})) = \sum_{m \in \text{Cl}(K)} \mathcal{P}(\phi(\beta \times m)) \end{aligned}$$

has coordinates in $\mathbf{Q}(\sqrt{D})$. When E has odd functional equation, we can use the first two facts of Theorem 2.8 to show that $P = \overline{P}$, so that P has coordinates in \mathbf{Q} . In this case we have $\psi_E = \psi_E \circ W_N$ which implies $\phi = \phi \circ W_N$, so that in \mathbf{C}/Λ we have

$$\overline{\phi(\beta \times m)} = \phi(W_N(\beta \times m)) = \phi(-\beta \times m n^{-1}) = \phi(\beta \times m^{-1} n),$$

which gives us that

$$\overline{P} = \sum_{m \in \text{Cl}(K)} \mathcal{P}(\overline{\phi(\beta \times m)}) = \sum_{m \in \text{Cl}(K)} \mathcal{P}(\phi(\beta \times m^{-1} n)) = \sum_{m \in \text{Cl}(K)} \mathcal{P}(\phi(\beta \times m)) = P.$$

⁽¹⁾Our convention is that the imaginary period is purely imaginary when the discriminant of E is positive, and in the negative discriminant case the real part of the imaginary period is $\Omega_{\text{re}}/2$. The fundamental volume Ω_{vol} is the area of the period parallelogram.

We can rewrite some of this by introducing some new notation.

Definition 2.9. — We write $\mathcal{H}_N^D(\beta)$ for subset of $\tau \in \mathcal{H}_N^D$ such that the associated form $f_\tau = (A, B, C)$ has $B \equiv \beta \pmod{2N}$. We write $\hat{\mathcal{H}}_N^D = \mathcal{H}_N^D/\Gamma_0(N)$, and noting that $\Gamma_0(N)$ acts on $\mathcal{H}_N^D(\beta)$, we write $\hat{\mathcal{H}}_N^D(\beta) = \mathcal{H}_N^D(\beta)/\Gamma_0(N)$.

Since $\hat{\mathcal{H}}_N^D(\beta)$ is in 1-1 correspondence with $\mathcal{Cl}(\mathbf{Q}(\sqrt{D}))$, we get that

$$P = \sum_{m \in \mathcal{Cl}(\mathbf{Q}(\sqrt{D}))} \mathcal{P}(\phi(\beta \times m)) = \sum_{\tau \in \hat{\mathcal{H}}_N^D(\beta)} \mathcal{P}(\phi(\tau)).$$

3. The Gross-Zagier theorem and an algorithm

We now have a plan of how to find a non-torsion point on a curve of analytic rank 1. We select an auxiliary negative fundamental discriminant D such that D is a square modulo $4N$, choose $\beta \in SS(D, N)$, find τ -representatives for $\hat{\mathcal{H}}_N^D(\beta)$, compute $\phi(\tau)$ for each, sum these in \mathbf{C}/Λ , map the resulting point to E via the Weierstrass parametrization, and try to recognize the result as a rational point. One problem is that we might get a torsion point. Another problem is that we won't necessarily get a generator, and thus the point might have inflated height, which would increase our requirements on real-number precision. The Gross-Zagier Theorem tells us what height to expect, and combined with the Birch–Swinnerton-Dyer Conjecture, we get a prediction of what height a generator should have. Our heights will be the “larger” ones, and are thus twice those chosen by some authors.

Theorem 3.1. — *Suppose $D < -3$ is a fundamental discriminant with D a square modulo $4N$ and $\gcd(D, 2N) = 1$. Then*

$$h(P) = \frac{\sqrt{|D|}}{4\Omega_{\text{vol}}} L'(E, 1) L(E_D, 1) \times 2^{\omega(\gcd(D, N))} \left(\frac{w(D)}{2} \right)^2.$$

Proof. — This is due to Gross and Zagier [15]. Here E_D is the quadratic twist of E by D , while $w(D)$ is the number of units in $\mathbf{Q}(\sqrt{D})$ and $\omega(n)$ is the number of distinct prime factors of n . \square

Calculations of Gross and Hayashi [16] indicate that this height formula is likely to be true for all negative fundamental discriminants D that are square mod $4N$.

We now write $P = lG + T$ where G is a generator⁽²⁾ and T is a torsion point, so that $h(P) = l^2 h(G)$. Then we replace $L'(E, 1)$ through use of the Birch–Swinnerton-Dyer conjecture [5] to get the following:⁽³⁾

⁽²⁾Note that we will actually get $\sqrt{\#\text{III}}$ times a generator from our method, since we cannot disassociate III from the regulator in the Birch–Swinnerton-Dyer formula.

⁽³⁾We use the convention that the Tamagawa number at infinity is equal to the number of connected components of E over \mathbf{R} — thus it is 1 for curves with negative discriminant and 2 for curves with positive discriminant.

Conjecture 3.2. — *With notations as above we have that*

$$l^2 = \frac{\Omega_{\text{re}}}{4\Omega_{\text{vol}}} \left(\prod_{p|N_\infty} c_p \cdot \#\text{III} \right) \frac{\sqrt{|D|}}{\#E(\mathbf{Q})_{\text{tors}}^2} L(E_D, 1) \cdot \left(\frac{w(D)}{2} \right)^2 2^{\omega(\gcd(D, N))}.$$

In particular, we note that we should use a quadratic twist E_D that has rank zero, so that $L(E_D, 1)$ does not vanish. The existence of such a twist is proven in [7]. Thus we have the following algorithm, which we shall work on improving.

Algorithm 3.3. — *Given a rational elliptic curve E of conductor N of analytic rank 1, find a non-torsion rational point.*

1. Compute $L'(E, 1)$ and find a fundamental discriminant $D < 0$ with D a square modulo $4N$ and $L(E_D, 1) \neq 0$, so that the index l is nonzero.
2. Choose $\beta \in SS(D, N)$ and compute (to sufficient precision) the complex number

$$z = \sum_{\tau \in \hat{\mathcal{H}}_N^D(\beta)} \phi(\tau) = \sum_{\tau \in \hat{\mathcal{H}}_N^D} \sum_{n=1}^{\infty} \frac{a_n}{n} e^{2\pi i n \tau}.$$

3. Let m be the gcd of l and the exponent of the torsion group of E . If the discriminant of E is positive, check if $\mathcal{P}(\dot{z})$ is close to a rational point on E for $u = 1, \dots, lm$ for both

$$\dot{z} = (m\text{Re}(z) + u\Omega_{\text{re}})/ml \quad \text{and} \quad \dot{z} = (m\text{Re}(z) + u\Omega_{\text{re}})/ml + \Omega_{\text{im}}/2.$$

If the discriminant of E is negative, let $o = \text{Im}(z)/\text{Im}(\Omega_{\text{im}})$ and check $\mathcal{P}(\dot{z})$ for $\dot{z} = (m\text{Re}(z) + u\Omega_{\text{re}})/ml + o\Omega_{\text{re}}/2$ over the same u -range.

One can compute the index l in parallel with the $\phi(\tau)$, since both involve computing the a_n of the elliptic curve E . However, this can cause problems if the index turns out to be zero (that is, if E_D has positive rank).

3.1. Step 2 of the Algorithm. — We now discuss how to do the second step efficiently. First note that we can sometimes pair $\phi(\tau)$ with its complex conjugate; recalling that $\phi = \phi \circ W_N$, by Theorem 2.8 in \mathbf{C}/Λ we have

$$(1) \quad \overline{\phi(\beta \times l)} = \phi(-\beta \times l^{-1}) = \phi(W_N(-\beta \times l^{-1})) = \phi(\beta \times (ln)^{-1}).$$

For $f = (A, B, C) \in \hat{\mathcal{H}}_N^D$ we write $\bar{f} = (A/N, -B, CN)$, so that when $g \sim \bar{f}$ in the class group we have $\phi(f) = \overline{\phi(g)}$ and thus $\phi(f) + \phi(g) = 2 \text{Re} \phi(f)$ in \mathbf{C}/Λ . We refer to this as **pairing** the forms.

For $\sum_n (a_n/n) e^{2\pi i n \tau}$ to converge rapidly, we wish for the imaginary parts of our representative τ 's to be large. It turns out the best we can do is essentially have the smallest imaginary part be about $1/N$ in size. We can achieve this via a trick of Delaunay, which introduces more Atkin-Lehner involutions.

Definition 3.4. — Let $Q|N$ with $\gcd(Q, N/Q) = 1$, and let $u, v \in \mathbf{Z}$ be such that $uQ^2 - vN = Q$. The Atkin-Lehner involution W_Q sends τ to $\frac{uQ\tau + v}{N\tau + Q}$.

This defines W_Q up to transformations by elements in $\Gamma_0(N)$. One can check that $W_Q(W_Q(\tau))$ is in the $\Gamma_0(N)$ -orbit of τ , and that the W_Q form an elementary abelian 2-group W of order $2^{\omega(N)}$. The important fact about the W_Q 's that we shall use is that $\psi_E = \pm\psi_E \circ W_Q$, so that $\phi(\tau) = \pm\phi(W_Q(\tau)) + \phi(W_Q(i\infty))$. The sign can be computed as $\epsilon_Q = \prod_{p|Q} \epsilon_p$ where ϵ_p is the local root number of E at p . Delaunay's idea is to maximise the imaginary part of τ over $\Gamma_0(N)$ and W rather than just $\Gamma_0(N)$; the difficulty is that the action of W_Q need not preserve β . However, we still have that

$$P = \sum_{\tau \in \mathcal{H}_N^R(\beta)} \phi(\tau) = \sum_{\tau \in \mathcal{H}_N^R(\beta)} \epsilon_Q \phi(W_Q(\tau)) + (\text{torsion point}).$$

For the analogue of the second part of Theorem 2.8, we need to consider what happens with to β . We define β_Q as follows. We make β_Q and β have opposite signs mod p^k for prime powers p^k with $p^k \parallel Q$ and $\gcd(p, D) = 1$, and else $\beta = \beta_Q$. In particular, we have that $Q = \gcd(\beta - \beta_Q, N)$ when N is odd. The desired analogue is now that

$$\phi(W_Q(\beta \times l)) = \epsilon_Q \phi(\beta_Q \times lq^{-1}) + \phi(W_Q(i\infty)) \quad \text{with} \quad q = \left[Q\mathbf{Z} + \frac{-\beta_Q + \sqrt{D}}{2}\mathbf{Z} \right].$$

The primes p which divide D are different since there is only one square root of D mod p ; thus β is preserved upon application of W_Q for Q that are products of such primes. For such Q , we can note the following with respect to complex conjugation. Suppose we have that $m \sim (ln)^{-1}$ so that by (1) we have $\overline{\phi(\beta \times l)} = \phi(\beta \times m)$. Then, using the fact that $q^{-1} = q$ in this case, in \mathbf{C}/Λ we have, up to torsion, that

$$\begin{aligned} \overline{\phi(W_Q(\beta \times l))} &\doteq \overline{\epsilon_Q \phi(\beta_Q \times lq^{-1})} = \epsilon_Q \phi(\beta_Q \times (lq^{-1}n)^{-1}) = \epsilon_Q \phi(\beta_Q \times (lqn)^{-1}) \doteq \\ &\doteq \phi(W_Q(\beta \times (ln)^{-1})) = \phi(W_Q(\beta \times m)). \end{aligned}$$

So we see that $(\beta \times l)$ can be paired iff $W_Q(\beta \times l)$ can be paired.

We now give the algorithm for finding good τ -representatives. The idea is to run over all forms (aN, b, c) of discriminant D with a small, mapping these via the appropriate Atkin-Lehner involution(s) to forms with fixed square root β , and doing this until the images cover the class group. Of course, the conjugation action is also considered.

Subalgorithm 3.5. — *Given D, N , find good τ -representatives.*

1. Choose $\beta \in SS(D, N)$. Set $U = \emptyset$ and $R = \emptyset$.
2. While $\#R \neq \#Cl((Q\sqrt{D}))$ do:
 3. Loop over a from 1 to infinity and $b \in SS(D, N)$ [lift b from $\mathbf{Z}/2N$ to \mathbf{Z}]:
 4. Loop over all solutions s of $Ns^2 + bs + (b^2 - D)/4N \equiv 0$ modulo a :
 5. Let $f = (aN, b + 2Ns, ((b + 2Ns)^2 - D)/4aN)$.
 6. Loop over all positive divisors d of $\gcd(D, N)$ [which is squarefree]:
 7. Let $g = W_Q(f)/Q$ where Q is d times the product of the $p^k \parallel N$ with $b \not\equiv \beta \pmod{p^k}$, so that $g \in \mathcal{H}_N^D(\beta)$.
 8. If the reductions of g and \bar{g} are both not in R then append them to R , and append f to U with weight ϵ_Q when $g \sim \bar{g}$ and with weight $2\epsilon_Q$ when $g \not\sim \bar{g}$.

With this subalgorithm, we get that $z = \sum_{f \in U} \text{weight}(f)\phi(\tau_f)$ in Step 2 of the main algorithm. We expect the maximal a to be of size $\#\text{Cl}(\mathbf{Q}(\sqrt{D}))/2\#W$. This subalgorithm makes “parameter selection” fast compared to the computation of the $\phi(\tau)$.

3.2. Step 3 of the Algorithm. — We now turn to the last step of our main algorithm, reconstructing a rational point on an elliptic curve from a real approximation. The most naïve method for this is simply to try to recognise the x -coordinate as a rational number. If our height calculation tells us to expect a point whose x -coordinate has a numerator and denominator of about H digits, the use of continued fractions will recognise it if we do all computations to about twice the precision, or $2H$ digits. We can note that by using a degree- n map to \mathbf{P}^{n-1} and n -dimensional lattice reduction, this can be reduced to $nH/(n-1)$ digits for every $n \geq 3$ — we will discuss a similar idea later when we consider combining descent with our Heegner point computations. But in this case we can do better; we are able to recognise our rational point with only H digits of precision due to a trick of Cremona, coming from an idea in a paper of Silverman [24]. The idea is that we know the canonical height of our desired point, and this height decomposes into local heights; we have

$$h(P) = h_\infty(P) + \sum_{p|N} h_p(P) + \log \text{denominator}(x(P)).$$

The height at infinity $h_\infty(P)$ can be approximated from a real-number approximation to P , and there are finitely many possibilities for each local height $h_p(P)$ depending the reduction type of E at p . We compute the various local heights to H digits of precision, and then can determine the denominator of $x(P)$ from this, our task being eased from the fact that it is square. Then from our real-approximation⁽⁴⁾ of P we can recover the x -coordinate, and from this we get P . Note that we need to compute $L'(E, 1)$ to a precision of H digits, but this takes only about $\sqrt{N}(\log H)^{O(1)}$ time. In practise, there can be many choices for the sums of local heights, and if additionally the index is large, then this step can be quite time-consuming. This can be curtailed a bit by doing the calculations for the square root of the denominator of the x -coordinate to only about $H/2$ digits, and then not bothering with the elliptic exponential step unless the result is sufficiently close to an integer.

3.3. Example. — We now give a complete example. Other explicit descriptions of computations with Heegner points appear in [3, 26, 17]. We take the curve given by $[1, -1, 0, -751055859, -7922219731979]$ for which the Heegner point has height $139.1747+$. We select $D = -932$, for which the class number is 12 and the index l is 4. We have $N = 11682$ and choose $\beta = 214$. Our first form is $(11682, 214, 1)$ to which we apply $W_1 = \text{id}$. The reduction of this is $(1, 0, 233)$, and it pairs with

⁽⁴⁾Elkies tells us that, given the height to precision H , the techniques of [14] (see Theorem 4 in particular) can reduce the needed precision of the real Heegner approximation to $o(H)$ as $H \rightarrow \infty$. The idea is that for a fixed C the equation $h_\infty((x, y, z)) + 2 \log z = C$ defines a *transcendental arc*, and thus the use of a sufficiently high degree Veronese embedding will reduce the needed precision substantially. This method in its entirety might not be that practical, though the use of height information in conjunction with the geometry of the curve should allow a useful reduction in precision.

itself under complex conjugation. Since we have $\gcd(D, N) = 2$, we can use W_2 without changing β ; we get the form (206717861394, 70769770, 6057) which reduces to the self-paired form (2, 2, 117). Our next form is (11682, 2338, 117) to which we apply W_{11} to get (122225810454, 230158978, 108351) which reduces to (11, 6, 22) and pairs with (11, -6, 22). Applying W_{22} gives a form which reduces to (11, -6, 22), so we ignore it. Next we have (11682, 2810, 169) to which we apply W_9 , getting a form that reduces to (9, 2, 26) and pairs with (9, -2, 26). Applying W_{18} gives a form that reduces to (13, -2, 18) and pairs with (13, 2, 18). Then we have (11682, 4934, 521) to which we apply W_{99} , getting a form that reduces to (3, -2, 78) and pairs with (3, 2, 78). And finally applying W_{198} we get a form that reduces to (6, -2, 39) and pairs with (6, 2, 39), and so we have all of our τ -representatives. We note that W_{11} , W_9 , and W_{18} switch the sign of the modular form, and thus the obtained forms get a weighting of -2 . The self-paired forms get a weighting of $+1$, and the other two forms get a weighting of $+2$. For the non-self-paired forms we must remember to take the real part of the computed $\phi(\tau)$ when we double it.⁽⁵⁾ The pairing turns is rather simple in this example, but need not be so perspicacious with respect to the class group. Note that we use only four distinct forms for our computations.

We need about 60 digits of precision if we use the Cremona-Silverman method to reconstruct the rational point, which means we must compute about 20000 terms of the L -series. The curve E has negative discriminant and no rational torsion points. We compute a real-approximation to the Heegner point in \mathbf{C}/Λ to be

$$z = 0.00680702983101357730368201485198918786991251635619740952608094.$$

We have $o = \text{Im}(z)/\text{Im}(\Omega_{\text{im}}) = 0$, and with $l = 4$ and $u = 2$ we get that

$$\dot{z} = 0.00891152819280235244790996808333469812474933020620405901507952,$$

to which we apply the Cremona-Silverman method of recovery. The curve E is annoying for this method, in that we have many possibilities for $h_p(P)$. The height of the Heegner point is given by

$$h(P) = 139.174739524758127811521877478222781093487974225206369462318,$$

and the height at infinity⁽⁶⁾ is given by

$$h_\infty(P) = 2.10306651755149369196435189022120441716979687181328497567075.$$

The reduction type at 2 is I_{25} , at 3 it is I_{13}^* , at 11 it is I_1 , and at 59 it is I_3 . Thus we have $13 \times 3 \times 1 \times 2$ choices for the local heights. It turns out that we have⁽⁷⁾ $h_p(P) = \frac{1}{6}v_p(\Delta) \log p$ for $p = 2, 11, 59$, while $h_3(P) = (13/6) \log 3$. The denominator of the x -coordinate is $12337088946900997614694947283^2$, and the numerator is

$$5908330434812036124963415912002702659341205917464938175508715.$$

⁽⁵⁾The self-paired forms f have $\phi(f) = \overline{\phi(f)}$ in \mathbf{C}/Λ but not necessarily in \mathbf{C} — the imaginary part cannot be ignored when the discriminant of E is negative and lm is odd.

⁽⁶⁾Note that Silverman [24] uses a different normalisation of height, and his choice of the parameter z when he computes the height at infinity corresponds to $\dot{z}/\Omega_{\text{re}}$ for us. Also, his method is only linearly convergent, while that given in [8, §7.5.7] is quadratically convergent.

⁽⁷⁾This follows *a posteriori* since P is nonsingular modulo these primes of multiplicative reduction.

3.4. Variants. — Next we mention a variant which, for the congruent number curve, has been investigated in depth by Elkies [13]. Here we fix a rank zero curve, say the curve $E : y^2 = x^3 - x$ of conductor 32, and try to find points on rank 1 quadratic twists E_D with $D < 0$. It can be shown that E_D will have odd functional equation for $|D| \equiv 5, 6, 7 \pmod{8}$. There is not necessarily a Gross-Zagier theorem in all these cases, and some involve mock Heegner points instead of Heegner points. However, we still have the prediction that $h(P) = \alpha_D L(E, 1) L'(E_D, 1)$ for some $\alpha_D > 0$. Elkies computes a point P in \mathbf{C}/Λ via a method similar to the above — however, he generally⁽⁸⁾ only attempts to determine if it is non-torsion, and thus need not worry as much about precision. There are about $\#\text{Cl}(\mathbf{Q}(\sqrt{D})) \approx \sqrt{|D|}$ conjugates of τ for which $\phi(\tau)$ needs to be computed; since we have an action of $\Gamma_0(32)$, computing each $\phi(\tau)$ takes essentially constant time, so we get an algorithm that takes about time $|D|^{1/2}$ to determine whether the computed point is non-torsion. Note that we don't obtain $L'(E_D, 1)$, which takes about $|D|$ time to compute, but only whether it is nonzero. MacLeod [18] investigated a similar family of quadratic twists, those of a curve of conductor 128. The relevant curves are $y^2 = (x + p)(x^2 + p^2)$ with $p \equiv 7 \pmod{8}$; with $p = 3167$ the height is 1022.64+. Some additional papers that deal with the theory and constructions in this case are those of Birch and Monsky [1, 2, 19, 20].

4. Combination with descent

To find Heegner points of large height, say 500 or more, it is usually best first to do a descent on the elliptic curve, as this will tend to reduce the size of the rational point by a significant factor.⁽⁹⁾ Upon doing a 2-descent, we need only $H/3$ digits of precision if we represent the covering curve as an intersection of quadrics in \mathbf{P}^3 and use 4-dimensional lattice reduction, and if we do a 4-descent we need only $H/12$ digits. We first explain how these lattice reduction methods work, and then show how to use them in our application. It might also be prudent to point out that if E has nontrivial rational isogenies, then one should work with the isogenous curve for which the height of the generator will be the smallest.⁽¹⁰⁾

4.1. Lattice Reduction. — Most of the theory here is due to Elkies [14]. We first describe a p -adic method — this is not immediately relevant to us as we do not know how to approximate the Heegner point in such a manner, but it helps to understand the idea. Let $F(W, X, Y) = 0$ be a curve in \mathbf{P}^2 . We wish to find rational points on F . Let $(1 : x_s : y_s)$ be a (nonsingular) point modulo some prime p , and lift this to a solution $(1 : x_0 : y_0)$ modulo p^2 . Then determine d such that any linear combination of $(1 : x_0 : y_0)$ and $(0 : p : dp)$ will be a solution mod p^2 (computing d

⁽⁸⁾Elkies also computes a generator of height 239.6+ for $y^2 = x^3 - 1063^2x$ in this manner.

⁽⁹⁾During the mid 1990s, Cremona and Siksek worked out a few examples using 2-descent.

⁽¹⁰⁾Because III might have different size for the various isogenous curves, we cannot always tell beforehand which curve(s) will have a generator of smallest height.

essentially involves taking a derivative). Then perform lattice reduction on the rows of the matrix

$$\begin{pmatrix} 1 & x_0 & y_0 \\ 0 & p & dp \\ 0 & 0 & p^2 \end{pmatrix}.$$

Finally search for global solutions to F by taking small linear combinations of the rows of the lattice-reduced matrix. If we choose p to be around B for some height bound B , upon looping through all local solutions modulo p we should find all global points whose coordinates are of size B ; in general we take p of size $B^{2/n}$ in projective n -space. This can be used, for instance, to search for points on a cubic model of an elliptic curve.⁽¹¹⁾

Over the real numbers the description is more complicated. Here we deal with the transformation matrix of the lattice reduction. If we wanted to do 2-dimensional reduction, that is, continued fractions, on a real number x_0 , we would perform lattice reduction on the rows of the matrix

$$M_2 = \begin{pmatrix} 1 & -x_0B \\ 0 & B \end{pmatrix}$$

to get good rational approximations to x_0 . We can note that $\overrightarrow{(1, x_0)}M_2 = \overrightarrow{(1, 0)}$ and that the transformation matrix T for which $\overrightarrow{TM_2}$ is lattice-reduced has the property that $\overrightarrow{(1, 0)}T$ is approximately proportional to $\overrightarrow{(1, x_0)}$. In four dimensions we take a point $(1 : x_0 : y_0 : z_0)$ on some curve, assuming that derivatives of y and z with respect to x are defined at this point. The matrix we use here is⁽¹²⁾

$$M_4 = \begin{pmatrix} 1 & -x_0B & (y'x_0 - y_0)B^2 & (-e(y'x_0 - y_0) + z'x_0 - z_0)B^3 \\ 0 & B & y'B^2 & (ey' - z')B^3 \\ 0 & 0 & B^2 & -eB^3 \\ 0 & 0 & 0 & B^3 \end{pmatrix}.$$

Here $e = z''/y''$ and all the derivatives are with respect to x and are to be evaluated at $(1 : x_0 : y_0 : z_0)$. Note that if we have computed $(1 : x_0 : y_0 : z_0)$ to H digits of precision, we must “lift” it to precision $3H$ to use this. Similar to the 2-dimensional case of above we have that $\overrightarrow{(1, x_0, y_0, z_0)}M_4 = \overrightarrow{(1, 0, 0, 0)}$, and $\overrightarrow{(1, 0, 0, 0)}T$ is approximately proportional to $\overrightarrow{(1, x_0, y_0, z_0)}$.

4.2. Results. — We now combine descent with the Heegner point method. We assume that we have a cover $C \rightarrow E$, and for each point $\mathcal{P}(\dot{z})$ given by the above algorithm we compute its real pre-images on C . For a 2-covering quartic, the x -coordinate has size $H/4$, but the y -coordinate on the quartic will be of size $H/2$. Either continued fractions on the x -coordinate or 3-dimensional lattice reduction on both coordinates and the curve requires a precision of $H/2$ digits — however, if our 2-cover is given as an intersection of quadrics in \mathbf{P}^3 , then we only need a precision of $(H/2)(2/3)$ since the Elkies method does better in higher dimension. For a 4-cover represented

⁽¹¹⁾This description is due to Elkies and is noted by Womack ([30, Section 2.9]).

⁽¹²⁾The 3-dimensional version is just the upper-left corner.

as an intersection of quadrics in \mathbf{P}^3 , the coordinates are of size $H/8$, and so we need a precision of $(H/8)(2/3)$ to recover our point.

We now give two examples of Heegner points of large height.⁽¹³⁾ First we consider E given by $[0, 1, 1, -4912150272, -132513750628709]$, for which $N = 421859$. Here the Heegner point is of height $3239.048+$. We refer the reader to [30] for how to do a 4-descent. The intersection of quadrics that gives the 4-cover is given by the symmetric matrices

$$\begin{pmatrix} 1 & 3 & 14 & 4 \\ 3 & 7 & 9 & 8 \\ 14 & 9 & -8 & 19 \\ 4 & 8 & 19 & 13 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 16 & -10 & 5 & -5 \\ -10 & 29 & -3 & -5 \\ 5 & -3 & -1 & 8 \\ -5 & -5 & 8 & 13 \end{pmatrix}.$$

We used $D = -795$ for which the index is 4. The class group has size 4; upon using the pairing from complex conjugation we need only 2 forms, which we can take to be $(421859, 234525, 32595)$ and $(421859, 384997, 87839)$. We need to use about $3239/12 \log(10) \approx 120$ digits of precision and take around 1.3 million terms of the L -series. For our approximation to a generator on \mathbf{C}/Λ we get

$$\dot{z} = 0.00825831518406814312450985646222558391095207954623175715662897127635126006560626891914983130574212343000780426018430276055.$$

We find the real pre-images of this on the 4-cover (this can be done via a resultant computation) and then via 4-dimensional lattice reduction we obtain the point

(905858492225062101133902424932526542192474474854331313031216338204053880670077944701302491852572823731202634266219944146702509489824532529044887987947859472355124939471295729.
5820784849940856724925010090474560451749158462165410106549337689496036041318008019546159331652386264579879746727095434743171075808134671399964513924870807157340785327661071757.
526601834730083187508441064291825052458007522956623515279464248174730240287018424148701587123000269221530600465998289112443045957965377530412500941202059568743656911281766881.
120566343955994724443268162651063750286159426913472902595469949397024821835521392662156637479278825638673289873881696613629128450397637909394604102581491799075589326751709650806).

which can then be mapped back to E . Even though we only used 120 digits of precision in our computation of a real approximation to the Heegner point, we can recover a point with approximately $3/2$ as many digits. Note that if we did not use descent, but recovered the point on the original curve using the Cremona-Silverman method, we would need 12 times the precision and 12 times as many terms in the L -series — this could be a total time factor of as much as 12^3 , depending upon the efficiency of our high-precision arithmetic. This computation (including 2-descent and 4-descent which each take a second) takes less than a minute.

Finally we give a more extreme example — this is the largest example which we have computed. The curve is from the database of Stein and Watkins [25]. Let E be given by $[0, 0, 1, -5115523309, -140826120488927]$, for which $N = 66157667$ and the Heegner point is of height $12557+$. The intersection of quadrics that gives the 4-cover is given by the two symmetric matrices

$$\begin{pmatrix} 0 & 1 & 3 & 3 \\ 1 & 5 & -1 & -6 \\ 3 & -1 & 8 & -2 \\ 3 & -6 & -2 & 16 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 12 & -21 & -10 & -68 \\ -21 & -13 & -7 & -27 \\ -10 & -7 & 3 & -7 \\ -68 & -27 & -7 & 15 \end{pmatrix}.$$

We select $D = -1435$ for which our 2 forms are $(66157667, 2599591, 25537)$ and $(66157667, 37610323, 5345323)$ and the index is 2. We need 460 digits of precision and

⁽¹³⁾These examples exemplify the experimental and heuristic correlation between large heights and large cancellation in $c_4^3 - c_6^2 = 1728\Delta$, since Ω_{re} can then be unusually small for a given $|\Delta|$.

- [4] W. Bosma, J. Cannon, C. Playoust, *The Magma algebra system. I. The user language*. In *Computational algebra and number theory* Proceedings of the 1st MAGMA Conference held at Queen Mary and Westfield College, London, August 23–27, 1993. Edited by J. Cannon and D. Holt, Elsevier Science B.V., Amsterdam (1997), 235–265. Cross-referenced as *J. Symbolic Comput.* **24** (1997), no. 3-4, 235–265. Online at magma.maths.usyd.edu.au
- [5] B. J. Birch, H. P. F. Swinnerton-Dyer, *Notes on elliptic curves. I. II.* *J. Reine Angew. Math.* **212** (1963), 7–25, **218** (1965), 79–108.
- [6] C. Breuil, B. Conrad, F. Diamond, R. Taylor *On the modularity of elliptic curves over \mathbf{Q} : wild 3-adic exercises*. *J. Amer. Math. Soc.* **14** (2001), no. 4, 843–939.
- [7] D. Bump, S. Friedberg, J. Hoffstein, *Nonvanishing theorems for L -functions of modular forms and their derivatives*. *Invent. Math.* **102** (1990), no. 3, 543–618.
- [8] H. Cohen, *A Course in Computational Algebraic Number Theory*. Grad. Texts in Math. **138**, Springer-Verlag, 1993.
- [9] H. Cohen, *Diophantine Equations, p -adic Numbers, and L -functions*, (to appear).
- [10] B. Conrad, F. Diamond, R. Taylor, *Modularity of certain potentially Barsotti-Tate Galois representations*. *J. Amer. Math. Soc.* **12** (1999), no. 2, 521–567.
- [11] H. Darmon, S.-W. Zhang, ed., *Heegner Points and Rankin L -series*. Math. Sci. Res. Inst. Publ., **49**, Cambridge Univ. Press, Cambridge, 2004. Available online from www.msri.org/publications/books/Book49/contents.html
- [12] F. Diamond, *On deformation rings and Hecke rings*. *Ann. of Math. (2)* **144** (1996), no. 1, 137–166.
- [13] N. D. Elkies, *Heegner point computations*. In *Algorithmic Number Theory*, Proceedings of the First International Symposium (ANTS-I) held at Cornell University, Ithaca, New York, May 6–9, 1994. Edited by L. M. Adleman and M.-D. Huang. Lecture Notes in Computer Science, **877**. Springer-Verlag, Berlin (1994), 122–133. N. D. Elkies, *Curves $Dy^2 = x^3 - x$ of odd analytic rank*. In *Algorithmic number theory*, Proceedings of the 5th International Symposium (ANTS-V) held at the University of Sydney, Sydney, July 7–12, 2002. Edited by C. Fieker and D. R. Kohel. Lecture Notes in Computer Science, **2369**. Springer-Verlag, Berlin (2002), 244–251. Available online at arxiv.org/math.NT/0208056
- [14] N. D. Elkies, *Rational points near curves and small nonzero $|x^3 - y^2|$ via lattice reduction*. In *Algorithmic number theory*, Proceedings of the 4th International Symposium (ANTS-IV) held at the Universiteit Leiden, Leiden, July 2–7, 2000. Edited by W. Bosma, Lecture Notes in Computer Science, **1838**. Springer-Verlag, Berlin (2000), 33–63. Available online at arxiv.org/math.NT/0005139
- [15] B. H. Gross, D. B. Zagier, *Heegner points and derivatives of L -series*. *Invent. Math.* **84** (1986), no. 2, 225–320.
- [16] Y. Hayashi, *Die Rankinsche L -Funktion und Heegner-Punkte für allgemeine Diskriminanten*. (German) [The Rankin L -function and Heegner points for general discriminants]. Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, 1993. Bonner Mathematische Schriften **259**. Universität Bonn, Mathematisches Institut, Bonn, 1994. viii+157 pp.

- Y. Hayashi, *The Rankin's L-function and Heegner points for general discriminants*. Proc. Japan Acad. Ser. A Math. Sci. **71** (1995), no. 2, 30–32.
- [17] E. Liverance, *Heights of Heegner points in a family of elliptic curves*. Ph. D. Thesis, Univ. of Maryland, 1993.
- [18] A. MacLeod, *A note on the curve $Y^2 = (X + p)(X^2 + p^2)$* . Rocky Mountain J. Math. **34** (2004), no. 1, 263–267.
- [19] P. Monsky, *Mock Heegner points and congruent numbers*. Math. Z. **204** (1990), no. 1, 45–67.
- [20] P. Monsky, *Three constructions of rational points on $Y^2 = X^3 \pm NX$* . Math. Z. **209** (1992), no. 3, 445–462.
- P. Monsky, *Errata: “Three constructions of rational points on $Y^2 = X^3 \pm NX$.”* Math. Z. **212** (1993), no. 1, 141. Corrects Lemma 4.7 and Theorem 4.8.
- [21] G. Shimura, *Abelian varieties with complex multiplication and modular functions*. Princeton Mathematical Series, **46**, Princeton University Press, 1998. Expanded version of [22].
- [22] G. Shimura, Y. Taniyama, *Complex multiplication of abelian varieties and its applications to number theory*. Publications of the Math. Soc. of Japan, **6**, The Math. Soc. of Japan, Tokyo, 1961, xi+159 pp. Revision of the first six chapters of *Modern Number Theory* (in Japanese), Kyōritsu Shuppan, Tokyo, 1957.
- [23] J. H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*. Springer-Verlag, 1994.
- [24] J. H. Silverman, *Computing rational points on rank 1 elliptic curves via L-series and canonical heights*. Math. Comp. **68** (1999), no. 226, 835–858.
- [25] W. A. Stein, M. Watkins, *A Database of Elliptic Curves—First Report*. In *Algorithmic number theory*, Proceedings of the 5th International Symposium (ANTS-V) held at the University of Sydney, Sydney, July 7–12, 2002. Edited by C. Fieker and D. R. Kohel. Lecture Notes in Computer Science, **2369**. Springer-Verlag, Berlin (2002), 267–275.
- [26] N. M. Stephens, *Computation of rational points of elliptic curves using Heegner points*. In *Number Theory and Applications*, Proceedings of the NATO Advanced Study Institute held in Banff, Alberta, April 27–May 5, 1988. Edited by R. A. Mollin. NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences, **265**. Kluwer Academic Publishers Group, Dordrecht (1989), 205–214.
- [27] G. Stevens, *Stickelberger elements and modular parametrizations of elliptic curves*. Invent. Math. **98** (1989), no. 1, 75–106.
- [28] R. Taylor, A. Wiles, *Ring-theoretic properties of certain Hecke algebras*. Ann. of Math. (2) **141** (1995), no. 3, 553–572.
- [29] A. Wiles, *Modular elliptic curves and Fermat's last theorem*. Ann. of Math. (2) **141** (1995), no. 3, 443–551.
- [30] T. Womack, *Explicit Descent on Elliptic Curves*. Ph. D. Thesis, Univ. of Nottingham, 2003.