MATHEMATICS OF COMPUTATION
VOLUME 57, NUMBER 196
OCTOBER 1991, PAGES 763–775

A FAMILY OF CYCLIC QUARTIC FIELDS ARISING FROM MODULAR CURVES

LAWRENCE C. WASHINGTON

ABSTRACT. We study a family of cyclic quartic fields arising from the covering of modular curves $X_1(16) \to X_0(16)$. An integral basis and a fundamental system of units are found. It is shown that a root of the quartic polynomial we construct is a translate of a cyclotomic period by an integer of the quadratic subfield of the quartic field.

Recently, O. Lecacheux [9, 10] and H. Darmon [4] showed how to use coverings of modular curves to obtain cyclic extensions of $\mathbb{Q}$. In particular, they were able to give a geometric construction of a family of cyclic quintic fields discovered by E. Lehmer [11]. The covering $X_1(N) \to X_0(N)$ (for $N > 2$) has degree $\phi(N)/2$ and group $(\mathbb{Z}/N\mathbb{Z})^\times/\{\pm 1\}$. For the quintic case, they took $N = 25$, which gave a cyclic covering of degree 10, then took the subcovering of degree 5. An important ingredient in the construction was the fact that $X_0(25)$ has genus 0. This also occurs for $N = 1, \ldots, 10, 12, 13, 16, 18$. These all give trivial or quadratic coverings except for $N = 7, 9, 13, 16, 18$. The values $N = 7, 9, 18$ yield cubic extensions and can be shown to yield the family of polynomials $X^3 - aX^2 - (a+3)X - 1$, namely the "simplest cubic fields" [17]. (However, it should be remarked that every cyclic cubic extension of $\mathbb{Q}$ comes from a polynomial of this form if a is allowed to be rational. Similarly, we are guaranteed that the quadratic extensions obtained from the coverings mentioned above correspond to polynomials of the form $X^2 - aX - 1$ with a rational.) The case $N = 13$ is treated by Lecacheux [9]. It might be suspected that the sextic fields she obtains are the same as the "simplest sextics" constructed by M.-N. Gras [6]. However, these latter fields were found by taking the fixed field of an element of order 6 in $\mathrm{PGL}_2(\mathbb{Q}) = \mathrm{Aut}(\mathbb{Q}(X))$. Therefore, they come from a covering of curves of genus 0. But $X_1(13)$ has genus 2. Alternatively, these sextic fields must be different because the discriminants of the quadratic, and cubic, subfields are different.

In the present paper, we study the case $N = 16$. As above, it might be hoped that this case would give a geometric construction of the quartic fields studied

Received March 14, 1990; revised October 22, 1990.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11R16.
Research supported in part by N.S.F.

by M.-N. Gras [5], but these again come from a covering of curves of genus 0, while $X_1(16)$ has genus 2 (and the family we construct yields a different set of discriminants). So we obtain a new family of cyclic quartics. In §1 we determine the discriminant and find an integral basis. In §2 we find a fundamental system of units. The regulators of these fields are small (which is what allows us to show that the units are fundamental), hence the class numbers will tend to be large. However, the conductors are composite, so these fields do not relate to Vandiver's conjecture, as was the case with other families of cyclic fields [3, 15, 16].

Emma Lehmer [11] made the remarkable observation that in the quadratic, cubic, and quartic "simplest" fields, the roots of the corresponding polynomials are units which are integer translates of cyclotomic periods. This led her to construct her family of quintic fields by showing that for suitable primes $p \equiv 1 \pmod 5$, the quintic cyclotomic period could be translated by an integer to obtain a unit. In [15] it was shown that a cyclic quadratic, cubic, or quartic field of prime conductor, having a unit which is an integer translate of a cyclotomic period, must be one of these "simplest" fields (in the quartic case we need the extra assumption that the unit have positive norm). Computations seem to indicate that a similar result holds in the quintic case.

For the quartic fields constructed in the present paper, it is not true that the units constructed are integer translates of periods, although this is not ruled out by the result of [15, see also 8, Proposition 3.14] since the conductor is composite. Instead, it is proved in §3 that the units are translates of periods by an integer in the quadratic subfield of the quartic field. Whether there is a relation between the modular construction of the fields and this connection with periods remains to be seen (cf. [4]).

In the last section of the paper we give the construction, via modular curves, of the family of quartic polynomials studied in the paper. However, once this family has been found, the construction is not needed to prove any of the properties. It would be interesting to use the modular construction in an intrinsic way to study the arithmetic of these fields. However, it should be mentioned that most of the calculations needed to discover the results of §1 were carried out (in Mathematica) using the q-expansions of the modular units. But once they were discovered, it was straightforward to verify them algebraically.

1. The quartic fields

Let $h \neq -1$ be an odd integer such that h, $h+2$, and h^2+4 are squarefree, and define

$$f_h(X) = X^4 - h^2X^3 - (h^3 + 2h^2 + 4h + 2)X^2 - h^2X + 1.$$

The discriminant of this polynomial is $h^2(h+2)^6(h^2+4)^3$; we determine the

discriminant of the associated field below. Let α_1 be a root of f_h. Then

$$\alpha_2 = \left(h + \frac{1}{h+2}\right) - \left(h^3 + h^2 + 3h + \frac{3}{h+2}\right)\alpha_1 + \left(-h^2 + h - 2 + \frac{3}{h+2}\right)\alpha_1^2 + \left(1 - \frac{1}{h+2}\right)\alpha_1^3$$

is also a root, and $\alpha_3 = 1/\alpha_1$ and $\alpha_4 = 1/\alpha_2$ are the remaining two roots.

It follows from the above that $K = \mathbb{Q}(\alpha_1)$ is a Galois extension of $\mathbb{Q}$. Clearly, f_h has no rational roots. Since the discriminant is never a square, it is easy to see that f_h cannot factor as the product of two quadratics (this can also be verified directly). Therefore, $K/\mathbb{Q}$ is a cyclic extension of degree 4. Writing f_h as $X^4 - cX^3 - qX^2 - cX + 1$, we find that the quadratic subfield k is $\mathbb{Q}(\sqrt{c^2 + 4q + 8}) = \mathbb{Q}(\sqrt{h^2+4})$.

Since $X^{-2}f_h(X)$ is a polynomial in $X + \frac{1}{X}$, it is easy to give closed formulas for the roots of f_h:

$$\frac{h^2 + (h+2)\sqrt{h^2+4} \pm \sqrt{2h(h+2)(h^2+4) + 2h^2(h+2)\sqrt{h^2+4}}}{4},$$

$$\frac{h^2 - (h+2)\sqrt{h^2+4} \pm \sqrt{2h(h+2)(h^2+4) - 2h^2(h+2)\sqrt{h^2+4}}}{4}.$$

Letting α_1 be the largest root, we immediately obtain the approximations (which can also be computed directly or from the q-expansions in §4)

$$\alpha_1 = h^2 + h + 1 + \frac{2}{h} - \frac{2}{h^2}\cdots, \qquad \alpha_2 = -h - 1 - \frac{1}{h} + \frac{2}{h^3}\cdots,$$

$$\alpha_3 = \frac{1}{h^2} - \frac{1}{h^3}\cdots, \qquad \alpha_4 = -\frac{1}{h} + \frac{1}{h^2}\cdots.$$

Let $\rho = \alpha_1 + i\alpha_2 - \alpha_3 - i\alpha_4$ (Lagrange resolvent). A straightforward calculation yields

$$\rho^4 = h^2(h+2)^2(h^2+4)(h-2i)^2.$$

Since $\mathbb{Q}(i, \alpha_1) = \mathbb{Q}(i, \rho)$, it is easy to determine the ramification for K: The primes dividing $h(h+2)$ are ramified in K/k but not in $k/\mathbb{Q}$. The primes dividing h^2+4 are totally ramified in $K/\mathbb{Q}$. It follows that K is the quartic field corresponding to the character $\chi = \psi_2\psi_4$, where ψ_4 is a quartic character of conductor h^2+4, and ψ_2 is the quadratic character of conductor $h(h+2)$. The conductor-discriminant formula immediately yields that the discriminant of K is

$$D = h^2(h+2)^2(h^2+4)^3.$$

The quadratic subfield k has fundamental unit

$$\varepsilon = \frac{h + \sqrt{h^2+4}}{2}.$$

An easy calculation shows that $\{1, \varepsilon, \alpha_1, \alpha_2\}$ is a basis for the ring of integers of K. Moreover, $\{1, \alpha_1, \varepsilon, \varepsilon\alpha_1\}$ is also an integral basis for K, which shows that $\{1, \alpha_1\}$ is an integral basis for K over k.

2. Fundamental units

Theorem. *$\{\varepsilon, \alpha_1, \alpha_2\}$ is a fundamental system of units for K when $|h| \geq 3$. When $h = 1$, the subgroup $\langle -1, \varepsilon, \alpha_1, \alpha_2\rangle$ is of index 2 in the full group of units.*

Proof. An easy calculation shows that the regulator R' obtained from these units is

$$2\log(\varepsilon)(\log^2(\alpha_1) + \log^2(|\alpha_2|)).$$

If I denotes the index of $\langle -1, \alpha_1, \alpha_2, \varepsilon\rangle$ in the full group of units, then $R' = IR$.

Let U_4 and U_2 be the unit groups for K and k, respectively. Let σ generate $G = \mathrm{Gal}(K/\mathbb{Q})$. The group ring $\mathbb{Z}[G]$ acts on U_4/U_2, and $\sigma^2 + 1$ is an annihilator. It follows that U_4/U_2 can be regarded as a module over the Gaussian integers $\mathbb{Z}[i]$, with i acting as σ. Since U_4/U_2 has $\mathbb{Z}$-rank 2, it must be isomorphic as a $\mathbb{Z}[i]$-module to $\mathbb{Z}[i] \oplus$ torsion. We want to show that the torsion does not occur. Note that $\mathbb{Z}$-torsion is the same as $\mathbb{Z}[i]$-torsion. Suppose $\beta \in U_4$ satisfies $\beta^n \in U_2$ for some $n > 1$. Then $(\sigma^2\beta)^n = \beta^n$, so $\sigma^2\beta = \pm\beta$, since ± 1 are the only roots of unity in the real field K. It follows that $\beta^2 \in k$, so we may assume $n = 2$. Consider the extension $k(\beta)/k$. Since β^2 is a unit, this subextension of K/k can ramify only at primes above 2, hence must be trivial. It follows that $\beta \in U_2$. Therefore, U_4/U_2 is isomorphic to $\mathbb{Z}[i]$ as a $\mathbb{Z}[i]$-module, so there exists a unit η such that η and $\eta' = \sigma(\eta)$ generate U_4/U_2 as an abelian group. Since $(\sigma^2 + 1)\eta \in U_2$, we have

$$\sigma^2(\eta) = \frac{\delta}{\eta}$$

for some $\delta \in U_2$.

An easy calculation using the basis $\{\eta, \eta', \varepsilon\}$ shows that the regulator is

$$R = 2\log(\varepsilon)\left(\left(\log|\eta| - \frac{1}{2}\log|\delta|\right)^2 + \left(\log|\eta'| + \frac{1}{2}\log|\delta|\right)^2\right).$$

In the spirit of [1], we consider the relative extension K/k. Let $D_{K/k}$ be the relative discriminant for the extension K/k. Since

$$D_K = \mathrm{Norm}_{k/\mathbb{Q}}(D_{K/k})D_k^2$$

and $D_k = h^2 + 4$, we have

$$\mathrm{Norm}_{k/\mathbb{Q}}(D_{K/k}) = h^2(h+2)^2(h^2+4).$$

Since $D_{K/k}$ divides $(\eta - \sigma^2\eta)^2$, we have

$$\mathrm{Norm}_{k/\mathbb{Q}}(D_{K/k})|\,\mathrm{Norm}_{k/\mathbb{Q}}\left(\eta - \frac{\delta}{\eta}\right)^2 = \left(\eta - \frac{\delta}{\eta}\right)^2\left(\eta' - \frac{\delta'}{\eta'}\right)^2.$$

Let

$$x = \mathrm{Max}\left(\frac{|\eta|}{|\delta^{\frac{1}{2}}|}, \frac{|\delta^{\frac{1}{2}}|}{|\eta|}\right), \qquad y = \mathrm{Max}\left(|\delta^{\frac{1}{2}}\eta'|, |\delta^{\frac{1}{2}}\eta'|^{-1}\right).$$

Then, since $|\delta||\delta'| = 1$, we find that

$$(*) \qquad h^2(h+2)^2(h^2+4) \le \left(\eta - \frac{\delta}{\eta}\right)^2\left(\eta' - \frac{\delta'}{\eta'}\right)^2 \le \left(x + \frac{1}{x}\right)^2\left(y + \frac{1}{y}\right)^2.$$

If $x \ge 3.25$, then $x + \frac{1}{x} \le 1.1x$, while if $1 \le x < 3.25$ we use the estimate $x + \frac{1}{x} \le 2x$. We use similar estimates for y. Assume first that at least one of $x \ge 3.25$ and $y \ge 3.25$ holds. Then, by Cauchy-Schwarz, we have

$$\log\left(\left(x + \frac{1}{x}\right)\left(y + \frac{1}{y}\right)\right) \le \log(2.2) + \log(x) + \log(y)$$

$$\le \log(2.2) + \sqrt{2}(\log^2(x) + \log^2(y))^{\frac{1}{2}} \le \log(2.2) + \left(\frac{R}{\log(\varepsilon)}\right)^{\frac{1}{2}}.$$

It follows that

$$(**) \qquad \log^2\left(\frac{|h(h+2)|(h^2+4)^{\frac{1}{2}}}{2.2}\right) \le \frac{R}{\log(\varepsilon)} = \frac{2}{I}(\log^2|\alpha_1| + \log^2|\alpha_2|).$$

Since the right side is approximately $\frac{10}{I}\log^2|h|$ and the left side is approximately $9\log^2|h|$, we see that $I < 2$, hence $I = 1$, for sufficiently large h. However, to work with all h, we need estimates on the size of α_1 and α_2. Since

$$f_h((h+1)^2) = h(h+2)^2(h^4 + 3h^3 + 2h^2 - 1),$$
$$f_h(h^2) = -(h^7 + 2h^6 + 4h^5 + 3h^4 - 1),$$
$$f_h(-h) = -(h^4 + 3h^3 + 2h^2 - 1),$$
$$f_h(-h-2) = h^4 + 5h^3 + 8h^2 + 8h + 9,$$

it follows that $|\alpha_1|$ is between h^2 and $(h+1)^2$, and $|\alpha_2|$ is between $|h|$ and $|h+2|$, for $|h| \ge 3$.

For $h > 0$ we obtain

$$I \le 2(\log^2((h+1)^2) + \log^2(h+2))/\log^2\left(\frac{|h(h+2)|(h^2+4)^{\frac{1}{2}}}{2.2}\right).$$

This yields $I < 2$ for $h \ge 5$. Similarly, $I < 2$ for $h \le -5$. Therefore $I = 1$ in these cases. When $h = 3$, the roots α_1 and α_2 can be computed numerically and substituted into (**) to obtain $I < 2$.

If both $|x| < 3.25$ and $|y| < 3.25$ then (*) yields

$$h^2(h+2)^2(h^2+4) < \left(3.25 + \frac{1}{3.25}\right)^4.$$

This implies that $h = 1$ or -3.

We have therefore proved that $I = 1$ except for possibly $h = 1$ and -3. These values may be treated individually. The regulators have been computed (see [2] and the associated unpublished table) for these fields. When $h = -3$, the regulator equals R', so $I = 1$. However, when $h = 1$ the regulator equals $R'/2$, so $I = 2$. In fact, $(-\alpha_1\alpha_2)^{1/2}$, which is a root of $X^4 - 3X^3 - X^2 + 3X + 1$, is a unit of the quartic field.

Remark. The referee has pointed out that the inequality $I < 2$ obtained above can also be obtained by the method of M.-N. Gras [5, p. 9, Proposition 5].

3. Relations with cyclotomic periods

Let $\zeta = \zeta_f = e^{2\pi i/f}$, where

$$f = h(h+2)(h^2+4).$$

Let σ be the generator of $\mathrm{Gal}(K/\mathbb{Q})$ satisfying $\sigma(\alpha_1) = \alpha_2$. We may assume that $\chi(\sigma) = i$, where χ is the Dirichlet character for K introduced in §1 (otherwise, replace χ by χ^{-1}). Define the cyclotomic periods $p_j \in K$ by

$$p_j = \sigma^{j-1}\,\mathrm{Tr}_{\mathbb{Q}(\zeta)/K}(\zeta), \quad 1 \le j \le 4.$$

Theorem. *Assume $h > 0$. Let $\delta = \mu(h(h+2)(h^2+4)) = \pm 1$ and $\eta = (\frac{h}{4})\delta = \pm 1$. Choose α_1 to be the largest root of $f_h(X)$. Then*

$$\alpha_1 - \frac{h^2 - \delta + ((h+2) + \eta)\sqrt{h^2+4}}{4}$$

is a cyclotomic period.

Remarks. A similar result can of course be obtained for $h < 0$, but we omit it. The result of the theorem says that if α_1 is expressed as a $\mathbb{Z}$-linear combination of periods, say $Ap_1 + Bp_2 + Cp_3 + Dp_4$, then $A = C$ and $B - D = \pm 1$, or $A - C = \pm 1$ and $B = D$.

Proof of the Theorem. The main part of the proof will be to determine the minimal polynomial for p_1 over the quadratic field k.

Define the (imprimitive) Gauss sums

$$g_j = \sum_{a=1}^{f} \chi^j(a)\zeta^a, \quad 1 \le j \le 4.$$

Then

$$p_1 = \tfrac{1}{4}(g_1 + g_2 + g_3 + g_4), \qquad p_2 = \tfrac{1}{4}(-ig_1 - g_2 + ig_3 + g_4),$$
$$p_3 = \tfrac{1}{4}(-g_1 + g_2 - g_3 + g_4), \qquad p_4 = \tfrac{1}{4}(ig_1 - g_2 - ig_3 + g_4).$$

(The first relation is easy to prove; the others can be deduced from the Galois action.) We want to study the periods, but first we consider these Gauss sums, since they seem to be easier to work with. Note that $g_4 = \mathrm{Tr}_{\mathbb{Q}(\zeta)/\mathbb{Q}}(\zeta)$. If p is prime and $p \nmid a$, then $\mathrm{Tr}_{\mathbb{Q}(\zeta_{pa})/\mathbb{Q}(\zeta_a)}(\zeta_p \zeta_a) = -\zeta_a$; it follows by induction on the number of prime factors of f that

$$g_4 = \mu(f),$$

where μ is the Möbius function.

Lemma. *Let ψ be a Dirichlet character* mod m. *If n is a positive integer with $m|n$ and with $(m, \frac{n}{m}) = 1$, define*

$$G_n = \sum_{\substack{0<a<n \\ (a,n)=1}} \psi(a)\zeta_n^a.$$

Then

$$G_n = \mu(n/m)\psi(n/m)G_m.$$

Proof. Write $mx + \frac{n}{m}y = 1$ for integers x, y. Note that $\psi(y)\psi(n/m) = 1$. Fix b with $(b, m) = 1$. Choose $b' \equiv b \pmod m$ with $(b', n) = 1$. Then

$$\sum_{\substack{0<a<n \\ (a,n)=1 \\ a\equiv b \,(\mathrm{mod}\, m)}} \zeta_n^a = \mathrm{Tr}_{\mathbb{Q}(\zeta_n)/\mathbb{Q}(\zeta_m)}(\zeta_n^{b'}) = \zeta_m^{yb'}\,\mathrm{Tr}(\zeta_{n/m}^{xb'}) = \zeta_m^{yb}\mu(n/m),$$

as above. Therefore,

$$G_n = \mu(n/m) \sum_{\substack{0<b<m \\ (b,m)=1}} \psi(b)\zeta_m^{yb} = \mu(n/m)\psi(n/m)G_m. \quad \square$$

In our case we have $\psi = \chi^2$, which is quadratic of conductor $m = h^2 + 4$, so $G_m = \sqrt{h^2+4}$. The lemma implies that

$$g_2 = \mu(h(h+2))\chi^2(h(h+2))\sqrt{h^2+4},$$

where χ^2 is regarded as a primitive character mod $h^2 + 4$. Note that if $h > 0$ then

$$\chi^2(h(h+2)) = \left(\frac{h^2+4}{h(h+2)}\right) = \left(\frac{h^2+4}{h}\right)\left(\frac{h^2+4}{h+2}\right) = 1\left(\frac{2}{h+2}\right) = -\left(\frac{-2}{h}\right),$$

the last equality following from a case-by-case consideration of the values of h (mod 8).

It remains to treat g_1 and g_3. Since K and $\mathbb{Q}(i)$ are disjoint over $\mathbb{Q}$, we may consider σ as an element of $\mathrm{Gal}(K(i)/\mathbb{Q}(i))$. Note that $\sigma(g_1) = -ig_1$. It follows that $g_1^2 \in k(i)$, $g_1^2 \notin \mathbb{Q}(i)$, and $g_1^4 \in \mathbb{Q}(i)$. Let ρ be the Lagrange resolvent from §1. Then $\sigma(\rho) = -i\rho$, so $g_1 = \beta\rho$ for some $\beta \in \mathbb{Q}(i)$. Therefore,

$$g_1^4 = \beta^4\rho^4 = \beta^4 h^2(h+2)^2(h^2+4)(h-2i)^2.$$

But $|g_1|^4 = f^2 = h^2(h+2)^2(h^2+4)^2 = |\rho|^4$, so $|\beta| = 1$. Since $h(h+2)(h^2+4)$ is odd and squarefree in $\mathbb{Z}$, it is also squarefree in $\mathbb{Z}[i]$. Therefore, ρ^4 is fourth-power-free in $\mathbb{Z}[i]$. Since $g_1^4 \in \mathbb{Z}[i]$, β must have trivial denominator. Therefore, $\beta \in \mathbb{Z}[i]$, so $\beta = \pm 1$ or $\pm i$. It follows that $\beta^4 = 1$ and

$$g_1^2 = \pm h(h+2)(h-2i)\sqrt{h^2+4}.$$

Since $\overline{g_1} = \chi(-1)g_3 = g_3$, we have

$$g_1^2 + g_3^2 = \pm 2h^2(h+2)\sqrt{h^2+4}.$$

Also, $g_1 g_3 = |g_1|^2 = f$. Therefore,

$$(g_1 + g_3)^2 = 2f + 2\gamma h^2(h+2)\sqrt{h^2+4},$$

where $\gamma = \pm 1$. We shall determine γ later.

Note that p_1 and p_3 are conjugate over k. Since $p_1 + p_3 = \frac{1}{2}(g_2 + g_4)$ and

$$p_1 p_3 = \frac{1}{16}((g_2+g_4)^2 - (g_1+g_3)^2),$$

we can compute the roots of the minimal polynomial of p_1 in terms of h. We obtain

$$\frac{\delta_1 + \delta_2\sqrt{h^2+4} \pm \sqrt{2f + 2\gamma h^2(h+2)\sqrt{h^2+4}}}{4},$$

where $\delta_1 = \mu(f)$ and $\delta_2 = \mu(h(h+2))\chi^2(h(h+2))$.

We may write the roots of $f_h(X)$ as

$$\frac{h^2 + \delta_3(h+2)\sqrt{h^2+4} + \delta_4\sqrt{2f + 2\delta_3 h^2(h+2)\sqrt{h^2+4}}}{4},$$

where $\delta_3, \delta_4 = \pm 1$. The largest root, namely α_1, is obtained by choosing $\delta_3 = \delta_4 = 1$. The root α_i obtained by setting $\delta_3 = \gamma$ and $\delta_4 = 1$ is α_1 if $\gamma = 1$, and is α_2 or α_4 if $\gamma = -1$. It cannot be $\alpha_3 = \sigma^2(\alpha_1)$ since σ^2 fixes $\sqrt{h^2+4}$. We find that

$$\alpha_i - p_j = \frac{h^2 - \delta_1 + ((h+2)\gamma - \delta_2)\sqrt{h^2+4}}{4},$$

where $i = 1, 2$, or 4, and $j = 1$ or 3. Applying σ or σ^{-1} if $\gamma = -1$, and taking into account the sign change of $\sqrt{h^2+4}$, we find that in all cases

$$\alpha_1 - \frac{h^2 - \delta_1 + ((h+2) - \gamma\delta_2)\sqrt{h^2+4}}{4}$$

is one of the cyclotomic periods p_1, p_2, p_3, p_4.

Since $\alpha_i - p_j$ is an algebraic integer, we must have

$$h^2 - \delta_1 \equiv (h+2) - \gamma\delta_2 \pmod 4.$$

It follows upon considering the four possibilities for δ_1, δ_2 that $\gamma = -(\frac{h}{4})\delta_1\delta_2$, where $(\frac{h}{4})$ is the Legendre symbol. This completes the proof of the theorem. □

4. Modular curves

Recall that

$$\Gamma_0(16) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \mid c \equiv 0 \pmod{16} \right\}$$

and

$$\Gamma_1(16) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(16) \mid a, d \equiv 1 \pmod{16} \right\}.$$

These groups act on the upper half plane $\mathscr{H}$ as fractional linear transformations. We set

$$X_0(16) = \Gamma_0(16)\backslash\mathscr{H} \cup \{\text{cusps}\}$$

and, similarly,

$$X_1(16) = \Gamma_1(16)\backslash\mathscr{H} \cup \{\text{cusps}\}.$$

Both $X_0(16)$ and $X_1(16)$ can be regarded as algebraic curves defined over $\mathbb{Q}$. Four of the six cusps of $X_0(16)$ are then rational over $\mathbb{Q}$, namely $0, \frac{1}{2}, \frac{1}{8}, \infty$. Six of the fourteen cusps of $X_1(16)$ are rational over $\mathbb{Q}$, namely $0, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \frac{1}{2}, \frac{1}{6}$. Call these $P_1, P_3, P_5, P_7, P_2, P_6$, respectively. In the quartic covering $X_1(16) \to X_0(16)$, the cusps P_1, P_3, P_5, P_7 lie above 0, and P_2, P_6 lie above $\frac{1}{2}$ (for details, see [14]).

Let $\wp(z, \tau)$ be the Weierstrass $\wp$-function for the lattice $\tau\mathbb{Z} + \mathbb{Z}$ (with $\tau \in \mathscr{H}$). For an integer s, let

$$\phi_s(\tau) = \wp\left(\frac{s}{16}, \tau\right).$$

Then $\phi_{rs} := \phi_r - \phi_s$ is a modular form of weight 2 on $\Gamma_1(16)$ for any two indices r, s. The Fourier expansion of each ϕ_s can be computed, and one finds that ϕ_s has a zero of order at least $\{st\}/d$ at the cusp P_t. Here, $d = (t, 16) = 1$ or 2 (= ramification index over the corresponding cusp of $X(16)$) and $\{st\}$ is defined by $\{st\} \equiv \pm st \pmod{16}$, $0 \le \{st\} \le 8$ (see [14]). It follows that the divisors of ϕ_3 and ϕ_4 satisfy

$$(\phi_3) \ge 3P_1 + 7P_3 + P_5 + 5P_7 + 3P_2 + P_6,$$
$$(\phi_4) \ge 4P_1 + 4P_3 + 4P_5 + 4P_7 + 4P_2 + 4P_6.$$

Therefore,

$$(\phi_{34}) \ge \min\{(\phi_3), (\phi_4)\} = 3P_1 + 4P_3 + P_5 + 4P_7 + 3P_2 + P_6.$$

But a nonzero modular form of weight $k = 2$ for $\Gamma_1(16)$ has exactly $k[\Gamma : \Gamma_1(16)]/12 = 16$ zeros modulo $\Gamma_1(16)$. Therefore,

$$(\phi_{34}) = 3P_1 + 4P_3 + P_5 + 4P_7 + 3P_2 + P_6.$$

Similarly,

$$(\phi_{54}) = 4P_1 + P_3 + 4P_5 + 3P_7 + 3P_2 + P_6.$$

Therefore,

$$\left(\frac{\phi_{34}}{\phi_{54}}\right) = -P_1 + 3P_3 - 3P_5 + P_7 .$$

Note that ϕ_{34}/ϕ_{54} has its zeros and poles only at the cusps, so it is what is known as a "modular unit" (see [7]).

Let $K_1(16)$ denote the field of meromorphic functions on $X_1(16)$, so $K_1(16)$ is the field of meromorphic functions $f(\tau)$ on $\mathscr{H}$ which are also meromorphic at the cusps and which are invariant under the change of variables $\tau \mapsto \gamma(\tau)$ for every $\gamma \in \Gamma_1(16)$. We define $K_0(16)$ similarly. The matrices

$$\begin{pmatrix} 3 & 2 \\ 16 & 11 \end{pmatrix}^j, \qquad 0 \le j \le 3,$$

form a set of coset representatives for $\Gamma_1(16)\backslash\Gamma_0(16)$ and, consequently, give the elements of $\mathrm{Gal}(K_1(16)/K_0(16))$. More explicitly, $\mathrm{Gal}(K_1(16)/K_0(16))$ is cyclic of order 4, generated by the map

$$\sigma : f(\tau) \mapsto f\left(\frac{3\tau + 2}{16\tau + 11}\right)$$

for $f \in K_1(16)$. Note that $\left(\begin{smallmatrix} 3 & 2 \\ 16 & 11 \end{smallmatrix}\right)$ maps P_1, P_3, P_5, P_7 to P_5, P_1, P_7, P_3 (modulo the action of $\Gamma_1(16)$), respectively.

An easy calculation yields $\sigma(\phi_{ij}/\phi_{kl}) = \phi_{11i,11j}/\phi_{11k,11l}$ (indices are taken mod 16), hence $\sigma(\phi_{34}/\phi_{54}) = \phi_{14}/\phi_{74}$, $\sigma^2(\phi_{34}/\phi_{54}) = \phi_{54}/\phi_{34}$, $\sigma^3(\phi_{34}/\phi_{54}) = \phi_{74}/\phi_{14}$. In terms of the Klein forms we have (see [7, p. 51]),

$$\phi_{34}/\phi_{54} = -\frac{\mathfrak{k}^2_{(0,\frac{5}{16})}}{\mathfrak{k}^2_{(0,\frac{3}{16})}}, \qquad \phi_{14}/\phi_{74} = -\frac{\mathfrak{k}^2_{(0,\frac{7}{16})}}{\mathfrak{k}^2_{(0,\frac{1}{16})}}.$$

This indicates that we should try taking square roots; in fact,

$$\alpha_1 := \left(\frac{\phi_{34}}{\phi_{54}}\frac{\phi_{14}}{\phi_{74}}\right)^{\frac{1}{2}} = \frac{\mathfrak{k}_{(0,\frac{5}{16})}}{\mathfrak{k}_{(0,\frac{3}{16})}}\frac{\mathfrak{k}_{(0,\frac{7}{16})}}{\mathfrak{k}_{(0,\frac{1}{16})}}$$

is a modular function for $\Gamma_1(16)$. This may be checked by using the transformation laws for Klein forms (see [7, pp. 27–28]). These laws also show that

$$\alpha_2 := \sigma(\alpha_1) = -\frac{\mathfrak{k}_{(0,\frac{7}{16})}}{\mathfrak{k}_{(0,\frac{5}{16})}}\frac{\mathfrak{k}_{(0,\frac{1}{16})}}{\mathfrak{k}_{(0,\frac{3}{16})}}, \quad \sigma^2(\alpha_1) = 1/\alpha_1, \quad \sigma^3(\alpha_1) = 1/\alpha_2 .$$

The Klein forms have product expansions [7, p. 29] in terms of $q = e^{2\pi i\tau}$, from which we obtain expansions for α_1 and α_2. It is convenient to change variables to $-1/16\tau$ (the Atkin-Lehner involution):

$$\alpha_1^* := \alpha_1\left(\frac{-1}{16\tau}\right) = q^{-2}\prod_{\substack{n\equiv\pm5,\pm7(16)\\ n>0}}(1-q^n)\prod_{\substack{n\equiv\pm1,\pm3(16)\\ n>0}}(1-q^n)^{-1},$$

$$\alpha_2^* := \alpha_1\left(\frac{-1}{16\tau}\right) = -q^{-1}\prod_{\substack{n\equiv\pm1,\pm7(16)\\ n>0}}(1-q^n)\prod_{\substack{n\equiv\pm3,\pm5(16)\\ n>0}}(1-q^n)^{-1}$$

(since $\left(\begin{smallmatrix} 0 & -1 \\ 16 & 0 \end{smallmatrix}\right)$ normalizes $\Gamma_1(16)$, these are functions on $X_1(16)$).

The curve $X_0(16)$ has genus 0, hence there is a Hauptmodul, namely a function H such that $K_0(16) = \mathbb{C}(H)$. We may take

$$H = \frac{2\sum_{n\in\mathbb{Z}} q^{(2n)^2}}{\sum_{n\in\mathbb{Z}} q^{(2n+1)^2}} = \frac{1}{q} + 2q^3 - q^7 + \cdots .$$

Then H is a modular function on $\Gamma_0(16)$ and has a simple pole at ∞ and no other poles. This may be seen as follows. Standard techniques show that both $\sum_{n\in\mathbb{Z}} q^{(2n)^2}$ and $\sum_{n\in\mathbb{Z}} q^{(2n+1)^2}$ are modular forms of weight $\frac{1}{2}$ on $\Gamma_0(16)$. The number of zeros of such a modular form is $\frac{1}{2}[\Gamma : \Gamma_0(16)]/12 = 1$, so $\sum_{n\in\mathbb{Z}} q^{(2n+1)^2}$ has exactly one zero, which is clearly at ∞. Therefore, H has one pole, which is likewise at ∞. A standard argument now shows that H is a Hauptmodul.

The minimal polynomial for α_1^* over $K_0(16)$ is

$$f_H(X) = (X - \alpha_1^*)(X - \alpha_2^*)(X - 1/\alpha_1^*)(X - 1/\alpha_2^*).$$

The coefficient of X^3, call it c^*, has q-expansion

$$-q^{-2} - 4q^2 - 2q^6 + 8q^{10} + \cdots ,$$

which agrees with the q-expansion of $-H^2$ at least through the q^{10} term. We want to show that they are in fact equal. The function α_1 has divisor $-2P_1 + P_3 - P_5 + 2P_7$ and α_2 has divisor $-P_1 - 2P_3 + 2P_5 + P_7$, so $\alpha_1 + \alpha_2 + 1/\alpha_1 + 1/\alpha_2$ has poles at most at P_1, P_3, P_5, P_7. Under the change of variables $\tau \mapsto -1/16\tau$ these cusps are mapped to cusps of $X_1(16)$ lying above the cusp ∞ of $X_0(16)$. Therefore, $c^* + H^2 = \alpha_1^* + \alpha_2^* + 1/\alpha_1^* + 1/\alpha_2^* + H^2$ is a function on $X_0(16)$ having poles at most at ∞. But $c^* + H^2$ has a zero of order at least 10 at ∞, hence has no poles, and therefore must vanish identically. Similarly, we find that the coefficient of X^2 is $-(H^3 + 2H^2 + 4H + 2)$. Therefore,

$$f_H(X) = X^4 - H^2X^3 - (H^3 + 2H^2 + 4H + 2)X^2 - H^2X + 1.$$

This polynomial has Galois group $\mathbb{Z}/4\mathbb{Z}$ over $\mathbb{C}(H)$. Its specialization to $H = h \in \mathbb{Z}$ yields the polynomial f_h of §1.

It might be worth noting that while the equation $f(H, X) = f_H(X) = 0$ gives an equation for $X_1(16)$ as an algebraic curve, it is easy to modify it into a more standard form. The extension $\mathbb{C}(H, \alpha_1^*)/\mathbb{C}$ contains the intermediate field $k = \mathbb{C}(H, \sqrt{H^2 + 4})$. This corresponds to the function field for the genus zero curve $u^2 - v^2 = 4$, which has the rational parametrization

$$u = 2\frac{1 + t^2}{1 - t^2}, \qquad v = \frac{4t}{1 - t^2}.$$

Therefore, $k = \mathbb{C}(t)$. Since

$$z := \alpha_1 + \frac{1}{\alpha_1} = \frac{H^2 + (H + 2)\sqrt{H^2 + 4}}{2},$$

the extension K/k is generated by

$$\sqrt{z^2-4}=\frac{-16t(t^2+1)(t^2-2t-1)}{(t-1)^4(t+1)^2}.$$

Therefore, $K=\mathbb{C}(t,\sqrt{t(t^2+1)(1+2t-t^2)})$. It follows that an equation of $X_1(16)$ as an algebraic curve is

$$Y^2=X(X^2+1)(1+2X-X^2).$$

This of course agrees with the fact that $X_1(16)$ has genus 2. The discriminant of the polynomial in X on the right-hand side is $-2048=-2^{11}$, which reflects the fact that $X_1(16)$ has good reduction outside 2.

The transformation

$$X=\frac{c+1}{c-1},\qquad Y=\frac{2z}{(c-1)^3}$$

changes the above equation into

$$z^2=(c^2+1)(c-1)(c^3-c^2-3c-1),$$

which was obtained by Beppo Levi [12, p. 113] in 1906.

It also seems worth mentioning that it is easy to use either of the above equations to find the rational points on $X_1(16)$. Levi did this using the latter equation; similarly, with the first equation we obtain by the standard type of argument that either X, X^2+1, $1+2X-X^2$ are all squares, or X is a square and the latter two factors are 2 times squares. The standard descent argument ([13, pp. 16-18]) shows that the only rational solutions to $a^2=b^4+1$ are $(\pm1,0)$, and the only rational solutions to $2a^2=b^4+1$ are $(\pm1,\pm1)$ (with all four choices of signs). We find that we must have $(X,Y)=(0,0),(\pm1,\pm2)$ (with all four choices of signs). There is also the point at infinity. Since there are six rational cusps, these six points must correspond to the cusps. Therefore, $X_1(16)$ has no noncuspidal rational points.

Acknowledgments

I would like to thank Don Zagier for several helpful conversations during the preparation of this paper.

Bibliography

1. A.-M. Bergé and J. Martinet, *Notions relatives de régulateurs et de hauteurs*, Acta Arith. (to appear).
2. J. Buchmann and J. v. Schmettow, *On the computation of unit groups and class groups of totally real quartic fields*, Math. Comp. **53** (1989), 387–397.
3. G. Cornell and L. C. Washington, *Class numbers of cyclotomic fields*, J. Number Theory **21** (1985), 260–274.
4. H. Darmon, *Note on a polynomial of Emma Lehmer*, Math. Comp. **56** (1991), 795–800.

5. M.-N. Gras, *Table numérique du nombre de classes et des unités des extensions cycliques de degré* 4 *de* $\mathbb{Q}$, Publ. Math. Fac. Sci. Besançon, fasc. 2 (1977/78).
6. ——, *Special units in real cyclic sextic fields*, Math. Comp. **48** (1987), 179–182.
7. D. Kubert and S. Lang, *Modular units*, Springer-Verlag, New York-Heidelberg-Berlin, 1981.
8. A. Lazarus, *The class number and cyclotomy of simplest quartic fields*, Thesis, Univ. of California, Berkeley, 1989.
9. O. Lecacheux, *Unités d'une famille de corps cycliques réels de degré* 6 *liés à la courbe modulaire* $X_1(13)$, J. Number Theory **31** (1989), 54–63.
10. ——, *Unités d'une famille de corps liés à la courbe* $X_1(25)$, Ann. Inst. Fourier (Grenoble) **40** (1990), 237–253.
11. E. Lehmer, *Connection between Gaussian periods and cyclic units*, Math. Comp. **50** (1988), 535–541.
12. B. Levi, *Saggio per una teoria aritmetica delle forme cubiche ternarie*, Atti Accad. Reale Sci. Torino **43** (1908), 99–120.
13. L. J. Mordell, *Diophantine equations*, Academic Press, London-New York, 1969.
14. A. Ogg, *Rational points on certain elliptic modular curves*, Proc. Sympos. Pure Math., vol. 24, Amer. Math. Soc., Providence, R.I., 1973, pp. 221–231.
15. R. Schoof and L. C. Washington, *Quintic polynomials and real cyclotomic fields with large class numbers*, Math. Comp. **50** (1988), 543–556.
16. E. Seah, H. C. Williams, and L. C. Washington, *The calculation of a large cubic class number with an application to real cyclotomic fields*, Math. Comp. **41** (1983), 303–305.
17. D. Shanks, *The simplest cubic fields*, Math. Comp. **28** (1974), 1137–1152.

Department of Mathematics, University of Maryland, College Park, Maryland 20742