Canad. Math. Bull. Vol. **28** (3), 1985

MODULAR PARAMETRIZATIONS OF ELLIPTIC CURVES

BY

D. ZAGIER

Dedicated to the memory of Robert Arnold Smith

ABSTRACT. Many — conjecturally all — elliptic curves $E/\mathbb{Q}$ have a "modular parametrization," i.e. for some N there is a map φ from the modular curve $X_0(N)$ to E such that the pull-back of a holomorphic differential on E is a modular form (newform) f of weight 2 and level N. We describe an algorithm for computing the degree of φ as a branched covering, discuss the relationship of this degree to the "congruence primes" for f (the primes modulo which there are congruences between f and other newforms), and give estimates for the size of this degree as a function of N.

Let X be a modular curve, i.e. a curve of the form $\Gamma\backslash\mathfrak{H} \cup \{\text{cusps}\}$ where $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ is a Fuchsian group of the first kind acting on the complex upper half plane $\mathfrak{H}$, and suppose that we have a map φ from X to some elliptic curve E over $\mathbb{C}$. On E there is a unique (up to scalar multiplication) holomorphic differential form; its pull-back under φ has the form $2\pi i f(\tau)\, d\tau$ where $f: \mathfrak{H} \to \mathbb{C}$ is a holomorphic cusp form of weight 2 on Γ. The situation of interest in number theory is when $\Gamma = \Gamma_0(N) \subset \mathrm{PSL}_2(\mathbb{Z})$, the curve E and the map φ are defined over $\mathbb{Q}$, and f is a Hecke eigenform with Fourier coefficients in $\mathbb{Z}$. Then the theory of Eichler-Shimura implies that the Hasse-Weil zeta-function of E equals the L-function of f; conversely, the Taniyama-Weil conjecture says that for *any* elliptic curve E over $\mathbb{Q}$ there exist φ and f related to E in this way, the integer N being the conductor of E.

In this paper we will not discuss the Taniyama-Weil conjecture. Instead, we suppose that the "modular parametrization" φ is given and discuss the question of computing its topological degree as a branched covering map between Riemann surfaces. This question is less trivial than appears at first sight because the Hurwitz formula gives no information about the degree (since the Euler characteristic of E is zero). There are many examples in the literature of elliptic curves over $\mathbb{Q}$ with a known modular parametrization (for example, all known elliptic curves with conductor $\leqslant 200$; see [1]), but in general the degree of φ is not given in these papers (except in the trivial case when the map arises from an isomorphism $E \simeq \Gamma'\backslash\mathfrak{H} \cup \{\text{cusps}\}$ for some Γ' between Γ and its normalizer, in which case $\deg(\varphi) = [\Gamma' : \Gamma]$; these are the "involutory curves" of

Received by the editors September 24, 1984 and, in revised form, December 10, 1984.
This work was partially supported by a grant from the National Science Foundation.
AMS Subject Classification: 14K07, 10D12, 10D23.

[8]). We will describe a general algorithm for calculating deg (φ) and a specific formula for the case $\Gamma = \Gamma_0(N)$, N prime, and give several examples. In particular, we compute the degree of φ (assuming its existence) for the particular elliptic curve $y^2 + y = x^3 - 7x + 6$ of conductor 5077, which is the curve of smallest known conductor with Mordell-Weil group (over $\mathbb{Q}$) of rank $\geqslant 3$. The degree of this particular φ (it turns out, prettily enough, to be 1984) is of interest because it figures in the effective lower bounds for class numbers of imaginary quadratic fields obtained from the work of Goldfeld [3] and Gross-Zagier [4] (cf. [9]). We also discuss connections between the primes dividing deg (φ), when φ is associated to an eigenform f of level N, and the "congruence primes" for f in the sense of Doi and Hida (the primes modulo which there are congruences between f and other eigenforms of level N) and give upper and lower bounds for the maximal growth of deg (φ) as a function of N.

1. **The modular parametrization φ.** Suppose given a map φ as above. The elliptic curve E has the form $\mathbb{C}/\Lambda$ for some lattice $\Lambda \subset \mathbb{C}$. Let z and τ be the coordinates in $\mathbb{C}$ and $\mathfrak{H}$. Then the holomorphic differential form dz on $\mathbb{C}$, being Λ-invariant, defines a form on E, and the pull-back $\varphi^*(dz)$ has the form $2\pi i f(\tau)\, d\tau$ for some cusp form f of weight 2 on Γ, i.e. a holomorphic function $f : \mathfrak{H} \to \mathbb{C}$ satisfying

$$f(\tau) = O(\mathrm{Im}\,(\tau)^{-1}), \qquad f\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^2 f(\tau) \qquad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma.$$

Conversely, starting with any cusp form f of weight 2 on Γ and setting

$$\varphi_1(\tau) = \int_{\tau_0}^{\tau} 2\pi i f(\tau')\, d\tau' \qquad (\tau \in \mathfrak{H})$$

(any $\tau_0 \in \mathfrak{H}$), we have $\varphi_1' = 2\pi i f$ and consequently

$$\frac{d}{d\tau}\left[\varphi_1\left(\frac{a\tau + b}{c\tau + d}\right) - \varphi_1(\tau)\right] = 2\pi i \left[(c\tau + d)^{-2} f\left(\frac{a\tau + b}{c\tau + d}\right) - f(\tau)\right] = 0,$$

so

$$\varphi_1(\gamma\tau) = \varphi_1(\tau) + C(\gamma) \qquad (\gamma \in \Gamma) \tag{1}$$

for some constant $C(\gamma) \in \mathbb{C}$. If the image of the map $C : \Gamma \to \mathbb{C}$, which is clearly a homomorphism, is contained in a lattice Λ, then the map $\varphi_1 : \mathfrak{H} \to \mathbb{C}$ induces a map $\varphi : \Gamma\backslash\mathfrak{H} \to \mathbb{C}/\Lambda$ and we get a modular parametrization of an elliptic curve. In the particular case when Γ contains the element $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, e.g. for $\Gamma = \Gamma_0(N)$, the cusp form f has a Fourier expansion $f(\tau) = \Sigma_{n \geqslant 1}\, a(n) e^{2\pi i n\tau}$ and (choosing $\tau_0 = \infty$) we find

$$\varphi_1(\tau) = -2\pi i \int_{\tau}^{i\infty} f(\tau')\, d\tau' = \sum_{n=1}^{\infty} \frac{a(n)}{n} e^{2\pi i n\tau}. \tag{2}$$

Since this is a rapidly convergent series (Hecke's estimate gives $|a(n)| \leqslant 20Mn$ if $|f(\tau)| \leqslant M/\mathrm{Im}(\tau)$), we can compute $\varphi_1(\gamma\tau)$ and $\varphi_1(\tau)$ (any $\tau \in \mathfrak{H}$) in (1) to any

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

desired degree of accuracy and hence determine $C(\gamma)$ numerically and, if $C(\Gamma)$ is contained in a lattice Λ, exactly.

Given any cusp form f, the *Petersson norm* of f is defined by

$$\|f\|^2 = \int_{\Gamma\backslash\mathfrak{H}} |f(\tau)|^2 \, du\, dv \qquad (\tau = u + iv).$$

Assuming that f corresponds as above to a modular parametrization $\varphi : X \to E = \mathbb{C}/\Lambda$, we can relate this Petersson norm to deg (φ) as follows:

$$\begin{aligned}\|f\|^2 &= \frac{i}{2} \int_{\Gamma\backslash\mathfrak{H}} f(\tau)\, d\tau \wedge \overline{f(\tau)}\, \overline{d\tau} \\ &= \frac{i}{8\pi^2} \int_{\Gamma\backslash\mathfrak{H}} \varphi^*(dz) \wedge \overline{\varphi^*(dz)} \\ &= \frac{i}{8\pi^2} \cdot \deg(\varphi) \cdot \int_E dz \wedge \overline{dz} \\ &= \frac{1}{4\pi^2} \cdot \deg(\varphi) \cdot \mathrm{Vol}(E),\end{aligned}$$

where $\mathrm{Vol}(E)$ is the area of a fundamental period parallelogram for the lattice Λ. Since this area is computable numerically (the period lattice can be calculated, using the Gauss arithmetic-geometric mean, from the Weierstrass equation of E), we have reduced the problem of computing deg (φ) to that of calculating $\|f\|$.

2. **Computation of the Petersson norm**. Let f be any cusp form of weight 2 on Γ and φ_1, $C : \Gamma \to \mathbb{C}$ as in §1. In this section we show how to compute $\|f\|$ in terms of the values of $C(\gamma)$ for generators γ of Γ. Let $\mathcal{F}$ be a fundamental domain for the action of Γ on $\mathfrak{H}$ which is a hyperbolic polygon (the vertices being interior or boundary points of $\mathfrak{H}$) having a finite number of sides which are identified in pairs in $\Gamma\backslash\mathfrak{H}$. We label the vertices P_j with j in an index set $J = \mathbb{Z}/r\mathbb{Z}$ in such a way that P_{j+1} is the successor of P_j in the natural orientation. Let e_j denote the edge P_jP_{j+1}, e_{j^*} the edge with which it gets identified, and $\gamma_j \in \Gamma$ the element that identifies them. Thus $* : J \to J$ is an involution on J and the γ_j are generators of Γ satisfying $\gamma_{j^*} = \gamma_j^{-1}$. Since the identification $\gamma_j : e_j \to e_{j^*}$ is orientation-reversing, we have $\gamma_j(P_j) = P_{j^*+1}$ (cf. Figure 1). The map $T : j \mapsto j^* + 1$ from J to itself breaks up J into finitely many orbits $[j] = \{j = T^\ell j, Tj, \ldots, T^{\ell-1}j\}$ in such a way that two vertices P_j and $P_{j'}$ are identified in $\Gamma\backslash\mathfrak{H}$ iff j and j' belong to the same orbit. We pick a base-point j_0 in each orbit and define a partial order on J by $j \prec j'$ if j and j' belong to the same orbit and $j = T^\alpha j_0$, $j' = T^\beta j_0$ with $0 \leq \alpha < \beta < \ell =$ size of orbit.

THEOREM 1. *With the above notations, we have the formula*

$$\|f\|^2 = \frac{1}{8\pi^2} \sum_{\substack{j, j' \in J \\ j \prec j'}} \mathrm{Im}(C(\gamma_j)\, \overline{C(\gamma_{j'})}).$$

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

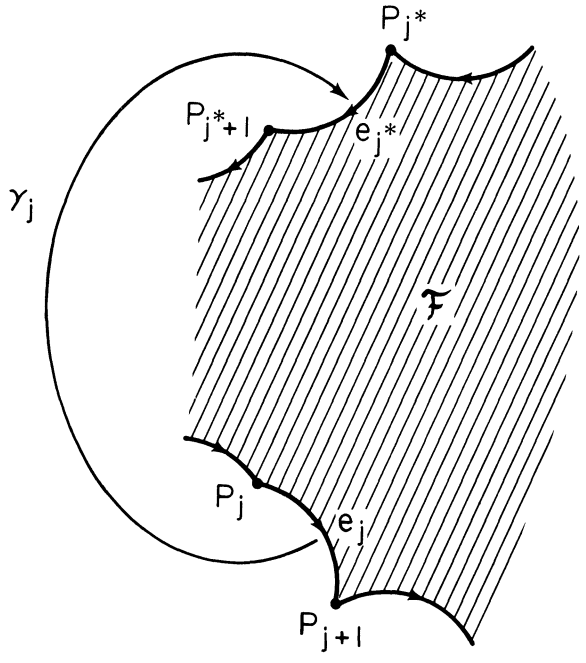

FIG. 1

In the situation when $C(\Gamma)$ is contained in a lattice Λ, so that φ_1 induces a map φ from $\Gamma\backslash\mathfrak{H}$ to $\mathbb{C}/\Lambda$, we choose an oriented basis ω_1, ω_2 of Λ (i.e. $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, $\mathrm{Im}(\omega_1\bar{\omega}_2) > 0$) and write $C(\gamma) = n_1(\gamma)\omega_1 + n_2(\gamma)\omega_2$ with n_1, n_2 homomorphisms from Γ to $\mathbb{Z}$; then Theorem 1 and the formula of §1 immediately give:

COROLLARY.

$$\deg(\varphi) = \frac{1}{2}\sum_{j<j'} [n_1(\gamma_j)n_2(\gamma_{j'}) - n_2(\gamma_j)n_1(\gamma_{j'})].$$

PROOF OF THEOREM 1. The beginning of the proof is suggested by a similar calculation for $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ in [5]. From $2\pi i f = \varphi_1'$ with φ_1 holomorphic we obtain

$$\begin{aligned}
\|f\|^2 &= \frac{i}{2}\iint_{\mathcal{F}} f(\tau)\overline{f(\tau)}\, d\tau \wedge d\bar{\tau} \\
&= \frac{1}{4\pi}\iint_{\mathcal{F}} d[\varphi_1(\tau)\overline{f(\tau)}\, d\bar{\tau}] \\
&= \frac{1}{4\pi}\int_{\partial\mathcal{F}} \varphi_1(\tau)\overline{f(\tau)}\, d\bar{\tau} \qquad \text{(Stokes' theorem)} \\
&= \frac{1}{4\pi}\sum_{j\in J}\int_{e_j} \varphi_1(\tau)\,\overline{f(\tau)}\,\overline{d\tau} \\
&= \frac{1}{8\pi}\sum_{j}{}'\left(\int_{e_j} + \int_{e_{j^*}}\right)\varphi_1(\tau)\overline{f(\tau)\, d\tau}.
\end{aligned}$$

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

Since e_{j^*} is the image of e_j under γ_j with the orientation reversed, and $f(\tau)\,d\tau$ is γ_j-invariant, we have

$$\int_{e_{j^*}} \varphi_1(\tau)\,\overline{f(\tau)\,d\tau} = -\int_{e_j} \varphi_1(\gamma_j\tau)\,\overline{f(\tau)\,d\tau},$$

so

$$\begin{aligned}\|f\|^2 &= \frac{1}{8\pi}\sum_j \int_{e_j} [\varphi_1(\tau) - \varphi_1(\gamma_j\tau)]\,\overline{f(\tau)\,d\tau} \\ &= \frac{-1}{8\pi}\sum_j C(\gamma_j)\int_{e_j}\overline{f(\tau)\,d\tau} \\ &= \frac{-i}{16\pi^2}\sum_j C(\gamma_j)[\overline{\varphi_1(P_{j+1})} - \overline{\varphi_1(P_j)}].\end{aligned}$$

In other words, by applying Stokes' theorem twice we have reduced a surface integral to a finite sum. We now simplify the last expression by replacing j by j^* in the first sum; since $\gamma_{j^*} = \gamma_j^{-1}$ and C is a homomorphism we have

$$\begin{aligned}\sum C(\gamma_j)\,\overline{\varphi_1(P_{j+1})} &= -\sum C(\gamma_j)\,\overline{\varphi_1(P_{j^*+1})} = -\sum C(\gamma_j)\,\overline{\varphi_1(\gamma_j P_j)} \\ &= -\sum C(\gamma_j)[\overline{C(\gamma_j)} + \overline{\varphi_1(P_j)}],\end{aligned}$$

and hence — since $\|f\|^2$ is real —

$$\|f\|^2 = \frac{1}{8\pi^2}\operatorname{Im}\left(\sum_j \overline{C(\gamma_j)}\,\varphi_1(P_j)\right).$$

Finally, we break up this sum into orbits under T. Let $[j_0] = \{j_0, Tj_0, \ldots, T^{\ell-1}j_0\}$ be a typical orbit with $T^\ell j_0 = j_0$ and note that $\Sigma_{j\in[j_0]} C(\gamma_j) = 0$ because $\Pi_{j\in[j_0]}\gamma_j$ fixes P_{j_0} and hence is (the identity or) an element of finite order. Hence

$$\begin{aligned}\sum_{j\in[j_0]} \overline{C(\gamma_j)}\,\varphi_1(P_j) &= \sum_{j\in[j_0]} \overline{C(\gamma_j)}\,[\varphi_1(P_j) - \varphi_1(P_{j_0})] \\ &= \sum_{j\in[j_0]} \overline{C(\gamma_j)} \sum_{j'\prec j} C(\gamma_{j'})\end{aligned}$$

since $P_j = (\Pi_{j'\prec j}\gamma_{j'})P_{j_0}$. The theorem follows.

3. **Explicit formulas when $\Gamma = \Gamma_0(N)$, N prime**. We now specialize the results of the last section to the group

$$\Gamma = \Gamma_0(N) = \left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \mathrm{PSL}_2(\mathbb{Z}) \mid c \equiv 0 \pmod N\right\}$$

with $N > 3$ prime. Let $\mathcal{F}_0$ be the standard fundamental domain for $\mathrm{PSL}_2(\mathbb{Z})$, i.e. the set of $\tau = u + iv \in \mathfrak{H}$ with $|u| \leqslant \frac{1}{2}$, $|\tau| \geqslant 1$. As a fundamental domain for Γ we could

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

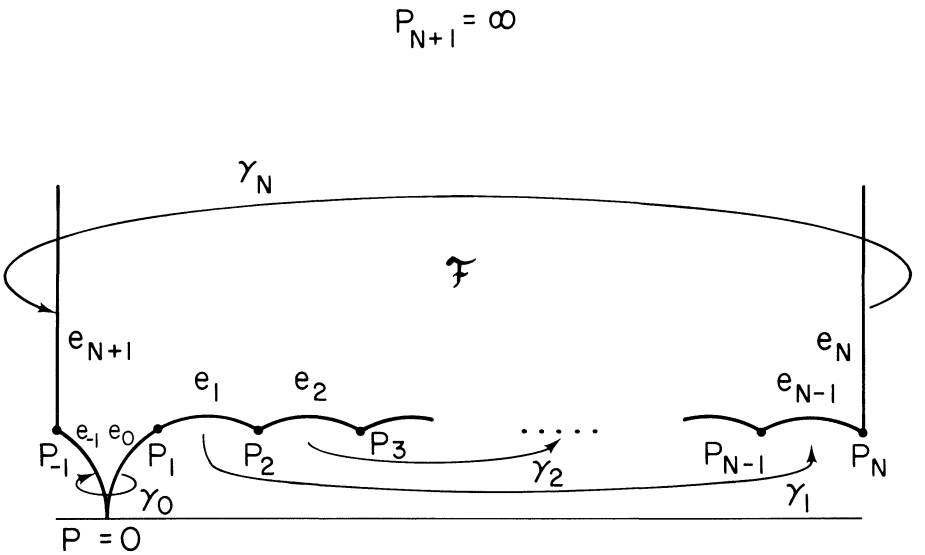

FIG. 2

take $\cup_{j=0}^{N} \alpha_j \mathcal{F}_0$, where the α_j are left coset representatives for Γ in $\mathrm{PSL}_2(\mathbb{Z})$, e.g. $\alpha_j = \begin{pmatrix} 0 & -1 \\ 1 & j \end{pmatrix}$ for $0 \leq j \leq N-1$ and $\alpha_N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. In fact we choose $\mathcal{F}$ to be the image of this fundamental domain under $w_N = N^{1/2}\begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$; this is also a fundamental domain because w_N normalizes Γ. Explicitly, $\mathcal{F} = \cup_{j=0}^{N-1} 1/N\,(\mathcal{F}_0 + j) \cup w_N \mathcal{F}_0$ (cf. Figure 2). The domain $\mathcal{F}$ has $N+3$ vertices (two of which are cusps), namely

$$P_0 = 0,\ P_j = \frac{2j - 1 + i\sqrt{3}}{2N}\ (1 \leq j \leq N),\ P_{N+1} = \infty,\ P_{N+2} = \frac{-1 + i\sqrt{3}}{2N}.$$

Thus in the notation of §2 we have $J = \mathbb{Z}/(N+3)\mathbb{Z}$. The involution $*$ on J is given by $0^* = N+2\ (= -1)$, $N^* = N+1$, and $jj^* \equiv -1 \pmod{N}$, $0 < j^* < N$ for $0 < j < N$. The corresponding identifications $\gamma_j : e_j \xrightarrow{\sim} e_{j^*}$ are

$$(3)\qquad \begin{cases} \gamma_0 = \begin{pmatrix} 1 & 0 \\ -N & 1 \end{pmatrix} : e_0 \to e_{-1}, \quad \gamma_N = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} : e_N \to e_{N+1}, \\ \gamma_j = \begin{pmatrix} j^* & -(jj^*+1)/N \\ N & -j \end{pmatrix} : e_j \to e_{j^*} \quad \text{for} \quad 0 < j < N. \end{cases}$$

The map $T : j \mapsto j^* + 1$ is described as follows:

$$0 \xrightarrow{T} 0,\ N+1 \xrightarrow{T} N+1,\ N \xrightarrow{T} N+2 \xrightarrow{T} 1 \xrightarrow{T} N,$$

$$j \xrightarrow{T} 1 - j^{-1} \xrightarrow{T} (1-j)^{-1} \xrightarrow{T} j \quad (1 < j < N),$$

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

where in the last formula the inverses are to be taken modulo N. Thus if $N \equiv 2 \pmod 3$ the map T has two fixed points 0 and $N+1$ (corresponding to the cusps) and $(N+1)/3$ orbits of length 3, while if $N \equiv 1 \pmod 3$ there are four fixed points and $(N-1)/3$ orbits of length 3, the two non-cuspidal fixed points being the roots of $j^2 - j + 1 \equiv 0 \pmod N$, $0 < j < N$ (corresponding to the elliptic fixed points of Γ). To apply Theorem 1, we need to compute $C(\gamma_j)$ for the various γ_j in (3). The elements γ_0 and γ_{N+1} are parabolic, so $C(\gamma_0) = C(\gamma_{N+1}) = 0$ and the two orbits $[0]$ and $[N+1]$ contribute nothing. From $\gamma_{-1} = \gamma_0^{-1}$, $\gamma_{N+1} = \gamma_N^{-1}$ and $\gamma_1\gamma_{N+2}\gamma_N = 1$ it follows that $C(\gamma_j) = 0$ also for the three j in the orbit $[1]$. For $1 < j < N$ we have $\gamma_j(P_j) = P_{Tj}$, so by definition $C(\gamma_j) = \varphi_1(P_{Tj}) - \varphi_1(P_j)$. If $f(\tau)$ has a Fourier development $\Sigma a(n)e^{2\pi i n\tau}$ with $a(n) \in \mathbb{R}$ and φ_1 is chosen as in (2), then we have $\varphi_1(P_j) = A(j) + iB(j)$ with

$$A(j) = \sum_{n=1}^{\infty} \frac{a(n)}{n} e^{-(\pi n\sqrt{3})/N} \cos\frac{(2j-1)n\pi}{N}, \tag{4}$$

$$B(j) = \sum_{n=1}^{\infty} \frac{a(n)}{n} e^{-(\pi n\sqrt{3})/N} \sin\frac{(2j-1)n\pi}{N}.$$

Then the contribution of an orbit $[j] = \{j, Tj, T^2j\}$ $(1 < j < N)$ in Theorem 1 is

$$\begin{aligned}
&\frac{1}{8\pi^2} \operatorname{Im}\left(C(P_j)\overline{C(P_{Tj})} + C(P_j)\overline{C(P_{T^2j})} + C(P_{Tj})\overline{C(P_{T^2j})}\right) \\
&\quad = \frac{1}{8\pi^2} \operatorname{Im}\left(\varphi_1(P_j)\overline{\varphi_1(P_{Tj})} + \varphi_1(P_{Tj})\overline{\varphi_1(P_{T^2j})} + \varphi_1(P_{T^2j})\overline{\varphi_1(P_j)}\right) \\
&\quad = \frac{1}{8\pi^2} \begin{vmatrix} 1 & A(j) & B(j) \\ 1 & A(Tj) & B(Tj) \\ 1 & A(T^2j) & B(T^2j) \end{vmatrix}.
\end{aligned}$$

We have proved:

THEOREM 2. *Let $f(\tau) = \Sigma_{n=1}^{\infty} a(n)e^{2\pi i n\tau}$, $a(n) \in \mathbb{R}$, be a cusp form of weight 2 on $\Gamma_0(N)$, $N > 3$ prime. For $j \in \mathbb{Z}/N\mathbb{Z}$ define real numbers $A(j)$, $B(j)$ by (4). Then*

$$\|f\|^2 = \frac{1}{8\pi^2} \sum_{\substack{j \in \mathbb{Z}/N\mathbb{Z} \smallsetminus \{0,1\} \\ (\mathrm{mod}\ T)}} \begin{vmatrix} 1 & A(j) & B(j) \\ 1 & A(Tj) & B(Tj) \\ 1 & A(T^2j) & B(T^2j) \end{vmatrix}, \tag{5}$$

where T is the automorphism $j \mapsto -j^{-1} + 1 \pmod N$ of order 3 on $\mathbb{Z}/N\mathbb{Z} \smallsetminus \{0, 1\}$.

Note that the cyclic group $\langle T \rangle$ of order 3 can be augmented by an involution $U : j \mapsto 1 - j$ to a group $G = \langle T, U \rangle$ of order 6 (isomorphic to the symmetric group on 3 letters) and that the determinant in (5) is invariant under U as well as T (because $A(Uj) = A(j)$, $B(Uj) = -B(j)$ and $UT = T^2U$). Hence we could also write the formula with "mod G" instead of "mod T" and $4\pi^2$ instead of $8\pi^2$.

Now suppose the periods $C(\Gamma)$ lie in a lattice $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$. Since f is real

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

$(\overline{f(-\bar{\tau})} = f(\tau))$, the period lattice of f is also real, so we can take Λ of the form

$$\Lambda = \mathbb{Z}\omega_+ + \mathbb{Z}i\omega_- \quad \text{or} \quad \Lambda = \mathbb{Z}\omega_+ + \mathbb{Z}(\tfrac{1}{2}\omega_+ + i\omega_-) \tag{6}$$

with $\omega_+, \omega_- > 0$. Then the numbers

$$\boldsymbol{a}(j) = \frac{1}{\omega_+} A(j), \quad \boldsymbol{b}(j) = \frac{1}{\omega_-} B(j) \qquad (j \in \mathbb{Z}/N\mathbb{Z}) \tag{7}$$

are integers, and we obtain (compare §1 and the Corollary to Theorem 1):

COROLLARY. *With the above notations, the degree of the map* $\varphi : X_0(N) = \Gamma_0(N)\backslash\mathfrak{H} \cup \{\text{cusps}\} \to \mathbb{C}/\Lambda$ *induced by* φ_1 *is given by*

$$\deg(\varphi) = \sum_{\substack{\mathbb{Z}/N\mathbb{Z}\smallsetminus\{0,1\} \\ \bmod G}} \begin{vmatrix} 1 & \boldsymbol{a}(j) & \boldsymbol{b}(j) \\ 1 & \boldsymbol{a}(Tj) & \boldsymbol{b}(Tj) \\ 1 & \boldsymbol{a}(T^2 j) & \boldsymbol{b}(T^2 j) \end{vmatrix}.$$

4. **Examples**. The situation of interest is when f is a (new) Hecke eigenform with coefficients $a(n) \in \mathbb{Z}$, $a(1) = 1$. Then the Eichler-Shimura theory [11, 12] implies that there is an elliptic curve E defined over $\mathbb{Q}$ such that E has good reduction at p and $|E(\mathbb{F}_p)| = p + 1 - a(p)$ for all primes $p \nmid N$; for any such E there is a map $\varphi : X_0(N) \to E$ (*Weil parametrization*) such that the pull-back under φ of a holomorphic differential on E is a multiple of $f(\tau)\, d\tau$. In general E is not unique, since any isogenous elliptic curve has the same properties, but among all Weil parametrizations (E, φ) there is a maximal one, the *strong Weil parametrization*, which dominates all the others. This curve has the property that its period lattice Λ is isomorphic to $C(\Gamma)$. In fact, a conjecture of Manin, which has been verified for all the curves we will look at (cf. [12]), says that the minimal period lattice of the strong Weil curve coincides with $C(\Gamma)$; here by "minimal period lattice" we mean the lattice of periods of the (Néron) canonical differential form $dx/(2y + a_1x + a_3)$ of a minimal model

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \qquad (a_1, \ldots, a_6 \in \mathbb{Z}) \tag{8}$$

("minimal" means that the discriminant Δ of (8) is minimal in absolute value among all equations of the form (8) for E). This lattice can be found easily from the coefficients of (8): it has the first or second form in (6) according as $\Delta > 0$ or $\Delta < 0$, and the positive real numbers $\omega_\pm$ can be computed rapidly using Gauss' arithmetic-geometric mean. (Conversely, given Λ one calculates the classical invariants $g_2(\Lambda)$ and $g_3(\Lambda)$ by the well-known Fourier expansions and obtains E in the Weierstrass form $y^2 = 4x^3 - g_2x - g_3$.) Since the series (2) converges rapidly (we have $|a(n)| \leqslant d(n)n^{1/2}$, where $d(n)$ is the number of divisors of n, so $a(n) = O(n^{1/2+\epsilon})$), we can easily compute the rational integers $\boldsymbol{a}(j)$, $\boldsymbol{b}(j)$ in (7), and then the last corollary leads to the value of $\deg(\varphi)$. The results of this computation for the Weil curves of prime conductor $N < 200$ (from [1]) and for the curve

$$y^2 + y = x^3 - 7x + 6 \tag{9}$$

TABLE 1. deg (φ) for Weil curves of small conductor

N	ϵ	a_1	a_2	a_3	a_4	a_6	Δ	ω_+	ω_-	deg (φ)
11	−	0	−1	1	−10	−20	-11^5	1.26921	1.45882	(1)
17	−	1	−1	1	−1	−14	-17^4	1.54708	1.37287	(1)
19	−	0	1	1	−9	−15	-19^3	1.35976	2.06355	(1)
37	+	0	0	1	−1	0	37	2.99346	2.45139	(2)
37	−	0	1	1	−23	−50	37^3	1.08852	1.76761	2
43	+	0	1	1	0	0	−43	5.46869	1.36318	(2)
53	+	1	−1	1	0	0	−53	4.68764	1.54059	(2)
61	+	1	0	0	−2	1	−61	6.13319	0.99721	(2)
67	−	0	1	1	−12	−21	−67	1.27377	3.02997	5
73	−	1	−1	0	4	−3	-73^2	2.36532	1.39639	3
79	+	1	1	1	−2	0	79	2.97540	2.01316	(2)
83	+	1	1	1	1	0	−83	3.37447	1.95716	(2)
89	−	1	1	0	4	5	-89^2	2.84461	1.09245	5
89	+	1	1	1	−1	0	−89	5.55263	1.14968	(2)
101	+	0	1	1	−1	−1	101	2.29512	2.72356	(2)
109	−	1	−1	0	−8	−7	−109	1.41103	2.97140	4
113	−	1	1	1	3	−4	-113^2	2.01837	1.42891	6
131	+	0	−1	1	1	0	−131	4.17161	1.48259	(2)
139	−	1	1	0	−3	−4	−139	1.73969	2.90067	6
163	+	0	0	1	−2	1	−163	5.51807	0.99371	6
179	−	0	0	1	−1	−1	−179	2.26020	2.55455	9
197	+	0	0	1	−5	4	197	2.83478	1.59772	10
5077	+	0	0	1	−7	6	5077	2.07584	1.48055	1984?

of prime conductor 5077 mentioned in the Introduction have been tabulated in Table 1. This table gives the value of N, the coefficients a_i in (8) and the corresponding discriminant Δ (which is always $\pm$ a power of N), the periods $\omega_\pm$ (which determine Λ by (6) according to $\operatorname{sgn}(\Delta)$), and the degree of the (strong) Weil parametrization φ. The number $\epsilon = \pm 1$ is defined by

$$f\left(-\frac{1}{N\tau}\right) = \epsilon N\tau^2 f(\tau) \qquad (\forall \tau \in \mathfrak{H}); \tag{10}$$

it is $+1$ if and only if the map φ factors through the projection $X_0^*(N) = X_0(N)/w_N$, and also determines the sign of the functional equation of the L-series of f (and hence conjecturally the parity of the rank of $E(\mathbb{Q})$). For N prime and <200 there is at most one eigenform with rational coefficients having given N and ϵ. The 12 degrees in parentheses in Table 1 are for the "involutory curves" of [8], where either $\epsilon = -1$ and $X_0(N) \to E$ is an isomorphism or $\epsilon = 1$ and $X_0^*(N) \to E$ is an isomorphism.

Finally, the last line of Table 1 is not proven since it has not yet been checked whether the curve E defined by (9) is a Weil curve.* However, we can compute the coefficients $a(n)$ of the L-series of E and then compute $A(j)$, $B(j)$ and hence (since $\omega_\pm$ are known) $\boldsymbol{a}(j)$, $\boldsymbol{b}(j)$ to high accuracy; if — as of course turns out to be the case

*See 'Note added in proof.'

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

— these are extremely close to rational integers, then we have strong confirmation of the Taniyama-Weil conjecture for E and a computation of the conjectural degree.

5. **Relation to congruence primes**. In this section we will show that if f is a normalized new form with integral Fourier coefficients and $\varphi : X_0(N) \to E$ is the strong Weil curve corresponding to f, then the congruence primes for f are precisely the primes dividing deg (φ). That this might be so was suggested to me by G. van der Geer and confirmed numerically by comparing the table of §4 with Table 1.1 of [2]. (To achieve agreement, one must add 2 to the list of "possible ℓ" in [2] whenever $f \in S_2^{\pm}(\Gamma_0(N))$ and $S_2^{\mp}(\Gamma_0(N))$ is non-empty, where S_2^{ϵ} denotes the space of forms satisfying (10), for the reason explained on p. 94 of [2].) The proof (of Theorem 3 below) was provided by K. Ribet, whom I would like to thank; the ideas involved are due to him and to Doi, Hida, Ohta, Mazur and others (cf. [10] and the references therein).

To give a precise formulation, let $\boldsymbol{S} = S_2(\Gamma_0(N)) \cap \mathbb{Z}[[q]]$ denote the set of cusp forms of weight 2 on $\Gamma_0(N)$ with integral Fourier coefficients and, for f as above, $\boldsymbol{L}$ the sublattice $\langle f \rangle^{\perp} \cap \boldsymbol{S}$ ($\langle f \rangle^{\perp}$ = span of eigenforms other than f = orthogonal complement of f w.r.t. the Petersson metric). Let r be the positive integer defined by one of the following equivalent conditions:

(i) r is the largest integer s.t. $\exists\, g \in \boldsymbol{L}$ with $f \equiv g \pmod{r}$;

(ii) $\{(f,h) | h \in \boldsymbol{S}\} = r^{-1}(f,f)\mathbb{Z}$, where $(\ ,\ )$ denotes the Petersson scalar product;

(iii) r is the exponent of the finite group $\boldsymbol{S}/(\mathbb{Z}f + \boldsymbol{L})$.

Thus p is a congruence prime iff $p|r$. We will show:

THEOREM 3. *Under the above assumptions,* $r = \deg(\varphi)$.

PROOF. **1**. The map φ induces a surjective map φ_* from $J = \mathrm{Jac}(X_0(N))$ to $\mathrm{Jac}(E) = E$ and a dual map $\varphi^* = (\varphi_*)^{\vee}$ from $E \simeq E^{\vee}$ to $J \simeq J^{\vee}$ (using the canonical identification of a Jacobian with its dual abelian variety); the map φ^* is injective because E is a strong Weil curve, and the composition $E \xrightarrow{\varphi^*} J \xrightarrow{\varphi_*} E$ is multiplication by $n = \deg(\varphi)$ on E. Let $A = \ker(\varphi_*)$, a codimension 1 abelian subvariety of J (it is connected because E is a strong Weil curve). Then A and $\varphi^*(E) \subset J$ intersect in a finite group which is the kernel of the map α defined by

$$\begin{array}{ccccccccc} 0 & \to & E & \xrightarrow{\varphi^*} & J & \to & A^{\vee} & \to & 0 \\ & & \cdot n \downarrow & & \| & & \uparrow \alpha & & \\ 0 & \leftarrow & E & \xleftarrow{\varphi_*} & J & \leftarrow & A & \leftarrow & 0 \end{array},$$

and it follows from this diagram that $A \cap \varphi^*(E)$ is isomorphic to the kernel of $E \xrightarrow{\cdot n} E$, i.e. to $(\mathbb{Z}/n\mathbb{Z})^2$; in particular, its exponent is n.

2. We have an isogeny $\beta : E \times A \to J$ given by $\beta(b,a) = \varphi^*(b) - a$ and an induced splitting $\mathrm{End}(J) \otimes \mathbb{Q} = \mathrm{End}(E) \otimes \mathbb{Q} \oplus \mathrm{End}(A) \otimes \mathbb{Q}$. Let $\boldsymbol{e} \in \mathrm{End}(J) \otimes \mathbb{Q}$ be the idempotent ($\boldsymbol{e}^2 = \boldsymbol{e}$) corresponding to $(1,0)$ under this splitting and m the denominator of $\boldsymbol{e}$ (= smallest integer with $m \cdot \boldsymbol{e} \in \mathrm{End}(J)$). Then $m = n$. Indeed, n divides

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

m because $m \cdot \boldsymbol{e}$ is $m \cdot 1$ on $\varphi^*(E)$ and 0 on A, so that multiplication by m is trivial on $\varphi^*(E) \cap A$, which we have just shown to have exponent n; conversely, since multiplication by n kills $A \cap \varphi^*(E) = \ker(\beta)$, the map $(\cdot n, 0)$ from $E \times A$ to itself factors through β, so $n \cdot \boldsymbol{e} \in \mathrm{End}(J)$ and $m|n$.

3. We have the Hecke algebra $\mathbb{T} = \mathbb{Z}[T_1, T_2, \ldots] \subset \mathrm{End}(J)$. By a theorem of Mazur [6], this inclusion is an isomorphism for N prime. Hence we can also define m as the denominator of $\boldsymbol{e}$ in $\mathbb{T}$.

4. The algebra $\mathbb{T}$ acts on $\boldsymbol{S}$, and a moment's thought shows that the number r defined above is simply the denominator in $\mathrm{End}(\boldsymbol{S})$ of the idempotent $\boldsymbol{e} \in \mathbb{T} \otimes \mathbb{Q} \subset \mathrm{End}(\boldsymbol{S}) \otimes \mathbb{Q}$ ($\boldsymbol{e}$ acts as the identity on f and as 0 on L). Since $\mathbb{T} \subset \mathrm{End}(\boldsymbol{S})$, it follows that $r|m$. Define a pairing $\langle\ ,\ \rangle : \boldsymbol{S} \times \mathbb{T} \to \mathbb{Z}$ by setting $\langle h, T\rangle$ equal to the first Fourier coefficient of $h|T$. Then the determinant of $\langle\ ,\ \rangle$ (w.r.t. $\mathbb{Z}$-bases of $\mathbb{T}$ and $\boldsymbol{S}$, which are both free $\mathbb{Z}$-modules of rank = genus of $X_0(N)$) is divisible by m/r, because $m \cdot \boldsymbol{e}$ is a primitive element of $\mathbb{T}$ with $h|m \cdot \boldsymbol{e} \in m/r\,\boldsymbol{S}$ for all $h \in \boldsymbol{S}$. But this determinant must be ± 1, because otherwise there would be an element $h \in \boldsymbol{S}$ and an integer $d > 0$ with $d \nmid h$ but $\langle h, T\rangle \equiv 0 \pmod{d}$ for all $T \in \mathbb{T}$, and this is impossible because $\langle h, T_j\rangle$ is the jth Fourier coefficient of h. Hence $m = r$.

Note that the assumption N prime was used only in Step 3, so we have the divisibility $r|\deg(\varphi)$ in any case. In the other direction, with the results of [10] one can see that $\deg(\varphi)$ always divides rN^i for some i.

6. **Growth of deg (φ)**. In this section we discuss upper and lower bounds for the growth of the degree of the modular parametrization of Weil curves as a function of the level.

To obtain lower bounds, we start with a modular form $f(\tau) = \Sigma\, a(n) e^{2\pi i n\tau}$ associated to a Weil curve E of conductor N and *twist* by a prime p not dividing N, i.e. pass to the new form $f^*(\tau) = \Sigma\,(n/p) a(n) e^{2\pi i n\tau}$. If Λ is the period lattice of f, then the periods of f^* lie in $\Lambda^* = 1/\delta\,\Lambda$, where $\delta = \sqrt{p}$ *if* $p \equiv 1 \pmod 4$ and $\delta = i\sqrt{p}$ *if* $p \equiv 3 \pmod 4$. This follows easily from the representation $f^*(\tau) = 1/\delta \sum_{k \pmod p} (k/p) f(\tau + k/p)$. The level of f^* is Np^2, as is easily checked. Its Petersson norm can be computed by Rankin's method, which gives the formula

$$\|f\|^2 = \frac{1}{48\pi}\,[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma_0(N)] \cdot \mathrm{Res}_{s=2} \sum_{n=1}^{\infty} \frac{|a(n)|^2}{n^s}$$

for any cusp form f of weight 2 and level N. Replacing f by f^* changes the index of $\Gamma_0(N)$ by a factor $p(p+1)$ and deletes the Euler p-factor of the Rankin L-series $\Sigma\,|a(n)|^2 n^{-s}$. Hence

$$\|f^*\|^2/\|f\|^2 = p(p+1)\left(\sum_{r=0}^{\infty} \frac{a(p^r)^2}{p^{2r}}\right)^{-1}.$$

The sum can be evaluated easily using $a(p^r) = (\alpha^{r+1} - \beta^{r+1})/(\alpha - \beta)$ with $\alpha + \beta = a(p)$, $\alpha\beta = p$, and we find that

$$\|f^*\|^2 = \|f\|^2 \cdot \frac{1}{p}(p-1)(p+1-a(p))(p+1+a(p)).$$

https://doi.org/10.4153/CMB-1985-044-8 Published online by Cambridge University Press

(Note that $p + 1 - a(p)$ is the number of points of E over the field of p elements and that $(p + 1 - a(p))(p + 1 + a(p))$ is the number of points over the field of p^2 elements.) Since the volume of Λ^* is $1/p$ times the volume of Λ, we find from the formula of §1 that the degree of the map

$$\varphi^* : X_0(Np^2) \to E^* = \mathbb{C}/\Lambda^*$$

induced by f^* is given by

$$\deg(\varphi^*) = (p - 1)(p + 1 - a(p))(p + 1 + a(p)) \deg(\varphi). \tag{11}$$

Now usually the curve E^* will be the strong Weil curve associated to f^*, but this need not always happen. For instance, if $N = 11$ then the periods of f^* will lie in a sublattice Λ' of Λ^* of index 5 and so the map φ^* factors through a map $\varphi' : X_0(11p^2) \to E' = \mathbb{C}/\Lambda'$ of degree $1/5 \deg(\varphi^*)$, which is an integer because of the congruence $a(p) \equiv p + 1 \pmod 5$. (This phenomenon was pointed out to me by G. Stevens.) However, the fact that E^* has a model over $\mathbb{Z}$ of discriminant Dp^6, where D is the discriminant of a minimal model of E, implies that Λ^* is the minimal period lattice for E^*, so if there is a non-trivial isogeny $E' \to E^*$ of curves defined over $\mathbb{Q}$ then the image of the minimal period lattice for E' cannot be contained in a multiple $k\Lambda^*$ $(k > 1)$, i.e. the isogeny is cyclic. Now the theorem of Mazur [7] on the non-existence of cyclic isogenies over $\mathbb{Q}$ of degree >163 implies that the degree of the strong Weil parametrization differs from (11) by a factor of at most 163, at least if we assume the truth of Manin's conjecture mentioned in §4. (I am indebted to B. Gross for pointing out this argument.) This proves:

THEOREM 4 (modulo Manin's conjecture). *There exist strong Weil curves of conductor $N \to \infty$ for which the degree of the Weil parametrization is $> cN^{3/2}$.*

In the other direction, by integrating $|f(\tau)|^2$ over the explicit fundamental domain of §3 and using the estimate $|a(n)| \le d(n)\sqrt{n}$ one obtains the estimate

$$\|f\|^2 \le \frac{1}{8\pi^4\sqrt{3}} N (\log^3 N + O(\log^2 N)) \qquad (N \to \infty, N \text{ prime}) \tag{12}$$

for the Petersson norm. The volume of the minimal period lattice of E is $|\Delta|^{-1/6}$ times a factor which is bounded away from zero and infinity if the j-invariant of E is bounded (if $|j| \to \infty$, this must be multiplied by $|j|^{-1/6}/\log|j|$). Hence the formula of §1 implies that $\deg(\varphi) = O(N^{1+\epsilon}|\Delta|^{1/6})$ for Weil curves of prime conductor and bounded j-invariant. In particular, if Δ is prime, so that $N = |\Delta|$, we get:

THEOREM 5 (assuming the Weil-Taniyama and Manin conjectures). *An elliptic curve over $\mathbb{Q}$ of prime discriminant N and bounded j-invariant has a modular parametrization $\varphi : X_0(N) \to E$ of degree $O(N^{7/6} \log^3 N)$.*

I do not know whether there exist elliptic curves of this sort with arbitrarily large N. This question boils down roughly to the question of representing a prime as the sum of a square and a cube of the same order of magnitude and may be amenable to an attack by sieve theory.

NOTE ADDED IN PROOF: It has now been verified by Mestre, using an idea of Serre, that $y^2 + y = x^3 - 7x + 6$ is a Weil curve of conductor 5077, and Manin's conjecture has been proved for this curve by Raynaud (it also follows from our calculations, which show that $C\colon \Gamma \to \Lambda$ is surjective). Hence the question mark in the last line of Table 1 can be removed.

REFERENCES

1. B. J. Birch and W. Kuyk, [eds.] *Modular Functions of One Variable IV*, Springer Lecture Notes 476, Berlin-Heidelberg-New York, 1975, Table I, pp. 81–113.

2. K. Doi and M. Ohta, *On some congruences between cusp forms on* $\Gamma_0(N)$, in Modular Functions of One Variable V, Springer Lecture Notes 601, Berlin-Heidelberg-New York, 1977, pp. 91–105.

3. D. Goldfeld, *The conjectures of Birch and Swinnerton-Dyer and the class numbers of quadratic fields*, Soc. Math. France, Astérisque 41–42 (1977), pp. 219–277.

4. B. Gross and D. Zagier, *Points de Heegner et dérivées de fonctions L*, C.R. Acad. Sc. Paris **297** (1983), pp. 85–87.

5. K. Haberland, *Perioden von Modulformen einer Variablen und Gruppencohomologie*, Math. Nachr. **112** (1983), pp. 245–282.

6. B. Mazur, *Modular curves and the Eisenstein ideal*, Publ. Math. I.H.E.S. **47** (1977), pp. 33–186.

7. B. Mazur, *Rational isogenies of prime degree*, Invent. Math. **44** (1978), pp. 129–162.

8. B. Mazur and H. P. F. Swinnerton-Dyer, *Arithmetic of Weil curves*, Invent. Math. **25** (1974), pp. 1–61.

9. J. Oesterlé, *Nombres de classes des corps quadratiques imaginaires*, Séminaire Bourbaki 1983–1984, Exposé 631.

10. K. Ribet, *Mod p Hecke operators and congruences between modular forms*, Invent. Math. **71** (1983), pp. 193–205.

11. G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton, 1971, Chapter VII.

12. H. P. F. Swinnerton-Dyer and B. J. Birch, *Elliptic curves and modular functions*, in Modular Functions of One Variable IV, Springer Lecture Notes 476, Berlin-Heidelberg-New York, 1975, pp. 2–32.

UNIVERSITY OF MARYLAND
COLLEGE PARK
MARYLAND, U.S.A.
AND
MAX-PLANCK-INSTITUT FÜR MATHEMATIK
BONN, FRG