The efficient calculation of Stark-Heegner points via overconvergent modular symbols

Henri Darmon
Robert Pollack

May 10, 2004

Contents

Introduction

Let E be an elliptic curve over $\mathbb{Q}$ of conductor N and let p be a prime which divides N exactly. Since E is modular, it corresponds to a cusp form f of weight two on Hecke's congruence group $\Gamma_0(N)$. The pth Fourier coefficient of this modular form, denoted a_p, is equal to 1 (resp. -1) if E has split (resp. non-split) multiplicative reduction at p. Let

$\mathcal{H}_p := \mathbb{P}_1(\mathbb{C}_p) - \mathbb{P}_1(\mathbb{Q}_p)$ denote the p-adic upper half plane, and fix a real quadratic field K in which the prime p is inert, together with an embedding of $\bar{K}$ into $\bar{\mathbb{Q}}_p \subset \mathbb{C}_p$.

Using certain periods attached to f, the article [Dar1] associates to any $\tau \in \mathcal{H}_p \cap K$ a so-called *Stark-Heegner point* $P_\tau \in E(K_p)$ and conjectures that this point is defined over a specific abelian extension—more precisely, a ring class field—of K. Stark-Heegner points appear to behave like classical Heegner points in many ways, except that the imaginary quadratic base field is replaced by a real quadratic field. An algorithm for computing them, in the case of elliptic curves of prime conductor, is described in [DG], where it is used to test the conjecture of [Dar1] numerically.

The calculations in [DG] raise the question of whether Stark-Heegner points lead to an efficient method (*conditional* on the conjectures of [Dar1]) for finding global points on elliptic curves comparable to the approach based on classical Heegner points, as it is described in [El] for example. The major obstacle to putting Stark-Heegner points to such a "practical" use is that the definition of P_τ rests on certain p-adic integrals whose direct evaluation as a limit of Riemann sums has exponential running time—namely, the number of arithmetic operations required to perform this evaluation with an accuracy of M significant digits is proportional to p^M. Since each extra digit of desired accuracy multiplies the running time of the algorithm roughly by p, the evaluation of P_τ following the approach of [DG] becomes intractable for even moderate values of M. This explains why [DG] was only able to verify the conjectures of [Dar1] to at most 8 or 9 digits of p-adic accuracy. In particular, the calculation of P_τ could almost never be used to independently *discover* a global point on $E(K)$, except in some instances where this point is of small height, when it could have been found just as easily by inspection. So while [DG] did produce convincing numerical evidence for the conjecture of [Dar1], by no means could the algorithms that it used be touted as a practical method for constructing global points on elliptic curves.

The main purpose of this note is to present a significant improvement to the algorithm of [DG] which runs in *polynomial time*, where the size of the problem is measured by the number M of desired p-adic digits of accuracy, the prime p being treated as a constant. This answers in the affirmative the question raised before Remark 1.7 in Section 1.2 of [DG].

The key to this new approach lies in the theory of *overconvergent modular symbols* developed in [PS1] and [PS2]. These modular symbols generalise the classical modular symbol $I_f : \mathbb{P}_1(\mathbb{Q}) \times \mathbb{P}_1(\mathbb{Q}) \longrightarrow \mathbb{Z}$ defined in terms of f by the rule

$$I_f\{r \to s\} := \frac{1}{\Omega^+} \operatorname{Re}\left(\int_r^s 2\pi i f(\tau) d\tau \right), \qquad r, s \in \mathbb{P}_1(\mathbb{Q}),$$

where Ω^+ is a suitable real period which can be chosen so that I_f becomes $\mathbb{Z}$-valued. This modular symbol defines the *Mazur–Swinnerton-Dyer measure* μ_f on $\mathbb{Z}_p^\times$ by the rule

$$\mu_f(a + p^r\mathbb{Z}_p) := a_p^{-r} \cdot I_f\{-a/p^r \to \infty\}, \quad a \in \mathbb{Z}. \tag{1}$$

Let $[x]$ denote the floor function of the real number x, and set

$$M' := \sup\{n \text{ such that } \operatorname{ord}_p(p^n/n) < M\}; \tag{2}$$

$$M'' := M + [\log(M')/\log(p)]. \tag{3}$$

A by-product of the theory of [PS1] and [PS2] is a method for computing the first M' moments of the measure μ_f,

$$\omega(a,k) := \int_{a+p\mathbb{Z}_p} (t-a)^k d\mu_f(t), \qquad 0 \le a \le p-1, \quad k = 0, 1, \ldots, M', \tag{4}$$

to an accuracy of $p^{-M''}$ in time which is polynomial in M. Section 1.3 explains how this data is enough to efficiently compute the Stark-Heegner points attached to E to an accuracy of p^{-M}—in fact, after the moments (4) have been precomputed and stored, the number of arithmetic operations required to evaluate one of the p-adic integrals involved in the definition of P_τ is $O(pM)$, a complexity which appears to be best possible.

The main contribution of the present paper resides in bringing together the ideas of [PS1], [PS2] and of [DG], and in the detailed glimpse into the phenomenology of Stark-Heegner points and Mordell-Weil groups over ring class fields of real quadratic fields made possible by the computational tools that emerge from this combination of ideas.

Section 1 recalls the definition of Stark-Heegner points and describes the new polynomial-time algorithm for computing them, while Section 3 describes the implementation of this algorithm as a Magma package called `shp` that can be downloaded from the world-wide web. Section 4 revisits the calculations of [DG], verifying them to hundreds instead of just 8 p-adic digits, filling in most of the mising data which the authors of [DG] were unable to supply due to the severely limited accuracy of their computations, and extending the range of positive discriminants for which the calculations are successfully performed.

We remark that the original motivation for the theory of overconvergent modular symbols arose from Coleman's generalisation of Hida's theory of p-adic families of ordinary eigenforms to the non-ordinary case. Since Dasgupta's thesis (cf. [Das], [DD]) it has become clear that Hida families are intimately connected to the theory of Stark-Heegner points. (That such families can be used to shed light on the theoretical, and not just computational, aspects of the Stark-Heegner point construction is a theme of the series of articles [BD1] and [BD2].)

1 Computing Stark-Heegner points

As in [DG], we confine our attention to the simplest case where the elliptic curve E has *prime* conductor p, so that K is any real quadratic field in which p is inert. As in the introduction, denote by f the weight two eigenform on $\Gamma_0(p)$ that is associated to E by Wiles' theorem, and write

$$\omega_f := 2\pi i f(\tau) d\tau$$

for the corresponding $\Gamma_0(p)$-invariant differential on the Poincaré upper half-plane $\mathcal{H}$. Let $\Gamma := \mathbf{PSL}_2(\mathbb{Z}[1/p])$, which acts on both $\mathcal{H}$ and $\mathcal{H}_p$ on the left by Möbius transformations, according to the rule

$$\gamma\tau = \frac{a\tau+b}{c\tau+d}, \qquad \text{where } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{5}$$

The Stark-Heegner point construction

$$\Gamma\backslash(\mathcal{H}_p \cap K) \longrightarrow E(K_p), \qquad \tau \mapsto P_\tau,$$

can be described in several stages.

1.1 Modular symbols

The group $\tilde{\Gamma} := \mathbf{PGL}_2(\mathbb{Z}[1/p]) \supset \Gamma$ is equipped with a homomorphism arising from the determinant

$$\det : \tilde{\Gamma} \longrightarrow \mathbb{Z}[1/p]^\times/(\mathbb{Z}[1/p]^\times)^2.$$

The target of this homomorphism is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, and given $\gamma \in \tilde{\Gamma}$ we define $|\gamma|_p, |\gamma|_\infty \in \{0,1\}$ by the rules

$$|\gamma|_p := \begin{cases} 0 & \text{if } \mathrm{ord}_p(\det(\gamma)) \text{ is even;} \\ 1 & \text{if } \mathrm{ord}_p(\det(\gamma)) \text{ is odd,} \end{cases} \qquad |\gamma|_\infty := \begin{cases} 0 & \text{if } \det(\gamma) > 0; \\ 1 & \text{if } \det(\gamma) < 0, \end{cases}$$

so that $\gamma \in \tilde{\Gamma}$ belongs to Γ if and only if $|\gamma|_p = |\gamma|_\infty = 0$.

In addition to the left action by Möbius transformations given by equation (5), it will occasionally be useful to consider the right action of the group $\mathbf{PGL}_2^+(\mathbb{Q})$ of matrices with positive determinant on either $\mathcal{H}_p$ or $\mathcal{H}$ given by the rule

$$\tau\gamma = -\gamma^{-1}(-\tau) = \frac{b+d\tau}{a+c\tau}, \qquad \text{where } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Let $\Sigma_0(p) \subset M_2(\mathbb{Q}_p)$ denote the semi-group of matrices defined by

$$\Sigma_0(p) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2(\mathbb{Z}_p) \mid a \in \mathbb{Z}_p^\times,\ c \in p\mathbb{Z}_p,\ \ ad - bc \neq 0 \right\},$$

and let V be any $\mathbb{Z}$-module equipped with a right action by $\Sigma_0(p)$.

Definition 1.1. A V-valued modular symbol is a map $\varphi : \mathbb{P}^1(\mathbb{Q}) \times \mathbb{P}^1(\mathbb{Q}) \longrightarrow V$, denoted by $(r,s) \mapsto \varphi\{r \to s\}$, satisfying

$$\varphi\{r \to s\} + \varphi\{s \to t\} = \varphi\{r \to t\} \quad \text{ for all } r,s,t \in \mathbb{P}^1(\mathbb{Q}).$$

The space of all V-valued modular symbols, denoted by $\mathrm{Symb}(V)$, has the structure of a right $\Sigma_0(p)$-module by the rule

$$(\varphi\big|\gamma)(r \to s) := \varphi(\gamma r \to \gamma s)\big|\gamma,$$

where $\varphi \in \mathrm{Symb}\, V$ and $\gamma \in \Sigma_0(p)$. If G is any subgroup of $\Sigma_0(p)^\times$, a V-valued modular symbol is said to be G-equivariant, or to be *on* G, if G fixes it under this action. The space of all such modular symbols is denoted $\mathrm{Symb}_G(V)$.

We will focus mainly on the case where $G = \Gamma_0(p) = \Sigma_0(p) \cap \Gamma$. The space $\mathrm{Symb}_{\Gamma_0(p)}(V)$ is equipped with a right action by the Hecke operators T_ℓ ($\ell \neq p$) and U_p defined by the formulae

$$\varphi\big|T_\ell := \varphi\big|\left(\begin{smallmatrix} \ell & 0 \\ 0 & 1 \end{smallmatrix}\right) + \sum_{a=0}^{\ell-1} \varphi\big|\left(\begin{smallmatrix} 1 & a \\ 0 & \ell \end{smallmatrix}\right), \qquad \varphi\big|U_p := \sum_{a=0}^{p-1} \varphi\big|\left(\begin{smallmatrix} 1 & a \\ 0 & p \end{smallmatrix}\right).$$

The $\Gamma_0(p)$-invariant differential form ω_f gives rise to a $\mathbb{C}$-valued modular symbol $\tilde{I}_f$ on $\Gamma_0(p)$ (where $\mathbb{C}$ is equipped with the trivial $\Sigma_0(p)$-action) by viewing $\mathbb{P}_1(\mathbb{Q})$ as the boundary of the extended upper half plane $\mathcal{H}^* := \mathcal{H} \cup \mathbb{P}_1(\mathbb{Q})$, and setting

$$\tilde{I}_f\{r \to s\} := \int_r^s \omega_f, \qquad r, s \in \mathbb{P}_1(\mathbb{Q}).$$

Note that the integral in this formula does converge, because f is a cusp form. The following theorem of Manin and Drinfeld makes it possible to convert $\tilde{I}_f$ into a $\mathbb{Z}$-valued modular symbol.

Proposition 1.2. *There exist unique periods Ω^+ and Ω^- in $\mathbb{R}^{>0}$ with the property that the functions I_f^+ and I_f^- on $\mathbb{P}_1(\mathbb{Q}) \times \mathbb{P}_1(\mathbb{Q})$ defined by*

$$I_f^+\{r \to s\} := \frac{1}{\Omega^+} \operatorname{Re}(\tilde{I}_f\{r \to s\}), \qquad I_f^-\{r \to s\} := \frac{1}{\Omega^-} \operatorname{Im}(\tilde{I}_f\{r \to s\}),$$

take values in $\mathbb{Z}$ and in no proper ideal of $\mathbb{Z}$.

The functions $I_f^+$and I_f^- belong to $\mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Z}) \subset \mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Z}_p)$, and are called the *even* and *odd modular symbols* attached to f respectively. The terminology of even and odd is justified by the further transformation properties

$$I_f^+\{-r \to -s\} = I_f^+\{r \to s\}, \qquad I_f^-\{-r \to -s\} = -I_f^-\{r \to s\} \tag{6}$$

satisfied by I_f^+ and I_f^-. Fix once and for all a choice of a "sign at infinity" $w_\infty \in \{1, -1\}$ and set

$$I_f = \begin{cases} I_f^+ & \text{if } w_\infty = 1; \\ I_f^- & \text{if } w_\infty = -1. \end{cases} \tag{7}$$

The case where $w_\infty = 1$ will be referred to as the *even case*, and that where $w_\infty = -1$, as the *odd case*.

1.2 p-adic measures

A *locally analytic distribution* on $\mathbb{P}_1(\mathbb{Q}_p)$ is a continuous $\mathbb{C}_p$-linear functional on the space of locally analytic $\mathbb{C}_p$-valued functions on $\mathbb{P}_1(\mathbb{Q}_p)$. If μ is such a distribution, and h is a locally analytic function, we write

$$\int_{\mathbb{P}_1(\mathbb{Q}_p)} h(t) d\mu(t) := \mu(h).$$

The distribution μ is completely determined by the values it takes on the characteristic functions $\mathbf{1}_U$ of the compact open subsets $U \subset \mathbb{P}_1(\mathbb{Q}_p)$, and it is customary to write

$$\mu(U) := \mu(\mathbf{1}_U) =: \int_U d\mu(t).$$

In this way μ gives rise to a finitely additive $\mathbb{C}_p$-valued function on the set of compact open subsets of $\mathbb{P}_1(\mathbb{Q}_p)$. The distribution μ is said to be a *measure* if $\mu(U)$ is p-adically bounded,

for all compact open $U \subset \mathbb{P}_1(\mathbb{Q}_p)$, and is said to be *integral* if it satisfies the stronger (and somewhat artificial, from the point of view of p-adic functional analysis, although this condition turns out to be useful in the Stark-Heegner point construction) condition

$$\mu(U) \text{ belongs to } \mathbb{Z}, \quad \text{for all compact open } U \subset \mathbb{P}_1(\mathbb{Q}_p).$$

Write $\mathcal{D}(\mathbb{P}_1(\mathbb{Q}_p))$ for the space of locally analytic distributions on $\mathbb{P}_1(\mathbb{Q}_p)$, and $\mathcal{M}(\mathbb{P}_1(\mathbb{Q}_p))$ for the submodule of integral measures on $\mathbb{P}_1(\mathbb{Q}_p)$.

Proposition 1.3. *There is a unique system $\mu_f\{r \to s\}$ of integral measures on $\mathbb{P}_1(\mathbb{Q}_p)$, indexed by $r, s \in \mathbb{P}_1(\mathbb{Q})$, and satisfying the rules*

1. $\mu_f\{r \to s\} + \mu_f\{s \to t\} = \mu_f\{r \to t\}$, *for all $r, s, t \in \mathbb{P}_1(\mathbb{Q})$;*
2. $\mu_f\{r \to s\}(\mathbb{P}_1(\mathbb{Q}_p)) = 0$, *for all $r, s \in \mathbb{P}_1(\mathbb{Q})$;*
3. $\mu_f\{r \to s\}(\mathbb{Z}_p) = I_f\{r \to s\}$, *for all $r, s \in \mathbb{P}_1(\mathbb{Q})$;*

as well as the following invariance property under $\gamma \in \tilde{\Gamma}$:

$$\mu_f\{\gamma r \to \gamma s\}(\gamma U) = a_p^{|\gamma|_p} w_\infty^{|\gamma|_\infty} \cdot \mu_f\{r \to s\}(U), \tag{8}$$

for all compact open subsets $U \subset \mathbb{P}_1(\mathbb{Q}_p)$.

Proof. The proof of this proposition is identical to that of Proposition 2.5 of [DD], which treats the case where f is replaced by certain weight two Eisenstein series. Crucial to both proofs is the fact that the eigenvalue of the Hecke operator U_p acting on f is $a_p = \pm 1$. Some of the motivation for introducing the system of measures $\mu_f\{r \to s\}$, based on the theory of p-adic integration and on an analogy with periods of Hilbert modular forms, is explained in [Dar1], Sections 1.1 and 1.2. □

Given the system $\mu_f\{r \to s\}$, write

$$\mu_f := \mu_f\{0 \to \infty\} \in \mathcal{M}(\mathbb{P}_1(\mathbb{Q}_p)). \tag{9}$$

The following lemma shows that this notation is consistent with the definition given in (1) of the Mazur–Swinnerton-Dyer measure attached to I_f.

Lemma 1.4. *For all $a \in \mathbb{Z}$,*

$$\mu_f(a + p^r\mathbb{Z}_p) = a_p^{-r} \cdot I_f\{-a/p^r \to \infty\},$$

so that the restriction of μ_f to $\mathbb{Z}_p^\times$ is the Mazur–Swinnerton-Dyer measure attached to I_f by equation (1).

Proof. This follows by letting $\gamma = \left(\begin{smallmatrix} p^r & a \\ 0 & 1 \end{smallmatrix}\right) \in \tilde{\Gamma}$ and noting that

$$\mu_f(a + p^r\mathbb{Z}_p) = \mu_f\{0 \longrightarrow \infty\}(\gamma\mathbb{Z}_p) = a_p^{-r} \cdot \mu_f\{-a/p^r \to \infty\}(\mathbb{Z}_p) = a_p^{-r} \cdot I_f\{-a/p^r \to \infty\},$$

where the second equality follows from equation (8) in Proposition 1.3, while the last follows from the defining property 3 of $\mu_f\{r \to s\}$ in that Proposition. □

1.3 Double integrals

The measures $\mu_f\{r \to s\}$ can be used to define certain $\mathbb{C}_p$-valued *double integrals* by choosing a p-adic logarithm

$$\log : \mathbb{C}_p^\times \longrightarrow \mathbb{C}_p$$

and setting, for $r, s \in \mathbb{P}_1(\mathbb{Q})$ and $\tau_1, \tau_2 \in \mathcal{H}_p$:

$$\int_{\tau_1}^{\tau_2}\int_r^s \omega_f := \int_{\mathbb{P}_1(\mathbb{Q}_p)} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f\{r \to s\}(t) \tag{10}$$

$$:= \lim_{\mathcal{C}=\{U_\alpha\}} \sum_\alpha \log\left(\frac{t_\alpha-\tau_2}{t_\alpha-\tau_1}\right) \mu_f\{r \to s\}(U_\alpha), \tag{11}$$

where the limit of Riemann sums in (11) is taken over finer and finer coverings of $\mathbb{P}_1(\mathbb{Q}_p)$ by mutually disjoint compact open subsets U_α, and t_α is a sample point in U_α. This p-adic "double integral" satisfies the additivity properties that are suggested by the notation:

$$\int_{\tau_1}^{\tau_2}\int_r^s \omega_f + \int_{\tau_1}^{\tau_2}\int_s^t \omega_f = \int_{\tau_1}^{\tau_2}\int_r^t \omega_f, \tag{12}$$

$$\int_{\tau_1}^{\tau_2}\int_r^s \omega_f + \int_{\tau_2}^{\tau_3}\int_r^s \omega_f = \int_{\tau_1}^{\tau_3}\int_r^s \omega_f, \tag{13}$$

as well as being invariant under translation by Γ:

$$\int_{\gamma\tau_1}^{\gamma\tau_2}\int_{\gamma r}^{\gamma s} \omega_f = w_\infty^{|\gamma|_\infty} a_p^{|\gamma|_p} \cdot \int_{\tau_1}^{\tau_2}\int_r^s \omega_f \quad \text{for all } \gamma \in \tilde{\Gamma}. \tag{14}$$

Because the measures $\mu_f\{r \to s\}$ are integral, a multiplicative refinement of (10) can be defined as in [Dar1] by formally exponentiating (11):

$$\prod_r^s\int_{\tau_1}^{\tau_2} \omega_f = \prod_{\mathbb{P}_1(\mathbb{Q}_p)} \left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f\{r \to s\}(t) := \lim_{\mathcal{C}=\{U_\alpha\}} \prod_\alpha \left(\frac{t_\alpha-\tau_2}{t_\alpha-\tau_1}\right)^{\mu_f\{r\to s\}(U_\alpha)}, \tag{15}$$

so that the expression appearing on the right of (15) is a limit of "Riemann products" instead of Riemann sums. Note that the multiplicative integral is a more precise invariant, since the additive integral can be recovered from it by the rule

$$\int_{\tau_1}^{\tau_2}\int_r^s \omega_f = \log\left(\prod_{\tau_1}^{\tau_2}\int_r^s \omega_f\right),$$

while the p-adic logarithm is never injective on $\mathbb{C}_p^\times$, even modulo torsion.

If n is any integer, we also define the integrals

$$\int_{\tau_1}^{\tau_2}\int_r^s n\omega_f, \quad \prod_{\tau_1}^{\tau_2}\int_r^s n\omega_f$$

in the natural way by replacing the measures $\mu_f\{r \to s\}$ by $n \cdot \mu_f\{r \to s\}$.

Since the p-adic "double integral" attached to f plays a key role in the definition of Stark-Heegner points, it becomes important to evaluate this integral to high accuracy in an efficient way.

The calculation of the double integrals. Let $\mathbb{Q}_{p^2}$ denote the quadratic unramified extension of $\mathbb{Q}_p$, let $\mathcal{O}$ denote its ring of integers, and let

$$\mathrm{red} : \mathbb{P}_1(\mathbb{Q}_{p^2}) \longrightarrow \mathbb{P}_1(\mathbb{F}_{p^2})$$

denote the natural reduction map. Consider the subsets of $\mathcal{H}_p$ defined by

$$\mathcal{H}_p(\mathbb{Q}_{p^2}) := \mathbb{P}_1(\mathbb{Q}_{p^2}) - \mathbb{P}_1(\mathbb{Q}_p); \tag{16}$$

$$\mathcal{H}_p^0 := \{\tau \in \mathbb{P}_1(\mathbb{Q}_{p^2}) \text{ such that } \mathrm{red}(\tau) \notin \mathbb{P}_1(\mathbb{F}_p)\} \subset \mathcal{H}_p(\mathbb{Q}_{p^2}). \tag{17}$$

We will content ourselves with explaining how to calculate the double integral and its multiplicative counterpart when the p-adic endpoints of integration τ_1 and τ_2 belong to $\mathcal{H}_p^0$, which turns out to be sufficient for our purposes.

Lemma 1.5. *If τ_1 and τ_2 belong to $\mathcal{H}_p^0$, then*

$$\oint_{\tau_1}^{\tau_2}\int_r^s \omega_f \text{ belongs to } \mathcal{O}^\times.$$

In particular, to compute it to an accuracy of p^{-M}, it is enough to compute

$$\oint_{\tau_1}^{\tau_2}\int_r^s \omega_f \pmod{p}, \tag{18}$$

and

$$\int_{\tau_1}^{\tau_2}\int_r^s \omega_f \pmod{p^{M-1}}. \tag{19}$$

The proof of this lemma is a direct consequence of the definitions. It is the main reason why it is convenient to work under the assumption that τ_1 and τ_2 belong to $\mathcal{H}_p^0$.

The calculation of (18) can be carried out in $O(p)$ operations using the formula

$$\oint_{\tau_1}^{\tau_2}\int_r^s \omega_f = \prod_{t=0}^{p-1}\left(\frac{t-\tau_2}{t-\tau_1}\right)^{\mu_f\{r\to s\}(t+p\mathbb{Z}_p)} \pmod{p}. \tag{20}$$

This running time is quite good when p is of reasonable size, and there is reason to believe that its efficiency cannot be improved upon in practice. We now turn to the more serious issue of calculating (19).

Continued fractions. Two elements $[a/b]$ and $[c/d]$ of $\mathbb{P}_1(\mathbb{Q})$, represented by fractions in lowest terms, are said to be *adjacent* if $ad - bc = \pm 1$. The convergents in the continued fraction expansion of $t \in \mathbb{Q}$ yield a sequence of adjacent rational numbers joining t to ∞; hence any two elements of $\mathbb{P}_1(\mathbb{Q})$ can be joined by such a sequence. (The usefulness of this elementary fact for calculations with modular symbols was already observed in [Man].) The additivity

property (12) reduces the problem of evaluating (19) to the special case where r and s are adjacent elements. But any pair of adjacent elements of $\mathbb{P}_1(\mathbb{Q})$ is $\mathbf{PSL}_2(\mathbb{Z})$-equivalent to the pair $(0,\infty)$:

$$\gamma r = 0, \quad \gamma s = \infty, \qquad \text{for some } \gamma \in \mathbf{PSL}_2(\mathbb{Z}).$$

Therefore by (14), we have

$$\int_{\tau_1}^{\tau_2}\int_{r}^{s} \omega_f = \int_{\gamma\tau_1}^{\gamma\tau_2}\int_{0}^{\infty} \omega_f.$$

Since the subset $\mathcal{H}_p^0$ is preserved under the action of $\mathbf{PSL}_2(\mathbb{Z})$, the problem of evaluating (19) has been reduced to that of efficiently evaluating to high accuracy, for any $\tau_1, \tau_2 \in \mathcal{H}_p^0$, the double integral

$$J(\tau_1,\tau_2) := \int_{\tau_1}^{\tau_2}\int_{0}^{\infty} \omega_f = \int_{\mathbb{P}_1(\mathbb{Q}_p)} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f(t), \tag{21}$$

where μ_f is defined by (9).

The main departure from the algorithm of [DG] lies in the approach that is followed to compute $J(\tau_1,\tau_2)$. In [DG], the region $\mathbb{P}_1(\mathbb{Q}_p)$ was broken up into $(p+1)p^{M-1}$ residue discs modulo p^M, and it was shown that the corresponding Riemann sum yields the integral of (21) to an accuracy of M significant p-adic digits. Since there are exponentially many discs on which the integrand needs to be evaluated, this approach quickly becomes intractable when M is even moderately large.

For any integers $k \geq 0$ and $0 \leq a \leq p-1$, define the *kth moment of μ_f around a* by the rule

$$\omega(a,k) := \int_{a+p\mathbb{Z}_p} (t-a)^k d\mu_f(t).$$

Our polynomial-time algorithm starts with the observation that $J(\tau_1,\tau_2)$ can be expressed succintly in terms of these moments. To see this, first break up $J(\tau_1,\tau_2)$ as a sum of $p+1$ contributions arising from the $p+1$ residue discs on $\mathbb{P}_1(\mathbb{Q}_p)$:

$$J(\tau_1,\tau_2) = J_\infty(\tau_1,\tau_2) + \sum_{a=0}^{p-1} J_a(\tau_1,\tau_2),$$

where

$$J_\infty(\tau_1,\tau_2) := \int_{\mathbb{P}_1(\mathbb{Q}_p)-\mathbb{Z}_p} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f(t), \quad J_a(\tau_1,\tau_2) := \int_{a+p\mathbb{Z}_p} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f(t).$$

To evaluate the term $J_\infty(\tau_1,\tau_2)$, observe that

$$d\mu_f(-1/t) = d\mu_f\{0\to\infty\}(-1/t) = d\mu_f\{\infty\to 0\}(t) = -d\mu_f(t).$$

Hence we can make the change of variables $t \mapsto -1/t$ to obtain:

$$J_\infty(\tau_1,\tau_2) = -\int_{p\mathbb{Z}_p} \log\left(\frac{1+t\tau_2}{1+t\tau_1}\right) d\mu_f(t) = J_\infty(\tau_2) - J_\infty(\tau_1),$$

where

$$J_\infty(\tau) = -\int_{p\mathbb{Z}_p} \log(1+t\tau)d\mu_f(t) = -\int_{p\mathbb{Z}_p} \sum_{n=1}^{\infty}(-1)^{n+1}\frac{(t\tau)^n}{n}d\mu_f(t) = \sum_{n=1}^{\infty}(-1)^n\frac{\omega(0,n)}{n}\tau^n.$$

Recall the integers M' and M'' defined in equations (2) and (3). Since $\omega(0,n) \equiv 0 \pmod{p^n}$ and since τ belongs to $\mathcal{O}^\times$, it follows that

$$J_\infty(\tau) = \sum_{n=1}^{M'}(-1)^n\frac{\omega(0,n)}{n}\tau^n \pmod{p^M}. \tag{22}$$

Formula (22) makes it possible to evaluate $J_\infty(\tau)$ in time which is polynomial in M, *provided* the values of the first M' moments $\omega(0,n)$ are known in advance to M'' significant p-adic digits.

The evaluation of the term $J_a(\tau_1,\tau_2)$ is similar. More precisely, we may write:

$$J_a(\tau_1,\tau_2) = J_a(\tau_2) - J_a(\tau_1),$$

where

$$\begin{aligned} J_a(\tau) &= \int_{a+p\mathbb{Z}_p} \log(t-\tau)d\mu_f(t) \\ &= \omega(a,0)\log(\tau-a) + \int_{a+p\mathbb{Z}_p} \log\left(1-\frac{t-a}{\tau-a}\right)d\mu_f(t) \\ &= \omega(a,0)\log(\tau-a) - \sum_{n=1}^{\infty}\frac{\omega(a,n)}{n}\left(\frac{1}{\tau-a}\right)^n. \end{aligned}$$

As before, since $\omega(a,n) \equiv 0 \pmod{p^n}$ and since $1/(\tau-a)$ belongs to $\mathcal{O}^\times$,

$$J_a(\tau) = \omega(a,0)\log(\tau-a) - \sum_{n=1}^{M'}\frac{\omega(a,n)}{n}\left(\frac{1}{\tau-a}\right)^n \pmod{p^M}.$$

Just as for $J_\infty(\tau)$, the evaluation of $J_a(\tau)$ can therefore be carried out in time which is polynomial in M, given the data of

$$\omega(a,j) = \int_{a+p\mathbb{Z}_p}(t-a)^j d\mu_f(t) \pmod{p^{M''}}, \qquad 0 \le a \le p-1, \quad j = 0,1,\ldots,M'. \tag{23}$$

A procedure to calculate this (finite amount of) data in polynomial time via the theory of overconvergent modular symbols is explained in Section 2.

1.4 Indefinite integrals

Assume that E is the strong Weil curve in its $\mathbb{Q}$-isogeny class. Let q be Tate's p-adic period attached to E, and let $n = \#E(\mathbb{Q})_{\text{tors}}$. As is explained in [Dar1], the following conjecture can be viewed as a refinement of a conjecture of Mazur, Tate and Teitelbaum [MTT] which was proved by Greenberg and Stevens in [GS].

Conjecture 1.6. *There exists a unique $\mathbb{Q}_{p^2}^\times/q^{\mathbb{Z}}$-valued function on $\mathcal{H}_p(\mathbb{Q}_{p^2})\times\mathbb{P}_1(\mathbb{Q})\times\mathbb{P}_1(\mathbb{Q})$, denoted*

$$(\tau,r,s)\mapsto \not\int_r^\tau\int^s n\omega_f, \tag{24}$$

and satisfying

1. $$\not\int_r^\tau\int^s n\omega_f \times \not\int_s^\tau\int^t n\omega_f = \not\int_r^\tau\int^t n\omega_f, \qquad \text{for all } r,s,t\in\mathbb{P}_1(\mathbb{Q});$$

2. $$\not\int_r^{\tau_2}\int^s n\omega_f \div \not\int_r^{\tau_1}\int^s n\omega_f = \not\int_{\tau_1}^{\tau_2}\int_r^s n\omega_f, \qquad \text{for all } \tau_1,\tau_2\in\mathcal{H}_p(\mathbb{Q}_{p^2}).$$

3. $$\not\int_{\gamma r}^{\gamma\tau}\int^{\gamma s} n\omega_f = \not\int_r^\tau\int^s n\omega_f, \qquad \text{for all } \gamma\in\Gamma.$$

This function is called the *indefinite integral* attached to f. Its uniqueness is not hard to establish. It is the existence which is more subtle: it implies the "exceptional zero conjecture" of [MTT] and the tame refinement of this conjecture that is explored in [Ds], but is in fact a bit stronger than these conjectures.

Following [DG], we now explain how (24) can be calculated in practice, assuming that it exists, in the special case where the p-adic endpoints τ_1 and τ_2 belong to $\mathcal{H}_p^0$. Firstly, the continued fraction trick discussed in Section 1.3, using the additivity property 1 of Conjecture 1.6, reduces the evaluation of (24) to the case where r and s are adjacent elements of $\mathbb{P}_1(\mathbb{Q})$. By property 3 of Conjecture 1.6, it is enough to evaluate expressions of the form

$$\not\int_0^\tau\int^\infty n\omega_f, \qquad \text{with } \tau\in\mathcal{H}_p^0.$$

The following manipulation reduces these expressions to the the double integral whose calculation was already discussed in Section 1.3.

$$\begin{aligned}\not\int_0^\tau\int^\infty n\omega_f &= \not\int_0^\tau\int^1 n\omega_f \times \not\int_1^\tau\int^\infty n\omega_f \\ &= \not\int_\infty^{-1/\tau}\int^{-1} n\omega_f \times \not\int_0^{\tau-1}\int^\infty n\omega_f = \not\int_{1-1/\tau}^{\tau-1}\int_0^\infty n\omega_f.\end{aligned}$$

1.5 Definition of P_τ

Let $\tilde{\Gamma}^+$ denote the group of matrices in $\mathbf{PGL}_2(\mathbb{Z}[1/p])$ with positive determinant. We are now ready to define the map

$$\tau\mapsto P_\tau, \qquad \Gamma\backslash(\mathcal{H}_p\cap K)\longrightarrow E(K_p)$$

underlying the Stark-Heegner point construction. The following lemma shows that in defining this map, one can restrict one's attention to the $\tau\in\mathcal{H}_p^0$.

Lemma 1.7. *The inclusion $\mathcal{H}_p^0 \subset \mathcal{H}_p$ induces a bijection*

$$\mathbf{PSL}_2(\mathbb{Z})\backslash(\mathcal{H}_p^0 \cap K) \longrightarrow \tilde{\Gamma}^+\backslash(\mathcal{H}_p \cap K). \tag{25}$$

Proof. The injectivity of the map (25) follows from the fact that the subgroup of $\tilde{\Gamma}^+$ which preserves $\mathcal{H}_p^0$ is $\mathbf{PSL}_2(\mathbb{Z})$, while the surjectivity is a consequence of the fact that any element of $\mathcal{H}_p(\mathbb{Q}_{p^2})$ is $\tilde{\Gamma}^+$-equivalent to an element in $\mathcal{H}_p^0$. Both of these facts are elementary. □

Since it is more convenient from a computational point of view to work with $\tau \in \mathcal{H}_p^0$, given any $\tau \in \mathcal{H}_p \cap K$, we will assume, after replacing it by an appropriate $\tilde{\Gamma}^+$-translate, that τ belongs to this subset of $\mathcal{H}_p$.

Let F_τ be the unique primitive integral binary quadratic form $F_\tau(x,y) = Ax^2+Bxy+Cy^2$ satisfying

$$F_\tau(\tau, 1) = 0, \quad A > 0.$$

The discriminant of this quadratic form is called the *discriminant of* τ. If D is a fixed (not necessarily fundamental) positive discriminant and $K = \mathbb{Q}(\sqrt{D})$ is the associated real quadratic field, then the set $\mathcal{H}_p^D$ of $\tau \in \mathcal{H}_p^0$ of discriminant D is stable under the action of $\mathbf{PSL}_2(\mathbb{Z})$, and

$$\mathcal{H}_p^0 \cap K = \bigcup_D \mathcal{H}_p^D,$$

where the union is taken over all discriminants of orders in K of conductor prime to p. Let $\mathcal{O}_D$ denote the order of K of discriminant D, defined by

$$\mathcal{O}_D := \begin{cases} \mathbb{Z}[\frac{\sqrt{D}}{2}] & \text{if } D \equiv 0 \pmod 4; \\ \mathbb{Z}[\frac{1+\sqrt{D}}{2}] & \text{if } D \equiv 1 \pmod 4. \end{cases}$$

The group $\mathcal{O}_D^\times$ and its subgroup $(\mathcal{O}_D)_1^\times$ of elements of norm one are free of rank one modulo torsion. Let u_D be a generator for $\mathcal{O}_D^\times$ in the even case, and a generator for $(\mathcal{O}_D)_1^\times$ in the odd case, and write

$$u_D = u + v\sqrt{D}, \quad \text{with } u, v \in \mathbb{Q}.$$

The matrix

$$\gamma_\tau := \begin{pmatrix} u + vB & -2vC \\ 2vA & u - vB \end{pmatrix} \in \mathbf{PGL}_2(\mathbb{Z})$$

fixes τ under Möbius transformations, and the properties of the indefinite integral spelled out in Conjecture 1.6 imply that the period

$$\tilde{P}_\tau := \int\!\!\!\!\!\!\diagup_r^\tau \int_r^{\gamma_\tau r} n\omega_f \in K_p^\times/q^{\mathbb{Z}}$$

does not depend on the choice of the base point $r \in \mathbb{P}_1(\mathbb{Q})$.

Definition 1.8. The *Stark-Heegner point* P_τ attached to τ is the image of $\tilde{P}_\tau$ in $E(K_p)$ by the Tate uniformization attached to E.

We now make a precise conjecture about the fields of definition of the Stark-Heegner points. Let $\mathrm{Pic}(\mathcal{O}_D)$ denote the Picard group of rank one projective modules over $\mathcal{O}_D$, and let $\mathrm{Pic}^+(\mathcal{O}_D)$ denote the group of oriented modules over $\mathcal{O}_D$. Class field theory identifies these groups with the Galois groups of certain abelian extensions of K, via the Artin map:

$$\mathrm{rec} : \mathrm{Pic}(\mathcal{O}_D) \longrightarrow \mathrm{Gal}(H_D/K), \quad \mathrm{Pic}^+(\mathcal{O}_D) \longrightarrow \mathrm{Gal}(H_D^+/K).$$

The extensions H_D and H_D^+ are called the *ring class field* and *narrow ring class field* attached to D respectively. The extension H_D is totally real, while H_D^+ is an extension of H_D whose Galois group is generated by complex conjugation. (Therefore it is of degree either 1 or 2, the latter case occuring if and only if the fundamental unit u_D of $\mathcal{O}_D^\times$ has norm 1.)

Let $h = [H_D : K]$ be the class number of D, i.e., the number of $\mathbf{SL}_2(\mathbb{Z})$-equivalence classes of primitive binary quadratic forms of discriminant D. Let $F_1, \ldots, F_h$ denote representatives for these quadratic forms, let $\tau_1, \ldots, \tau_h$ denote the corresponding elements of $\mathcal{H}_p^D/\mathbf{PSL}_2(\mathbb{Z})$, and let $P_j := P_{\tau_j}$ be the associated Stark-Heegner points. The following conjecture is a slightly more concrete reformulation of Conjecture 5.3 of [Dar1].

Conjecture 1.9. *1. In the even case, the h points $P_1, \ldots, P_h$ belong to $E(H_D)$ and the natural action of $\mathrm{Gal}(H_D/K)$ preserves this collection of points.*

2. In the odd case, the $2h$ points $\pm P_1, \ldots, \pm P_h$ belong to $E(H_D^+)$ and the action of $\mathrm{Gal}(H_D^+/H_D)$ preserves this collection of points. Furthermore, if τ_∞ denotes complex conjugation,

$$\tau_\infty P_j = -P_j. \tag{26}$$

(In particular the P_j are of order 1 or 2 when $H_D^+ = H_D$.)

Conjecture 1.9 is the statement which we have attempted to verify numerically.

1.6 Recognizing p-adic numbers as rational numbers

The ideas developed in the previous sections allow us to compute the points

$$P_j = (x_j, y_j), \quad j = 1, \ldots, h,$$

of Conjecture 1.9 to a p-adic accuracy of p^{-M} in time which is polynomial in M.

In testing Conjecture 1.9, two types of experiment are typically performed, for a given E and D.

1. *Trace computations.* In the even case, attempt to recognize the local point

$$P(D) := P_1 + \cdots + P_h \tag{27}$$

as a global point in $E(K)$.

2. *Class field computations.* In this type of experiment, which is most interesting when H_D is not abelian over $\mathbb{Q}$, the polynomial

$$h_D(t) := \prod_{j=1}^{h}(t - x_j) \tag{28}$$

with coefficients in K_p is tentatively identified as a polynomial with coefficients in K. It can then be checked whether the resulting polynomial in $K[t]$ has H_D as splitting field, and whether the roots of this polynomial are the x-coordinates of global points belonging to $E(H_D)$ in the even case, and to $E(H_D^+)$ in the odd case.

In both types of experiment, it is crucial to be able to efficiently recognize a rational number a given a p-adic approximation α of it satisfying

$$|a/\alpha - 1| \leq p^{-M}$$

for some M. In the archimedean setting where α is a real instead of a p-adic number, this can be dealt with using the *continued fraction* expansion of α (cf. [Co], Sec. 1.3.4 for example). While "p-adic continued fractions" have been proposed in the literature and used to recognize rational numbers of small height (cf. Section 3.2 of [Ru], for example), such methods do not appear to work in general and their range of applicability is not well understod. It therefore seems preferable to adopt a more robust approach based on lattice reduction in the p-adic setting. Firstly, write

$$a := p^e\beta,$$

where β is an element of $\mathbb{Z}_p^\times$ which is known to an accuracy of p^{-M}. (I.e., only the image $\bar{\beta}$ of β in $(\mathbb{Z}/p^M\mathbb{Z})^\times$ is given.) The problem of recognizing β as a rational number $b = r/s$ is tackled by letting (r, s) be a shortest vector in the lattice $L_{\beta,M} \subset \mathbb{Z}^2$ defined by

$$L_{\beta,M} := \{(x, y) \text{ such that } p^M \text{ divides } x - \beta y.\}$$

spanned by the vectors $(\tilde{\beta}, 1)$ and $(p^M, 0)$, where $\tilde{\beta} \in \mathbb{Z}$ is any representative of the congruence class $\bar{\beta}$. Note that $L_{\beta,M}$ contains at most one *primitive* vector (r, s) (up to sign) satisfying

$$|r|, |s| < \frac{1}{\sqrt{2}}p^{M/2}.$$

If the coordinates of a shortest vector $v_0 = (r, s)$ in $L_{\beta,M}$ are smaller than this bound by at least a few orders of magnitude, then one can be reasonably confident that $\beta = r/s$, although of course the program can never actually *prove* such an equality, β being only given with finite accuracy.

Finding the shortest vector in a rank two lattice of discriminant p^M is a task which can be handled efficiently via the LLL algorithm (cf. [Co], Sec. 2.6 and 2.7.3). This procedure is so efficient that recognizing p-adic numbers as rational numbers takes up a negligible part of the Stark-Heegner point calculations.

2 Computing the moments of Mazur's measure

Thanks to the previous section, the problem of efficiently computing Stark-Heegner points has been reduced to that of calculating the data

$$\omega(a,j) = \int_{a+p\mathbb{Z}_p} (t-a)^j d\mu_f(t) \pmod{p^{M''}}, \qquad 0 \leq a \leq p-1, \quad j = 0,1,\ldots,M', \quad (29)$$

attached to μ_f, in *polynomial time*.

2.1 Overconvergent modular symbols

The algorithm for doing this rests on the notion of *overconvergent modular symbols* which we now briefly recall, following [PS1], [PS2]. Let $\mathcal{D}(\mathbb{Z}_p)$ be the space of (locally analytic) $\mathbb{Q}_p$-valued distributions on $\mathbb{Z}_p$, equipped with the structure of a right $\Sigma_0(p)$-module via the rule

$$(\mu\big|\gamma)(h(t)) = \mu(h(t\gamma)).$$

Definition 2.1. An *overconvergent modular symbol* of level p is a $\mathcal{D}(\mathbb{Z}_p)$-valued modular symbol on $\Gamma_0(p)$.

The space $\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p))$ of overconvergent modular symbols is an (infinite-dimensional) p-adic Frechet space equipped with a Hecke-equivariant map

$$\rho : \mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p)) \longrightarrow \mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Q}_p)$$

to the space of $\mathbb{Q}_p$-valued modular symbols by taking "total measure", i.e., by setting for $\Phi \in \mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p))$,

$$\rho(\Phi)\{r \to s\} := \Phi\{r \to s\}(\mathbb{Z}_p) = \int_{\mathbb{Z}_p} d\Phi\{r \to s\}(t).$$

The map ρ will be referred to as the *specialization* map.

If X is any $\mathbb{C}_p$-vector space equipped with an action of a linear operator U_p, let $X^{(<h)}$ denote the subspace of X on which U_p acts with slope less than h, i.e., the direct sum of all pseudo-eigenspaces for U_p whose associated eigenvalue λ satisfies $\mathrm{ord}_p(\lambda) < h$. The following proposition gives control on the subspace of overconvergent symbols of slope strictly less than 1.

Theorem 2.2 (Stevens). *The Hecke-equivariant map*

$$\rho : \mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p))^{(<1)} \longrightarrow \mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Q}_p)^{(<1)}$$

is an isomorphism.

Proof. See [St, Theorem 7.1]. □

The modular symbol I_f defined in Proposition 1.2 and equation (7) after it can be viewed as an element of $\mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Q}_p)$, where $\mathbb{Q}_p$ is endowed with the trivial $\Sigma_0(p)$-action. Since f is an eigenform, I_f is a Hecke-eigensymbol with the same eigenvalues as f; that is,

$$I_f\big|T_\ell = a_\ell I_f, \quad \text{for all } \ell \neq p, \qquad \text{and } I_f\big|U_p = a_p I_f.$$

In particular, since $a_p = \pm 1$, I_f is an eigensymbol of slope zero. One obtains the following corollary of Theorem 2.2.

Corollary 2.3. *The symbol $I_f \in \mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Q}_p)$ lifts uniquely to a U_p-eigensymbol $\Phi_f \in \mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p))$, satisfying*

$$\rho(\Phi_f) = I_f, \quad \Phi_f\big|U_p = a_p\Phi_f.$$

The following proposition relates the locally analytic distributions $\Phi_f\{r \to s\}$ on $\mathbb{Z}_p$ to the measures $\mu_f\{r \to s\}$ on $\mathbb{P}_1(\mathbb{Q}_p)$ defined in Proposition 1.3.

Proposition 2.4. *For all $r, s \in \mathbb{P}_1(\mathbb{Q})$ and all locally analytic g on $\mathbb{Z}_p$,*

$$\int_{\mathbb{Z}_p} g(t)d\Phi_f\{r \to s\}(t) = \int_{\mathbb{Z}_p} g(-t)d\mu_f\{r \to s\}(t).$$

Proof. Let $\mu_f^\#\{r \to s\} : \mathbb{P}_1(\mathbb{Q}) \times \mathbb{P}_1(\mathbb{Q}) \longrightarrow \mathcal{D}(\mathbb{Z}_p)$ be defined by the rule

$$\mu_f^\#\{r \to s\}(g(t)) := \int_{\mathbb{Z}_p} g(-t)d\mu_f\{r \to s\}(t).$$

A direct calculation shows that $\mu_f^\#$ belongs to $\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p))$, and that

$$\rho(\mu_f^\#) = I_f, \qquad \mu_f^\#\big|U_p = a_p\mu_f^\#.$$

It follows from Corollary 2.3 that $\mu_f^\# = \Phi_f$, as was to be shown. □

Corollary 2.5. $\mu_f = w_\infty\Phi_f\{0 \to \infty\}$.

Proof. This follows from the transformation property for I_f under the matrix $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ given in equation (6). □

Thanks to this corollary, the data (29) is equivalent to the corresponding data for the distribution $\Phi_f\{0 \to \infty\}$. In order to lighten the notations, let us ignore the sign discrepancy between μ_f and $\Phi_f\{0 \to \infty\}$ that arises in the odd case and simply write

$$\mu_f := \Phi_f\{0 \to \infty\}$$

from now on.

Computing (29) is clearly equivalent to computing

$$\int_{a+p\mathbb{Z}_p} t^j \, d\mu_f(t) \pmod{p^{M''}}, \qquad 0 \leq a \leq p-1, \quad j = 0, 1, \ldots, M'.$$

To extract this data from Φ_f, note that

$$\Phi_f = \frac{1}{a_p}\Phi_f\big|U_p,$$

and therefore

$$\mu_f = \Phi_f\{0 \to \infty\} = \frac{1}{a_p}\sum_{a=0}^{p-1}\Phi_f\{a/p \to \infty\}\big|\left(\begin{smallmatrix}1 & a\\ 0 & p\end{smallmatrix}\right).$$

The matrix $\left(\begin{smallmatrix}1 & a\\ 0 & p\end{smallmatrix}\right)$ sends $\mathcal{D}(\mathbb{Z}_p)$ to the space of distributions supported on $a + p\mathbb{Z}_p$, so that

$$\int_{a+p\mathbb{Z}_p} t^j d\mu_f(t) = \frac{1}{a_p}(\Phi_f\{a/p \to \infty\}\big|\left(\begin{smallmatrix}1 & a\\ 0 & p\end{smallmatrix}\right))(t^j) \tag{30}$$

$$= \frac{1}{a_p}\Phi_f\{a/p \to \infty\}((a+pt)^j) \tag{31}$$

$$= \frac{1}{a_p}\sum_{r=0}^{j}\binom{j}{r}a^{j-r}p^r\Phi_f\{a/p \to \infty\}(t^r). \tag{32}$$

Note the factor of p^r occuring in the above formula, which implies that to compute (29), it suffices to compute

$$\Phi_f\{a/p \to \infty\}(t^r) \pmod{p^{M''-r}}, \qquad r = 0, 1, \ldots, M'.$$

In what follows, we will explain how to compute this data, with $\{a/p \to \infty\}$ replaced by an arbitrary path $\{r \to s\}$.

2.2 Iterating U_p

The symbol Φ_f is efficiently computed by realizing it as the limit of a sequence of symbols obtained by repeatedly applying U_p to an initial approximate solution Φ.

Let $\mathcal{D}_0(\mathbb{Z}_p)$ be the subspace of $\mathcal{D}(\mathbb{Z}_p)$ consisting of distributions all of whose moments are integral; that is,

$$\mathcal{D}_0(\mathbb{Z}_p) = \{\mu \in \mathcal{D}(\mathbb{Z}_p) \;\big|\; \mu(t^j) \in \mathbb{Z}_p \text{ for all } j \geq 0\}.$$

Proposition 2.6. *Let Φ be any element of $\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}_0(\mathbb{Z}_p))$ satisfying $\rho(\Phi) = I_f$. Then*

$$a_p^{-n}\Phi\big|U_p^n - \Phi_f \text{ belongs to } p^n\,\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}_0(\mathbb{Z}_p)).$$

In particular, the sequence $\{a_p^{-n}\Phi\big|U_p^n\}$ converges p-adically to Φ_f.

Proof. A direct calculation (see [PS2]) reveals that if $\Psi \in \mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}_0(\mathbb{Z}_p))$ is in the kernel of the specialization map ρ, then

$$\Psi\big|U_p \text{ belongs to } p\,\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}_0(\mathbb{Z}_p)).$$

Since the symbol $\Phi - \Phi_f$ lies in the kernel of specialization and since $a_p^{-1}\Phi_f\big|U_p = \Phi_f$, it follows that

$$a_p^{-n}\Phi\big|U_p^n - \Phi_f = a_p^{-n}(\Phi - \Phi_f)\big|U_p^n \text{ belongs to } p^n \operatorname{Symb}_{\Gamma_0(p)}(\mathcal{D}_0(\mathbb{Z}_p)),$$

as was to be shown. □

Proposition 2.6 forms the basis of the following polynomial time algorithm to compute the data (29).

Step 1 Find *any* symbol Φ in $\operatorname{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p))$ lifting I_f.

Step 2 Apply the operator $a_p^{-1}U_p$ to Φ repeatedly. Each application produces an improved lift of I_f that is p-adically closer to Φ_f. In fact, each iteration introduces at least one extra digit of p-adic accuracy.

Step 3 Evaluate such a lift at the paths from a/p to ∞ to yield approximations to the moments of the original measure μ_f.

We discuss each step in turn.

2.3 Lifting modular symbols

The first step of the algorithm is to find some symbol Φ of $\operatorname{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p))$ lifting I_f. This is achieved by means of the following explicit presentation of the space of V-valued modular symbols on $\Gamma_0(p)$.

Theorem 2.7. *Assume that the image of $\Gamma_0(p)$ in $\mathbf{PSL}_2(\mathbb{Z})$ is torsion-free. Then for some $m \geq 1$, there exist*

$$r_i, s_i \in \mathbb{P}_1(\mathbb{Q}), \quad \gamma_i \in \mathbf{SL}_2(\mathbb{Z}), \qquad i = 1, \ldots, m,$$

such that:

1. *Any modular symbol $\varphi \in \operatorname{Symb}_{\Gamma_0(p)}(V)$ satisfies the relation*

$$\varphi\{0 \to \infty\}\big|\left(\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right) - 1\right) = \sum_{j=1}^{m} \varphi\{r_i \to s_i\}\big|(\gamma_i - 1).$$

2. *Conversely, given elements $v_1, \ldots, v_m, v_\infty \in V$ satisfying*

$$v_\infty\big|\left(\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right) - 1\right) = \sum_{j=1}^{m} v_i\big|(\gamma_i - 1),$$

there is a unique modular symbol $\varphi \in \operatorname{Symb}_{\Gamma_0(p)}(V)$ such that

$$\varphi\{r_i \to s_i\} = v_i \text{ and } \varphi\{0 \to \infty\} = v_\infty.$$

Proof. See [PS2]. □

Remark 2.8. The proof of Theorem 2.7 is constructive, and gives an efficient way of producing the r_i, s_i and γ_i. There is also a more general version that does not assume that the image of $\Gamma_0(p)$ in $\mathbf{PSL}_2(\mathbb{Z})$ is torsion-free. Instead of having a list of paths satisfying one relation, two or three additional relations may appear depending on the order of the torsion subgroup of $\Gamma_0(p)$. For the remainder of this section, it will be assumed for simplicity that we are in the torsion-free case. However, all of the constructions of this section can be made to work in the presence of torsion (see [PS2] for details), and it is this more general version that has been implemented in the `shp` package.

To produce a $\mathcal{D}(\mathbb{Z}_p)$-valued modular symbol from Theorem 2.7, it is necessary to solve the "difference equation"

$$\mu\big|\Delta = \nu, \tag{33}$$

where $\Delta = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) - 1$ and $\mu, \nu \in \mathcal{D}(\mathbb{Z}_p)$. Since $\rho(\mu\big|\Delta) = 0$ for any $\mu \in \mathcal{D}(\mathbb{Z}_p)$, a necessary condition to solve this equation is for ν to have total measure zero. Unfortunately, in $\mathcal{D}(\mathbb{Z}_p)$ this condition is not also sufficient. To fix this problem, we will pass to a larger space of distributions where this condition does in fact become sufficient.

For r a real number greater than 1, let $A[\mathbb{Z}_p, r]$ denote the space of power series with coefficients in $\mathbb{Q}_p$ that are convergent on the disc of radius r in $\mathbb{C}_p$ around 0. Let $D[\mathbb{Z}_p, r]$ denote the (continuous) $\mathbb{Q}_p$-dual of this space. Finally, set $\mathcal{D}^\dagger := \operatorname*{inj\,lim}_{r>1} \mathcal{D}[\mathbb{Z}_p, r]$.

The span of the functions $\{t^j\}_{j=0}^\infty$ is dense in the space of locally analytic functions on $\mathbb{Z}_p$ and in $A[\mathbb{Z}_p, r]$ for any r. Thus, a distribution μ in $\mathcal{D}(\mathbb{Z}_p)$ or in $\mathcal{D}^\dagger$ is uniquely determined by its sequence of moments $\{\mu(t^j)\}_{j=0}^\infty$. One advantage to working in $\mathcal{D}^\dagger$ is that there is a simple criterion to test when a sequence of elements in $\mathbb{Q}_p$ actually arises as the moments of $\mu \in \mathcal{D}^\dagger$.

Proposition 2.9. *Let $\{\theta_n\}$ be a sequence of elements of $\mathbb{Q}_p$ such that for every $r > 1$,*

$$\big|\theta_n\big|_p \text{ is } o(r^n) \text{ as } n \to \infty.$$

Then there exists a unique distribution $\mu \in \mathcal{D}^\dagger$ such that $\mu(t^n) = \theta_n$.

We are now in a position to present a solution to the difference equation in $\mathcal{D}^\dagger$.

Proposition 2.10. *If $\nu \in \mathcal{D}^\dagger$ is a distribution of total measure zero, then there is a unique solution $\mu \in \mathcal{D}^\dagger$ of equation (33). Moreover,*

$$\mu(t^n) = \sum_{j=0}^{n} \frac{\nu(t^{j+1})\binom{n}{j} B_{n-j}}{j+1},$$

where B_k is the k-th Bernoulli number.

Proof. See [PS2] □

Remark 2.11. The presence of denominators in the formula of Proposition 2.10 is the reason why (33) may fail to have solutions in $\mathcal{D}(\mathbb{Z}_p)$. This is because any $\mu \in \mathcal{D}(\mathbb{Z}_p)$ has bounded moments, while the distribution solving the difference equation need not share this property.

Theorem 2.7 and Proposition 2.10 allow a $\mathbb{Q}_p$-valued modular symbol to be explicitly lifted to an overconvergent modular symbol. Namely, if φ belongs to $\mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Q}_p)$, set

$$\nu_i = \varphi\{r_i \to s_i\} \cdot \delta_0 \in \mathcal{D}^\dagger,$$

where δ_0 is the Dirac distribution based at zero. Then set

$$\nu = \sum_{i=0}^{m} \nu_i \big| (\gamma_i - 1),$$

which has total measure zero. By Proposition 2.10, there exists a unique distribution $\nu_\infty \in \mathcal{D}^\dagger$ such that

$$\nu_\infty \big| \Delta = \nu,$$

and, by Theorem 2.7, there exists a unique symbol $\Phi' \in \mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}^\dagger)$ satisfying

$$\Phi'\{r_i \to s_i\} = \nu_i \text{ and } \Phi'\{0 \to \infty\} = \nu_\infty.$$

Unfortunately, it is not necessarily the case that $\rho(\Phi') = \varphi$, since the fact that

$$\rho(\Phi')\{r_i \to s_i\} = \varphi(r_i \to s_i) \quad \text{ for } i = 1, \ldots, m$$

does not imply that

$$\rho(\Phi')\{0 \to \infty\} = \varphi\{0 \to \infty\}.$$

To fix this problem, it is verified in [PS2] that $\rho(\Phi') - \varphi$ is an *Eisenstein* modular symbol. Thus, choosing any prime ℓ such that $a_\ell \neq \ell + 1$ (which is possible since f is a cusp form; in fact, the choice $\ell = 2$ would always do), the symbol

$$\Phi := (a_\ell - (\ell + 1))^{-1} \Phi' \big| (T_\ell - (\ell + 1))$$

produces the desired lift of φ.

Remark 2.12. While our goal was to find a lift of I_f to a symbol in $\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}(\mathbb{Z}_p))$, we have only described how to lift it to an element of the larger space $\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{D}^\dagger)$. However, it is easy to see that the proof of Proposition 2.6 remains valid if $\mathcal{D}(\mathbb{Z}_p)$ is replaced with $\mathcal{D}^\dagger$. Thus, no generality is lost by working in this larger space of modular symbols.

2.4 Finite approximation modules

Carrying out our algorithm in practice requires a method of approximating an overconvergent modular symbol by a finite amount of data, so that it can be stored on a computer. Theorem

2.7 represents a V-valued modular symbol φ by a finite list of elements of V. What is needed is a way to approximate $\mu \in \mathcal{D}^\dagger$.

Our current characterization of a distribution is by its sequence of moments. Of course, such a description contains an infinite amount of data in two different ways: there are infinitely many moments and each moment is a p-adic number which requires an infinite amount of information to be given to full accuracy.

Consider the subspace $\mathcal{D}_0 \subseteq \mathcal{D}^\dagger$ of distributions all of whose moments are integral. A natural attempt to approximate an element of $\mathcal{D}_0$ would be to fix an integer N and consider the first N moments of the distribution mod p^N. Unfortunately, the action of $\Sigma_0(p)$ does not preserve these approximations; that is, for $\gamma \in \Sigma_0(p)$ and $\mu \in \mathcal{D}_0$, the data

$$\mu(t^j) \pmod{p^N}, \qquad j = 0, 1, \ldots, N-1$$

does not determine the corresponding data for $\mu\big|\gamma$.

The basic problem with this approach is that the subspace of $\mathcal{D}_0$ consisting of distributions whose first N moments vanish is not preserved by $\Sigma_0(p)$. However, in [PS1], the following $\Sigma_0(p)$-stable filtration of $\mathcal{D}_0$ is introduced:

$$\mathrm{Fil}^r(\mathcal{D}_0) := \left\{ \mu \in \mathcal{D}_0 \;\middle|\; \mu(t^j) \in p^{r-j}\mathbb{Z}_p \right\}.$$

The stability of this filtration implies that

$$\mathcal{F}(N) := \mathcal{D}_0/\,\mathrm{Fil}^N(\mathcal{D}_0) \cong \left(\mathbb{Z}/p^N\mathbb{Z}\right) \times \left(\mathbb{Z}/p^{N-1}\mathbb{Z}\right) \times \cdots \times (\mathbb{Z}/p\mathbb{Z})$$

is naturally a $\Sigma_0(p)$-module. In other words, the data of

$$\mu(t^j) \pmod{p^{N-j}}, \qquad j = 0, 1, \ldots, N-1$$

determines the corresponding data for $\mu\big|\gamma$ for all $\gamma \in \Sigma_0(p)$. The module $\mathcal{F}(N)$ is referred to as the *N-th finite approximation module* attached to $\mathcal{D}_0$.

Because of the denominators that appear in solving the difference equation, we will also need to consider approximations of distributions whose moments are not all integral (or even bounded!). To do this, let

$$\mathcal{K}_0 := \{\mu \in \mathcal{D}^\dagger \;\big|\; p^j\mu(t^j) \in \mathbb{Z}_p\}.$$

Note then that

$$\mathcal{F}(N) := \mathcal{D}_0/\,\mathrm{Fil}^N(\mathcal{D}_0) \cong \mathcal{D}_0/\mathcal{D}_0 \cap p^N\mathcal{K}_0 \cong (\mathcal{D}_0 + p^N\mathcal{K}_0)/p^N\mathcal{K}_0.$$

Thus, to represent a distribution $\mu \in \mathcal{D}^\dagger$ on a computer "to accuracy N", we find the smallest integer r such that the first N moments of $p^r\mu$ are all integral. Since $p^r\mu$ belongs to $\mathcal{D}_0 + p^N\mathcal{K}_0$, it then makes sense to project this distribution to $\mathcal{F}(N)$.

The following proposition describes how to lift symbols with values in these finite approximation modules. Set $r(N) = [\log(N)/\log(p)]$.

Proposition 2.13. *If φ belongs to $p^{r(N)}\,\mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Z}/p^N\mathbb{Z})$, then there exists a symbol $\Phi \in \mathrm{Symb}_{\Gamma_0(p)}(\mathcal{F}(N))$ such that $\rho(\Phi) = \varphi$. Moreover, this lift can be explicitly described by the formulas of Proposition 2.10.*

Proof. See [PS2]. □

2.5 Computing the moments

We now describe an algorithm that computes the data of (29) in polynomial time. The time complexity of this algorithm will be analyzed in section 2.6.

Let N be the smallest integer such that

$$N - r(N) \geq M'' \geq M' + 1$$

and let $\overline{I}_f$ be the image of the symbol I_f in $\mathrm{Symb}_{\Gamma_0(p)}(\mathbb{Z}/p^N\mathbb{Z})$. By Proposition 2.13, we can explicitly form some lift

$$\overline{\Phi} \in \mathrm{Symb}_{\Gamma_0(p)}(\mathcal{F}(N))^{w_\infty}$$

of $p^{r(N)}\overline{I}_f$. Repeatedly applying the operator $a_p^{-1}U_p$ to $\overline{\Phi}$ produces a sequence of elements of $\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{F}(N))$ which stabilizes to some element $\overline{\Phi}_f$ in no more than $N+1$ iterations (by Proposition 2.6 and Remark 2.12). The element $\overline{\Phi}_f$ is precisely the image of $p^{r(N)}\Phi_f$ in $\mathrm{Symb}_{\Gamma_0(p)}(\mathcal{F}(N))$.

Next, we evaluate $\overline{\Phi}_f$ at the paths $\{a/p \to \infty\}$ for each a between 0 and $p-1$. Because of the way in which the distributions are stored, $\overline{\Phi}_f\{a/p \to \infty\}$ is just the sequence of values

$$p^{r(N)}\Phi_f\{a/p \to \infty\}(t^j) \pmod{p^{N-j}}, \quad j = 0, \ldots, N-1.$$

Canceling the extra powers of p yields the values $\Phi_f\{a/p \to \infty\}(t^j)$ modulo $p^{N-r(N)-j}$ for $0 \leq j \leq N$. Since $N - r(N) \geq M'' \geq M' + 1$, the moment $\Phi_f\{a/p \to \infty\}(t^j)$ has therefore been computed modulo $p^{M''-j}$, for $0 \leq j \leq M'$. Formula (32) can be used to relate this data to the moments $\int_{a+p\mathbb{Z}_p} t^j \, d\mu_f \bmod p^{M''}$, from which the data (29) is readily recovered.

2.6 Complexity analysis

We now analyze the running time complexity of computing (29) following the strategy described in Section 2.5.

Proposition 2.14. *The procedure described in Section 2.5 computes the moments*

$$\omega(a,j) \pmod{p^{M''}}, \qquad 0 \leq a \leq p-1, \quad j = 0, 1, \ldots, M'$$

in $O(M^3p^3 \log M \log p)$ arithmetic operations on integers of size on the order of p^M.

Proof. The most time-consuming part in this computation lies in the iteration of the U_p operator; that is, the number of arithmetic operations required to iterate U_p on a lift of I_f (measured as a function of p and M) dominates the number of operations required to carry out the other parts of the algorithm. Moreover, the most time intensive part of iterating U_p is accounted for by the right actions of elements in $\Sigma_0(p)$ on distributions. For this reason, we will simply count the number of right actions performed and then analyze the time complexity of a single right action.

We first compute the number of right actions required to apply U_p once to an $\mathcal{F}(N)$-valued modular symbol Φ. To do this, we must compute $(\Phi\big|U_p)\{r_i \to s_i\}$ for the paths

$\{r_i \to s_i\}$ of Theorem 2.7. (Note that there are on the order of p such paths.) For each i, we have

$$(\Phi\big|U_p)\{r_i \to s_i\} = \sum_{a=0}^{p-1} \Phi\{\gamma_a r_i \to \gamma_a s_i\}\big|\gamma_a,$$

where $\gamma_a = \left(\begin{smallmatrix} 1 & a \\ 0 & p \end{smallmatrix}\right)$. To compute $\Phi\{\gamma_a r_i \to \gamma_a s_i\}$, we use Manin's continued fraction algorithm to write

$$\Phi\{\gamma_a r_i \to \gamma_a s_i\} = \sum_j \Phi\{a_{ij} \to b_{ij}\},$$

where a_{ij} and b_{ij} are adjacent rational numbers. The number of terms in the above sum is on the order of $\log(p)$. By the proof of Theorem 2.7 in [PS2], for arbitrary adjacent rational numbers r and s, we have

$$\Phi\{r \to s\} = \sum_i \Phi\{r_i \to s_i\}\big|\alpha_i$$

for $\alpha_i \in \mathbb{Z}[\Gamma_0(p)]$, where each α_i is composed of a sum of no more than 2 basic elements in $\Gamma_0(p)$. Thus to apply U_p once to an $\mathcal{F}(N)$-valued modular symbol requires on the order of $p^3 \log(p)$ right actions by elements $\gamma \in \Sigma_0(p)$.

The operator U_p needs to be iterated at most $N+1$ times in order to form a $\mathcal{F}(N)$-valued U_p-eigensymbol. To compute the moments with enough accuracy to obtain the associated Stark-Heegner points to accuracy p^{-M}, we can take N to be some integer with size on the order of the size of M. Thus, to complete this part of the algorithm requires on the order of $Mp^3 \log(p)$ right actions on elements in $\mathcal{F}(N)$.

Finally, an efficient method for computing $\mu\big|\gamma$ for $\mu \in \mathcal{D}^\dagger$ is given in [PS2]. The main idea is that there is a $\Sigma_0(p)$-equivariant isomorphism between $\mathcal{D}^\dagger$ and the space of log-differentials on the complement of the unit disc in $\mathbb{C}_p$. One has that a distribution $\mu \in \mathcal{D}^\dagger$ corresponds to

$$\omega = \sum_{j=0}^{\infty} a_j z^{-j} \frac{dz}{z}$$

where $a_j = \mu(t_j)$. To act by γ on ω, one simply acts on the variable z and then expands out. To do this in $\mathcal{F}(N)$ takes $N^2 \log(N)$ arithmetic operations. Moreover, since we are working in $\mathcal{F}(N)$, we never need to work with integers greater than p^N.

Proposition 2.14 follows from this analysis. □

3 The `shp` package

The Stark-Heegner point package has been implemented in the language `Magma` and can be downloaded from the address

`http://www.math.mcgill.ca/darmon/programs/programs.html`.

After following the instructions for downloading the `shp` package, the user will have a copy of the necessary programs and data, stored in a directory called `shp_package`.

A. To compute Stark-Heegner points on any elliptic curve of prime conductor $p = N \leq 100$, (and, in particular, to repeat any of the calculations that are reported on in Section 4) the user simply needs to go into the `shp` subdirectory in `shp_package`, invoke `magma`, and type

`load shp`N`;`

from the Magma command prompt. The script file `shp`N instructs `magma` to load certain programs and data (such as a table of pre-computed moments) needed to carry out the calculations of Stark-Heegner points on the strong Weil curve of conductor $p = N$. After being prompted to enter some of the parameters for the computation, the user then has access to a number of functions, such as those allowing the computation of p-adic double integrals. All of these are described in the on-line documentation for the `shp` package. The main command that can be used to test Conjecture 1.9 is

`HP, P, hD := stark_heegner_points(E, D, Qp)`,

which takes as input the elliptic curve `E` of prime conductor p, a positive discriminant `D` satisfying $(\frac{D}{p}) = -1$, and a fixed precision p-adic field `Qp`, and returns the following data:

1. A vector `HP` of length $h = h(D)$ containing the h distinct Stark-Heegner points $P_1, \ldots, P_h$ in `E(Qp(`$\sqrt{\texttt{D}}$`))`.

2. In the even case, the point `P` $= P(D)$ of equation (27). The program attempts to recognize the x and y coordinates of $P(D)$ as elements of K, and tests whether the resulting pair corresponds to a global point in $E(K)$. If this identification is not successful, the point $P(D)$ is returned as an element of $E(\mathbb{Q}_{p^2})$ and a warning message is printed. Otherwise the global point in $E(K)$ that was found (and which is, with overwhelming likelihood, the point $P(D)$, although the program of course cannot guarantee this!) is returned. (In the odd case, the point $P(D)$ is set to be the point at infinity, a definition that is motivated by equation (26) in Conjecture 1.9.)

3. The variable `hD` is assigned the value of the degree h polynomial $h_D(t)$ of equation (28). The program attempts to identify $h_D(t)$ as a polynomial in $K[t]$, and returns such a polynomial satisfied by the x_j to the calculated degree of acuracy.

B. To work with elliptic curves of conductor > 100, or to increase the accuracy of the calculations beyond 100 p-adic digits, the user will first need to compute the sequence of moments attached to Mazur's measure for the desired elliptic curve. This can be done by going into the `moments` subdirectory in `shp_package`, invoking `magma`, and typing

`load moments.magma;`

from the Magma command prompt. The user will then be prompted to supply the parameters for the computation such as the prime p, an identifier for the elliptic curve following the conventions of the tables of Cremona if there are several isogeny classes of elliptic curves of conductor p, and the number M of digits of desired accuracy. The program produces the data (29) and stores it in an array. Typically this array of moments will be stored in a file where it can later be accessed repeatedly for different Stark-Heegner point calculations (on the same curve, but with varying real quadratic discriminants.)

4 Numerical examples

This section summarises some of the new experimental evidence for Conjecture 1.9 that was obtained with the help of the `shp` package, with special emphasis on discriminants of class number > 2 for which points defined over the relevant class field would be difficult to compute without the algorithms of this paper.

The curve $X_0(11)$. Let f be the unique normalised cusp of weight 2 on $\Gamma_0(11)$. The first $M = 100$ moments of the corresponding 11-adic measure μ_f were computed to an 11-adic accuracy of 11^{-101}, taking around 2 minutes on a fast workstation. The curve $X_0(11)$, also denoted by $11A1$ in the tables of Cremona, is described by the Weierstrass equation

$$11A : y^2 + y = x^3 - x^2 - 10x - 20,$$

and is the curve on which the Stark-Heegner points appear to be the best behaved (i.e., to be of *smallest height*).

The `shp` package was used to verify the data in Tables 1 and 2 of [DG] to 100 digits. It was also used to fill in many of the missing entries in Table 2 that the authors of [DG] were previously unable to identify as global points. For example, for the first missing entry of this table, corresponding to $D = 101$, the command `stark_heegner_points(E,101,Qp)` produces, in a few seconds, the point $P_{\frac{1+\sqrt{101}}{2}}$ with an accuracy of 11^{-100} and recognizes it as an 11-adic approximation to the global point $(x/t^2, y/t^3)$, with $t = 15711350731963510$ and

$$\begin{aligned} x &= 10816241366446925396670846851168 49, \\ y &= -19391462977749218369160989980706200 47276215775500 \\ &\quad -45034813271762519727132587561686024065704563549 3\sqrt{101}. \end{aligned}$$

The large height of this point explains the difficulties encountered by [DG] in recognizing it as a global point. Thanks to the `shp` package, the traces to K of the Stark-Heegner points were efficiently identified as global points for all values of $D \leq 200$, with the notable exception of $D = 173$, which has narrow class number one. After increasing the 11-adic accuracy from 100 to 200 digits, (necessitating a moment pre-computation that took roughly 20 minutes) the programs were also able to successfully identify the point $P_{\frac{1+\sqrt{173}}{2}}$.

The smallest positive discriminant of class number 3 in which 11 is inert is $D = 316 = 2^2 79$. The x-coordinates of the three distinct Stark-Heegner points attached to this discriminant appear to satisfy the polynomial with (relatively!) small coefficients

$$\begin{aligned} h_{316}(x) &= 72766453768745463520694728094967184x^3 \\ &\quad -71914415566181323559220215097240264940x^2 \\ &\quad +26530295357490354135744648963823312705 16x \\ &\quad -15333781783601940675857202851550615143803, \end{aligned}$$

whose splitting field is indeed the Hilbert class field of $\mathbb{Q}(\sqrt{79})$. A number of calculations of a similar sort were performed with larger discriminants, relying occasionally on an accuracy

of up to 400 significant 11-adic digits; the scope of these calculations is summarized in Tables 2 and 3 below.

The numerical investigations of Stark-Heegner points on $X_0(11)$ suggest that the logarithmic height of these points grows quickly as D increases, in contrast to what happens with some of the other curves that were considered, such as the curve $X_0(37)^+$ of conductor 37, as is illustrated by Table 1 below.

The curve $X_0(37)^+$. The modular curve $X_0(37)$ is of genus two, and its quotient $X_0^+(37)$ by the Atkin-Lehner involution is an elliptic curve, denoted by $37A$ in the tables of Cremona. Given by the equation

$$E : y^2 + y = x^3 - x,$$

it is notable for being the elliptic curve of smallest conductor for which the sign in the functional equation of $L(E, s)$ is -1, and for being unique in its $\mathbb{Q}$-isogeny class. (All the elliptic curves of prime conductor less than 37 are equal to the Eisenstein quotients of the corresponding modular curve, and in particular have non-trivial rational isogenies.) This last fact may explain why the Stark-Heegner points, which this time were computed to only 50 significant 37-adic digits, are of small height, so that they can be recognized as global points even for large discriminants with sizeable class numbers, as is illustrated in the following table.

D	h	The polynomial $h_D(x)$
257	3	$x^3 - 66x^2 - 24x + 8$
316	3	$x^3 - 4x^2 - 2x + 4$
328	4	$4x^4 - 20x^3 - 11x^2 + 8x - 1$
401	5	$81x^5 - 657x^4 + 1195x^3 - 173x^2 - 976x + 527$
473	3	$169x^3 - 120x^2 - 180x + 136$
505	4	$128881x^4 - 635475x^3 - 580801x^2 - 66795x - 1495$
520	4	$x^4 - 2904x^3 + 126x^2 + 3456x + 1296$
568	3	$x^3 - 7x^2 + 10x - 2$
577	7	$x^7 - 29x^6 + 245x^5 - 633x^4 + 515x^3 - 15x^2 - 18x - 1$
621	3	$9x^3 - 9x^2 - 24x - 4$
624	4	$x^4 - 30x^3 + 120x^2 - 126x + 9$
672	4	$9x^4 - 6x^3 - 14x^2 - 2x + 1$
680	4	$81x^4 - 340x^3 + 328x^2 - 9$
689	4	$49x^4 - 1751x^3 + 9925x^2 - 8493x - 1017$
697	6	$1600x^6 - 784x^5 - 728x^4 + 111x^3 + 53x^2 - 3x - 1$
1093	5	$256x^5 - 10160x^4 + 3569x^3 + 163x^2 - 122x + 4$
1129	9	$27889x^9 - 10266200x^8 + 71385938x^7 + 201496372x^6 + 77385436x^5$ $-83339876x^4 - 17366802x^3 + 19161226x^2 - 3925233x + 251669$
1297	11	$961x^{11} - 4035x^{10} - 3868x^9 + 19376x^8 + 13229x^7 - 27966x^6 - 21675x^5$ $+11403x^4 + 11859x^3 + 1391x^2 - 369x - 37$
1761	7	$x^7 - 84x^6 + 1294x^5 - 2406x^4 + 49x^3 + 1020x^2 + 102x - 27$
1996	5	$16x^5 - 297x^4 + 1956x^3 - 5574x^2 + 7076x - 3293$
2029	7	$256x^7 - 2288x^6 + 3409x^5 + 5568x^4 - 8444x^3 - 5150x^2 + 4515x + 1862$
2308	7	$2209x^7 - 1247663x^6 - 897885x^5 + 3874015x^4 + 3905023x^3$ $-1589845x^2 - 2559414x - 597415$
3604	12	$187333969x^{12} - 17948390102x^{11} + 33568503259x^{10} + 22354978168x^9$ $-64830216081x^8 + 11548363590x^7 + 28537911009x^6 - 14626940684x^5$ $+1207770356x^4 + 510260996x^3 - 85472644x^2 - 2324816x + 722036$

Table 1: Some values of $h_D(x)$ for $E = X_0(37)^+$.

It is interesting to contrast these calculations with what occurs for the elliptic curve $37B$ described by the Weierstrass equation

$$E : y^2 + y = x^3 + x^2 - 23x - 50,$$

whose L-function has sign 1. For example, the polynomial $h_{577}(x)$ for this curve appears to be given by

$$\begin{aligned} h_{577}(x) \;=\; & 2936231528590386481x^7 - 48339370980156932397 80x^6 \\ & +2186527328027961799999 82x^5 - 2359684636847762502980 28x^4 \\ & -8023973384386324566210757x^3 - 1193483980495390619139462x^2 \\ & +8873997051178478426466846 0x + 1361578540700670673826 71979, \end{aligned}$$

with coefficients significantly larger than those appearing on the line $D = 577$ in Table 1.

Summary of calculations for other curves. In addition to the elliptic curves of conductors 11 and 37 already discussed, our numerical experiments focussed on the following strong Weil curves, designated following the conventions of the tables of Cremona:

$$\begin{aligned}
17A &: y^2 + xy + y = x^3 - x^2 - x - 14,\\
19A &: y^2 + y = x^3 + x^2 - 9x - 15,\\
43A &: y^2 + y = x^3 + x^2,\\
53A &: y^2 + xy + y = x^3 - x^2,\\
61A &: y^2 + xy = x^3 - 2x + 1,\\
67A &: y^2 + y = x^3 - 12x - 21,\\
73A &: y^2 + xy = x^3 - x^2 + 4x - 3,\\
83A &: y^2 + xy + y = x^3 + x^2 + x,\\
389A &: y^2 + y = x^3 + x^2 - 2x.
\end{aligned}$$

Table 2 below summarises some of the calculations that were done to test Conjecture 1.9 numerically for each curve.

p	E	M	s	D	$h=1$	$h=2$	$h=4$
11	$11A$	400	1	≤ 200	33	10	1
17	$17A$	100	1	≤ 100	19	3	0
19	$19A$	100	1	≤ 100	19	3	0
37	$37A$	100	-1	≤ 200	37	8	0
37	$37B$	100	1	≤ 50	9	0	0
43	$43A$	100	-1	≤ 50	11	0	0
53	$53A$	100	-1	≤ 50	10	0	0
61	$61A$	100	-1	≤ 50	10	1	0
67	$67A$	100	1	≤ 50	11	0	0
73	$73A$	100	1	≤ 50	10	1	0
83	$83A$	100	-1	≤ 50	7	0	0
389	$389A$	20	1	≤ 50	8	1	0

Table 2: Range of numerical verifications for small D.

The fourth column of Table 2 lists the sign s in the functional equation for $L(E/\mathbb{Q}, s)$, and the fifth gives the range of positive discriminants in which p is inert for which Conjecture 1.9 was tested. The last three columns indicate the number of values of D in the tested ranges with class number 1, 2 and 4, these being the only class numbers that occured in these ranges.

Table 3 below lists, next to each curve, the first few discriminants of class number three for which p is inert, and for which Stark-Heegner point calculations, carried out to an accuracy of p^{-M}, led to a successful identification of the polynomial $h_D(t)$ and of the Stark-Heeegner points themselves as points in $E(H_D)$. Tables 4 and 5 do the same for discriminant D of class number 5 and ≥ 7 respectively.

p	E	M	D
11	$11A$	400	$316, 321, 404$
17	$17A$	100	$148, 316$
19	$19A$	100	$148, 257, 316, 469$
37	$37A$	100	See Table 1
37	$37B$	100	$257, 316, 473$
43	$43A$	100	$148, 257, 321, 469$
53	$53A$	100	$257, 316, 321, 404, 469$
61	$61A$	100	$148, 316, 404, 592, 621$
67	$67A$	100	$229, 316, 321, 404$
73	$73A$	100	$229, 321, 404, 469$
83	$83A$	100	$257, 316, 321, 404, 469, 473$
389	389	20	148

Table 3: Some numerical verifications with $h(D) = 3$

p	E	M	D
37	$37A$	100	See Table 1
37	$37B$	100	401
43	$43A$	100	1093
53	$53A$	100	$401, 817, 1093$
61	$61A$	100	$401, 817, 1393, 1429, 1604$
67	$67A$	100	$401, 817$
73	$73A$	100	817
83	$83A$	100	$1093, 1429$
389	$389A$	20	401

Table 4: Some numerical verifications with $h(D) = 5$

p	E	M	D
37	$37A$	100	See Table 1
37	$37B$	100	577
43	$43A$	100	$577, 1009, 1297^{h=11}$
53	$53A$	100	1009
61	$61A$	100	$577, 1009, 1761, 1129^{h=9}$
67	$67A$	100	577
73	$73A$	100	$577, 1009$
83	$83A$	100	$577, 1009, 1129^{h=9}, 1297^{h=11}$
389	$389A$	20	577

Table 5: Some numerical verifications with $h(D) \geq 7$

Remarks.
1. Generally speaking, the Stark-Heegner points seem to be of smaller height on the curves with $s = -1$, than on those for which the sign s is equal to 1. But they are also notably small on the curve $389A$ with $s = 1$, in spite of the larger size of this conductor.

2. The curve $E = 398A$ was singled out because it is the elliptic curve of smallest conductor whose rank over $\mathbb{Q}$ is equal to 2. For any value of D, the point $P(D)$ is expected to be a torsion point on $E(K)$, by analogy with what happens with classical Heegner points (and as should occur if, as conjectured in [Dar1] and [DG], Stark-Heegner points satisfy an analogue of the Gross-Zagier formula). All our experimental results agree with this prediction, to the calculated degree of 389-adic accuracy. Note that when $h(D) > 1$, the Stark-Heegner point calculations, carried out to 20 significant 389-adic digits, always resulted in finding a point of infinite order on $E(H_D)$, for the values of D listed in the last lines of Tables 3, 4 and 5.

3. It would have been too tedious to list the polynomial $h_D(x)$ corresponding to each entry in the rightmost columns of Tables 3, 4 and 5. However, the reader can repeat these calculations independently after downloading the `shp` package, using the script files that are included to facilitate this task. For example, experimentally verifying Conjecture 1.9 for the second entry in the penultimate line of Table 5 can be accomplished by typing the sequence of two commands

```
load shp83;
```

```
HP, P, hD := stark_heegner_points(E,577,Qp);
```

from the magma command prompt. The calculation in this case, which is fairly typical, takes about two minutes (with 20 significant 83-adic digits, a precision that apparently turns out to be sufficient in this case for identifying both P and $h_D(t)$ as global objects).

4. In [DD] it is observed that when the cusp form f is replaced by the logarithmic derivative of a modular unit—a weight two Eisenstein series—the Stark-Heegner point construction leads to a refinement of the p-adic Gross-Stark units in ring class fields of real quadratic fields. The last chapter of [Das] gives polynomial time algorithms for computing these p-units. Because the periods of Eisenstein series can be written down explicitly, the approach followed in [Das] does not rely on iteration of the U_p operator and is even more efficient that the method described in this paper.

References

[BD1] M. Bertolini and H. Darmon. *Hida families and rational points on elliptic curves.* In progress.

[BD2] M. Bertolini and H. Darmon. *The rationality of Stark-Heegner points over genus fields of real quadratic fields.* In progress.

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

[Dar1] H. Darmon. *Integration on $\mathcal{H}_p \times \mathcal{H}$ and arithmetic applications*. Ann. of Math. (2) **154** (2001), no. 3, 589–639.

[Dar2] H. Darmon. *Rational points on modular elliptic curves*. CBMS Regional Conference Series in Mathematics, **101**. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2004.

[Das] S. Dasgupta. PhD thesis, Berkeley. In progress.

[DD] H. Darmon and S. Dasgupta. *Elliptic units for real quadratic fields*. Submitted.

[DG] H. Darmon and P. Green. *Elliptic curves and class fields of real quadratic fields: algorithms and evidence*. Experiment. Math. **11** (2002), no. 1, 37–55.

[Ds] E. de Shalit. *On the p-adic periods of $X_0(p)$*. Math. Ann. **303** (1995), no. 3, 457–472.

[El] Noam D. Elkies. *Heegner point computations*. Algorithmic number theory (Ithaca, NY, 1994), 122–133, Lecture Notes in Comput. Sci., 877, Springer, Berlin, 1994.

[GS] R. Greenberg and G. Stevens. *p-adic L-functions and p-adic periods of modular forms*. Invent. Math. **111** (1993), no. 2, 407–447.

[Man] J.I. Manin. *Parabolic Points and Zeta Functions of Modular Curves.* Izv. Akad. Nauk SSSR Ser. Mat. **36** (1972), no. 1, 19–66.

[MTT] B. Mazur, J. Tate, and J. Teitelbaum. *On p-adic analogues of the conjectures of Birch and Swinnerton-Dyer*. Invent. Math. **84** (1986), no. 1, 1–48.

[PS1] R. Pollack and G. Stevens. *The "missing" p-adic L-function*, in preparation.

[PS2] R. Pollack and G. Stevens. *Explicit computations with overconvergent modular symbols*, in preparation.

[Ru] K. Rubin. *p-adic variants of the Birch and Swinnerton-Dyer conjecture for elliptic curves with complex multiplication. p*-adic monodromy and the Birch and Swinnerton-Dyer conjecture (Boston, MA, 1991), 71–80, Contemp. Math., **165**, Amer. Math. Soc., Providence, RI, 1994.

[St] G. Stevens. *Overconvergent Modular Symbols*. unpublished notes.