Panoramas & Synthèses
36, 2012, p. 81–97

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.

***Résumé* (Quelques remarques sur le calcul des point de Heegner). —** Nous donnons une vue d'ensemble de la théorie des points de Heegner pour les courbes elliptiques, puis décrivons diverses idées nouvelles qui permettent le calcul des points rationnels des courbes de rang 1. En particulier, nous discutons l'idée de Cremona (suivant Silverman) de reconnaître un point rationnel *via* sa hauteur, l'idée de Delaunay d'utiliser les involutions d'Atkin-Lehner lors de la sélection des paramètres auxiliaires, et l'idée d'Elkies combinant descente et réduction de réseau, qui peuvent diminuer drastiquement la précision requise pour mener à bien le calcul.

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 Section 8.6 in Henri Cohen's latest book [**9**], and also owes a great debt to Christophe Delaunay. The ideas in the section about lattice reduction are largely due to Noam Elkies. The author also thanks R. Breu and the Arizona Winter School students of 2006 for some comments. We do not delve deeply into the theory of

2000 *Mathematics Subject Classification.* — 11G05, 11G40, 14G05.
***Key words and phrases.* —** Elliptic curves, Heegner points, descent, lattice reduction.

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 (France) for financial support. The algorithms described here have been implemented in the Magma computer algebra system [**5**]; 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.

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 τ 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 associated form $f_\tau = (A, B, C)$. 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. □

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 $\Delta\big(\gamma(\tau)\big) = \Delta(\tau)$ since the discriminant is fixed. A computation shows that $\gamma \in \Gamma_0(N)$ and $\tau \in \mathcal{H}_N^D$ imply $\gamma(\tau) \in \mathcal{H}_N^D$. □

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

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

Definition 2.5. — We let $\mathcal{S}(D, N)$ be the set of square roots mod $2N$ of D mod $4N$

Upon writing $\mathcal{C}l(K)$ for the class group of a number field K, we note (omitting details) that we have a bijection between the sets $\mathcal{H}_N^D/\Gamma_0(N)$ and $\mathcal{S}(D, N) \times \mathcal{C}l\big(\mathbf{Q}(\sqrt{D})\big)$, given by mapping $[\tau] \in \mathcal{H}_N^D/\Gamma_0(N)$ to $(B \bmod 2N) \times [\mathbf{Z} + \tau\mathbf{Z}]$ where $f_\tau = (A, B, C)$, and in the other direction, given $\beta \times l \in \mathcal{S}(D, N) \times \mathcal{C}l\big(\mathbf{Q}(\sqrt{D})\big)$ we take

$(A, B, C) \in l$ with $N|A$ and $B \equiv \beta \pmod{2N}$, and then $\tau = \frac{-B+\sqrt{D}}{2A}$. We state this as:

Lemma 2.6. — *The sets $\mathcal{H}_N^D/\Gamma_0(N)$ and $\mathcal{S}(D,N) \times \mathcal{C}l\big(\mathbf{Q}(\sqrt{D})\big)$ are in bijection.*

From now on we let E be a global minimal model of a elliptic curve (over $\mathbf{Q}$) of conductor N, and take D to be a negative fundamental discriminant such that D is a square mod $4N$. We let $\mathbf{H}^\star$ 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 $\big(\wp(z), \wp'(z)\big)$ on E.

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

Proof. — This is due to Wiles and others [**32, 30, 12, 10, 6**]. We let ϕ be the associated map from $\mathbf{H}^\star/\Gamma_0(N)$ to $\mathbf{C}/\Lambda$. Explicitly, we have that $\tau \in \mathbf{H}^\star$ gets mapped to the complex point $\phi(\tau) = 2\pi i \int_{i\infty}^{\tau} \psi_E \, d\tau = \sum_n (a_n/n) e^{2\pi i n \tau}$, 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 [**29**] and [**27**] — for curves of rank 0, the Manin constant might be nonzero, as with $X_1(11)$, though the optimal curve under an $X_0(N)$-parametrisation is expected always to have trivial Manin constant). □

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

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

Proof. — This is the theorem of complex multiplication of Shimura [**24**, §6.8]. We outline the proof of the initial statement, working *via* the modular j-function. We have that $j(\tau)$ is in the Hilbert class field H (see [**25**, II, 4.3]) and similarly with $j(N\tau)$. Thus $X_0(N)/H$ contains the moduli point m_τ 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 m_τ under $\hat{\phi}$ has coordinates in H. □

Note that $\mathcal{P}\big(\overline{\phi(\tau)}\big) = \overline{\mathcal{P}\big(\phi(\tau)\big)}$, so that there is no danger of confusing complex conjugation in $\mathbf{C}/\Lambda$ with complex conjugation of the coordinates of the point on E. Using the third fact of Theorem 2.8, we can trace $\mathcal{P}\big(\phi(\tau)\big)$ down to a point

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 $\Omega_{\mathrm{re}}/2$. The fundamental volume Ω_{vol} is the area of the period parallelogram.

in $\mathbf{Q}(\sqrt{D})$. Indeed, writing $K = \mathbf{Q}(\sqrt{D})$ (and H for its Hilbert class field) we get that

$$P = \mathrm{Trace}_{H/K}\big(\mathcal{P}\big(\phi(\tau)\big)\big) = \sum_{\sigma\in\mathrm{Gal}(H/K)} \mathcal{P}\big(\phi(\tau)\big)^{\sigma} = \sum_{m\in\mathcal{C}l(K)} \mathcal{P}\big(\phi(\beta\times l)\big)^{\mathrm{Artin}(m)}$$
$$= \sum_{m\in\mathcal{C}l(K)} \mathcal{P}\big(\phi\big(\beta\times lm^{-1}\big)\big) = \sum_{m\in\mathcal{C}l(K)} \mathcal{P}\big(\phi(\beta\times m)\big)$$

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 $\mathbf{C}/\Lambda$ we have

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

which gives us that

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

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

Definition 2.9. — We write $\mathcal{H}_N^D(\beta)$ for the 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 $\mathcal{C}l\big(\mathbf{Q}(\sqrt{D})\big)$, we get

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

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 $\beta\in\mathcal{S}(D,N)$, find τ-representatives for $\hat{\mathcal{H}}_N^D(\beta)$, compute $\phi(\tau)$ for each, sum these in $\mathbf{C}/\Lambda$, 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; we denote it by $\hat{h}$. We write E_D for the quadratic twist of E by D, and $w(D)$ is the number of roots of unity in $\mathbf{Q}(\sqrt{D})$, with $\omega(n)$ being the number of distinct prime factors of n.

Theorem 3.1. — *Let $E/\mathbf{Q}$ be a elliptic curve of analytic rank 1. Suppose $D < -3$ is a fundamental discriminant with D a square modulo $4N$ and $\gcd(D, 2N) = 1$ Then*[2]

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

Proof. — This is due to Gross and Zagier [**16**]. □

Calculations of Gross and Hayashi [**17, 18**] indicate that this height formula is likely to be true for all negative fundamental discriminants D that are square mod $4N$, at least if $a_p = -1$ for all primes p that divide $\gcd(D, N)$.

We now write $P = lG + T$ where G is a generator[3] of $E(\mathbf{Q})$ modulo torsion and T is a torsion point, so that $\hat{h}(P) = l^2 \hat{h}(G)$. Then we rewrite $L'(E, 1)$ using the Birch–Swinnerton-Dyer conjecture [**4**] to get the following:

Conjecture 3.2. — *With notations as above we have*[4]

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

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. — *Let $E/\mathbf{Q}$ have analytic rank 1. Find a non-torsion rational point.*

1. *Compute $L'(E, 1)$ to sufficient precision to verify that the given curve really is of analytic rank 1. 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 $\beta \in \mathcal{S}(D, N)$ 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 $u = 1, \ldots, lm$ for both*

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

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

2. The reader can note that the latter $2^{\omega(\gcd(D,N))}\big(\frac{w(D)}{2}\big)^2$ term is trivial in the case of Gross and Zagier, but needs to be included in the general conjecture.

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

4. 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.