Glasgow Math. J. **48** (2006) 351–363. © 2006 Glasgow Mathematical Journal Trust.
doi:10.1017/S0017089506003223. Printed in the United Kingdom

CONGRUENCE AND NON-CONGRUENCE SUBGROUPS OF THE (2,3,7)-GROUP

W. W. STOTHERS

Department of Mathematics, University of Glasgow, Glasgow, G12 8QW, Scotland, U.K.
e-mail: wws@maths.gla.ac.uk

(Received 9 December, 2003; revised 10 April, 2006; accepted 2 July, 2006)

Abstract. Let Δ denote the (2,3,7)-group. We establish an upper bound for the number of congruence subgroups of index n and a lower bound for the total number of subgroups of index n. Since the latter grows more quickly, there exist non-congruence subgroups of index n for all n greater than some n_0.

2000 *Mathematics Subject Classification.* 10D05, 20H10.

1. Introduction. Let Δ denote the (2,3,7)-group, i.e.

$$\Delta = \langle x, y, z : x^2 = y^3 = z^7 = xyz = 1 \rangle.$$

Let $\mu = 2\cos(2\pi/7)$, and $K = \mathbb{Q}(\mu)$. We see in Section 2 that K and Δ are closely related.

In [**1**] Cohen obtained families of subgroups of Δ by studying finite quotients of $\mathbb{Z}[\mu]$. This generalised work of Macbeath [**5**]. We shall see that Cohen's subgroups are *congruence* subgroups in an appropriate sense. In Section 3, we obtain them naturally from a quaternion algebra over K.

In [**2**] Conder showed that, for all but finitely many values of n, the alternating group A_n is a quotient of Δ. This uses unpublished ideas of Higman. We shall see that these are *non-congruence* subgroups.

To help us distinguish the classes of subgroups, we prove in Section 4 a result which relates the *index* of a congruence subgroup to its *level* (an ideal of $\mathbb{Z}[\mu]$). This is analogous to the central result in [**12**], and the proof uses essentially the same idea.

We will establish an *upper* bound for the number of congruence subgroups of index n, and a lower bound for the total number of subgroups of index n. Since the latter grows more quickly, there exist non-congruence subgroups of index n for all n greater than some n_0.

We use Condor's results [**2**] to show that we may take $n_0 = 167$. In fact, Condor's results also give non-congruence subgroups of smaller index. The smallest is 15. As we shall see, Sinkov's subgroups [**9**] are non-congruence of index 14. This is the lowest possible index for a *non-congruence* subgroup.

In [**8**], Shimura considered congruence subgroups. We will follow his approach in Section 3. For our counting arguments and work on non-congruence subgroups, we need more detail than that given in [**8**]. We therefore begin with a discussion of a certain quadratic form. This overlaps with some of Shimura's work. See, in particular, [**8**, Example 3.19, pp. 83–84].

REMARK. It is easy to see that Δ is a quotient of the classical modular group Γ, which can be defined as

$$\Gamma = \langle t, p : t^2 = p^3 = 1 \rangle.$$

It follows that each subgroup of Δ lifts to a subgroup of Γ, but it is important to note that congruence subgroups do not, in general, lift to congruence subgroups of Γ (see Section 3).

2. Algebraic preliminaries. Let $\omega = \exp(2\pi i/7)$. Then $\mu = \omega + \omega^{-1}$, so that μ is a root of the (irreducible) cubic

$$f(x) = x^3 + x^2 - 2x - 1. \tag{1}$$

In fact, μ is the only positive root, and $\mu > 1$. We sum up the results we require about K in the following theorem. We merely sketch the proof since most of it is a straight-forward exercise in algebraic number theory.

DEFINITION. For any non-zero ideal A of $\mathbb{Z}[\mu]$, we write $N(A)$ for the cardinality of the ring $\mathbb{Z}[\mu]/A$.

THEOREM 2.1. (i) $\mathbb{Z}[\mu]$ *is the ring of integers of* K.
(ii) $1 - \mu$ *is a unit of* $\mathbb{Z}[\mu]$, *and its conjugates in* K *are positive.*
(iii) *Let* p *be a rational prime. The ideal* (p) *in* $\mathbb{Z}[\mu]$
  (a) *splits as* $P_1P_2P_3$, *with* $\mathbb{Z}[\mu]/P_k \cong GF(p)$, $N(P_k) = p$, *if* $p \equiv \pm 1 \pmod 7$,
  (b) *ramifies as* $(2 - \mu)^3$, *with* $\mathbb{Z}[\mu]/(2 - \mu) \cong GF(7)$, $N((2 - \mu)) = 7$, *if* $p = 7$,
  (c) *remains prime, with* $\mathbb{Z}[\mu]/(p) \cong GF(p^3)$, $N((p)) = p^3$, *otherwise.*
(iv) *Each totally positive unit in* $\mathbb{Z}[\mu]$ *is a square.*

Sketch of proof. (i) It is well-known (e.g. [**14**]) that an integer α in $\mathbb{Q}(\omega)$ can be written as $\sum_{k=1}^{6} a_k\omega^k$, with $a_k \in \mathbb{Z}$. If $\alpha \in K$, then $a_k = a_{7-k}$ $(k = 1, 2, 3)$, and the result follows.

(ii) follows from formula (1) and the remark after it.

(iii) is an application of a theorem of Kummer (see [**14**, p. 317]).

For (iv), we observe that K has 3 archimedean places, all real, so the unit group of $\mathbb{Z}[\mu]$ has 3 generators. Hence the square units have index 8. The result follows by considering signs of μ and its conjugates $\mu^2 - 2$, $1 - \mu - \mu^2$.

Let Q be the quaternion algebra associated with the form

$$F(a_1, a_2, b_1, b_2) = a_1^2 + b_1^2 + (1 - \mu)\big(a_2^2 + b_2^2\big),$$

with $a_1, a_2, b_1, b_2 \in K$. We write a typical element as $[a_1, a_2, b_1, b_2]$. The identity is $[1, 0, 0, 0]$ and

$$[0, 0, 1, 0]^2 = (\mu - 1)[1, 0, 0, 0].$$

Since $(\mu - 1) > 0$, we can put $j = (\mu - 1)^{1/2}$ (so $j > 0$). Then we view Q as a $K(j)$-algebra (with $[0, 0, 1, 0]$ corresponding to j). We have the matrix description:

$$Q = \left\{ \begin{pmatrix} a_1 + ja_2 & b_1 + jb_2 \\ -b_1 + jb_2 & a_1 - jb_2 \end{pmatrix} : a_1, a_2, b_1, b_2 \in K \right\}$$

For brevity, we shall usually use the quadruple notation.

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

THEOREM 2.2. *The $\mathbb{Z}[\mu]$-module $\mathbb{B}$ generated by*

$$u_1 = [1, 0, 0, 0], \qquad u_2 = [0, 0, -1, 0],$$
$$u_3 = \tfrac{1}{2}[1, -1, \mu^2 + \mu - 1, 0], \quad u_4 = \tfrac{1}{2}[\mu^2 + \mu - 1, 0, -1, 1]$$

is a maximal order of Q.

Proof. To show that $\mathbb{B}$ is an order, we need only verify that for $m, n \in \{1, \ldots, 4\}$, $u_m u_n \in \mathbb{B}$. This is routine.

To show that $\mathbb{B}$ is maximal, we compute the discriminant (i.e. $\det || \operatorname{tr} u_m u_n ||$). This turns out to be $(\mu - 1)^2$, a unit of $\mathbb{Z}[\mu]$ (by Theorem 2.1(ii)). Since any order strictly containing $\mathbb{B}$ would have as discriminant a proper divisor, $\mathbb{B}$ is maximal.

Let $\overline{\Delta} = \{X \in \mathbb{B} : \det(x) = 1\}$. Observe that each u_k belongs to $\overline{\Delta}$, so the group generated by the u_k lies in $\overline{\Delta}$.

THEOREM 2.3. *With u_2, u_3, u_4 as in Theorem* 2.2,
(i) $\overline{\Delta} = \langle u_2, u_3, u_4 \rangle$.
(ii) $\Delta \cong \overline{\Delta}/\{\pm I\}$.

Proof. This result is *stated* in [**7**, p. 46]. We shall sketch a proof here.
We begin by observing that (from calculations)

$$u_2^2 = u_3^3 = u_4^7 = I, \quad u_4 = u_2 u_3.$$

Indeed, by considering the Möbius transformations corresponding to u_1, u_2, u_3 and their fixed points, $< u_2, u_3, u_4 > /\{\pm I\}$ is isomorphic to Δ.

From [**7**, p. 40], $\overline{\Delta}/\{\pm I\}$ is discontinuous. Since the (2,3,7)-triangle group is maximal among discontinuous groups, the results follow.

REMARKS. (i) To follow Shimura's line of proof, we need Theorem 2.1(iv) to see that the totally positive units of $\mathbb{B}$ have the form αu, with α a unit of $\mathbb{Z}[\mu]$, and $u \in \overline{\Delta}$.

(ii) This representation of Δ appeared in [**4**], though Fricke used the order $2\mathbb{B}$, with a corresponding operation $X * Y = \frac{1}{2} XY$. By putting restrictions modulo 2 Fricke obtained subgroups of index 9 and 63. In the language of Section 3, these are congruence subgroups.

3. Congruence Subgroups.

DEFINITION. For a prime P of $\mathbb{Z}[\mu]$, we write K_P (resp. $\mathbb{Z}[\mu]_P$) for the P-completion of K (resp. $\mathbb{Z}[\mu]$).

LEMMA 3.1. *For each finite non-zero prime P of $\mathbb{Z}[\mu]$, there is a K_P-algebra isomorphism $\phi_P : Q \otimes_K K_P \to M_2(K_P)$ such that $\overline{\Delta}$ is mapped into $SL_2(\mathbb{Z}[\mu]_P)$.*

Proof. From [**3**, p. 54], we see that $Q \otimes_K K_P$ is isomorphic to $M_2(K_P)$ if and only if

$$F(a_1, a_2, b_1, b_2) = 0 \tag{2}$$

has a non-trivial zero over K_P.

If $P \neq 2\mathbb{Z}[\mu]$, then Hensel's Lemma can be applied to (2), starting with a zero modulo P. Since $(1 - \mu)$ is a unit, this is easy to find.

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

If $P = 2\mathbb{Z}[\mu]$, we must start with a zero modulo $8\mathbb{Z}$. We may take $a_1 = a_2 = 1$, $b_1 = \mu^2 + \mu - 1$, $b_2 = 2t$, where $t^2(1-\mu) \equiv 1 (\mathrm{mod}\ 2)$. (The element t exists since each element of $\mathbb{Z}[\mu]/2\mathbb{Z}[\mu]$ is a square.)

Since $M_2(K_P)$ has one conjugacy class of maximal order; and this includes $M_2(\mathbb{Z}[\mu]_P)$, we may assume that $\phi_P(\mathbb{B}) = M_2(\mathbb{Z}[\mu]_P)$.

We observe that, for any 2×2 matrix X,

$$(\mathrm{tr}(X)I - X)X = \det(X)I.$$

Applying (the K_P-algebra isomorphism) ϕ_P to each side, we see that ϕ_P preserves determinants. The final part of the lemma follows.

Although we do not require an explicit ϕ_P, it is worth noting that we can obtain one, at least when P contains to an *odd* rational prime.

Case 1: $(\mu - 1)$ is a quadratic residue modulo P.

By Hensel's Lemma, we can find $j_P \in k_P$ with $j_P^2 = \mu - 1$. Then a suitable ϕ_P is obtained by mapping j to j_P.

Case 2: $(\mu - 1)$ is a quadratic non-residue modulo P.

Then $(c, d) \mapsto c^2 - (\mu - 1)d^2$ is a norm map from the quadratic extension of $\mathbb{Z}[\mu]/P$ to $\mathbb{Z}[\mu]/P$. Thus, by Hensel's Lemma, we can find $c, d \in k_P$ with $c^2 - (\mu - 1)d^2 = -1$. Put

$$A = \begin{pmatrix} 1 & c \\ j & -cj \end{pmatrix}.$$

For $a_1, a_2, b_1, b_2 \in \mathbb{Z}[\mu]_P$, we have

$$A \begin{pmatrix} a_1 + ja_2 & b_1 + jb_2 \\ -b_1 + jb_2 & a_1 - ja_2 \end{pmatrix} A^{-1} = \begin{pmatrix} a_1 - b_1 & a_2 - b_2 \\ (a_2 + b_2)(\mu - 1) & a_1 + b_1 \end{pmatrix}.$$

The required ϕ_P is now obvious.

We now return to the view of Q as a quadratic space over K with form given by (1). We write Q_P for the corresponding quadratic space over K_P.

THEOREM 3.2. *Suppose that T is a finite set of finite primes of $\mathbb{Z}[\mu]$ and that, for each $P \in T$, we have an $X_P \in Q_P$ with* $\det(X_P) = 1$. *For any positive integer r, there is an $X \in \overline{\Delta}$ with*

$$X \equiv X_P (\mathrm{mod}\ P^r) \quad (P \in T).$$

Proof. This is a special case of a theorem in [**6**, p. 314]. To see that this applies, we observe that

(i) Q has dimension 4 over K.

(ii) Since $(1 - \mu)$ has a positive conjugate, $\det(X) = 0 \Rightarrow X = 0$, so Q is "regular".

(iii) The set of finite primes of $\mathbb{Z}[\mu]$ is a "Dedekind set", and Q_P is "isotropic" where P is the infinite prime corresponding to the identity embedding of K in $\mathbb{R}$.

DEFINITIONS. For any non-zero ideal A of $\mathbb{Z}[\mu]$, the *principal congruence subgroup of level A in $\overline{\Delta}$* is defined by

$$\overline{\Delta}(A) = \{X \in \overline{\Delta} : X - I \in A\mathbb{B}\},$$

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

and in Δ by

$$\Delta(A) = \overline{\Delta}(A) \cdot \{\pm I\}/\{\pm I\}.$$

In some respects, the $\overline{\Delta}(A)$ are easier to handle. For example,

$$A = \prod_k P_k^{r(k)} \Rightarrow \overline{\Delta}(A) = \cap_k \overline{\Delta}\big(P_k^{r(k)}\big).$$

Clearly, $\overline{\Delta}(A) \lhd \overline{\Delta}$ and $\Delta(A) \lhd \Delta$. Using ϕ_P, we have maps

$$\overline{\psi}(P^r) : \overline{\Delta}/\overline{\Delta}(P^r) \to SL_2(\mathbb{Z}[\mu]/P^r).$$

By the Chinese Remainder Theorem in $\mathbb{Z}[\mu]$,

$$A = \prod_k P_k^{r(k)} \Rightarrow SL_2(\mathbb{Z}[\mu]/A) = \prod_k SL_2\big(\mathbb{Z}[\mu]/P_k^{r(k)}\big) \tag{3}$$

(The product on the right of (3) is direct.) Then we have a map

$$\overline{\psi}_A : \overline{\Delta}/\overline{\Delta}(A) \to SL_2(\mathbb{Z}[\mu]/A).$$

Finally, we define maps

$$\psi_A : \Delta/\Delta(A) \to PSL_2(\mathbb{Z}[\mu]/A)$$

by

$$\psi_A(\{\pm X\}) = \{\pm\overline{\psi_A}(X)\}.$$

THEOREM 3.3. *For each non-zero ideal A of $\mathbb{Z}[\mu]$, the maps $\overline{\psi_A}$ and ψ_A are isomorphisms.*

Proof. We need only show that the maps are surjective.

The result for $A = P^r$ follows easily from Theorem 3.2 (with $T = \{P\}$) once we observe that, if $\det(X) \equiv 1 (\mathrm{mod}\, P^s)$, then we can find $Y \equiv X (\mathrm{mod}\, P^s)$ with $\det(Y) \equiv 1 (\mathrm{mod}\, P^{s+1})$. Hence each element of $SL_2(\mathbb{Z}[\mu]/P^r)$ corresponds to an element of $SL_2(\mathbb{Z}[\mu]_P)$

For $A = \prod_{k=1}^n P_k^{r(k)}$, we apply Theorem 3.2 with $T = \{P_1, \ldots, P_n\}$ and $r = \max\{r(k)\}$.

We now determine the indices of the $\overline{\Delta}(A)$, and hence of the $\Delta(A)$. After Theorem 3.3, this amounts to finding the order of $SL_2(\mathbb{Z}[\mu]/A)$. For P prime, we know that

$$|\overline{\Delta} : \overline{\Delta}(A)| = |SL_2(\mathbb{Z}[\mu]/P)| = N(P)(N(P)^2 - 1). \tag{4}$$

THEOREM 3.4. *If P is a prime ideal of $\mathbb{Z}[\mu]$, and r is a positive integer, then*

$$|\overline{\Delta}(P^r) : \overline{\Delta}(P^{r+1})| = N(P)^3$$

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

Proof. Let π generate P over P^2. If $X \in \overline{\Delta}(P^r)$, then, as $X \equiv I (\mathrm{mod}\, P^r)$,

$$\phi_p(X) \equiv \begin{pmatrix} 1+\pi^r a & \pi^r b \\ \pi^r c & 1+\pi^r d \end{pmatrix} (\mathrm{mod}\, P^{r+1}).$$

As $\det(X) = 1$, $d \equiv -a (\mathrm{mod}\, P)$. The result now follows from Theorem 3.2.

COROLLARY 3.5. *If* $A = \prod_k P_k^{r(k)}$, *then*
(i) $|\overline{\Delta} : \overline{\Delta}(A)| = \prod_k N(P)^{3r(k)-2}(N(P)^2 - 1)$, *and*
(ii) $|\Delta : \Delta(A)| = \varepsilon(A)|\overline{\Delta} : \overline{\Delta}(A)|$, *where*

$$\varepsilon(A) = \begin{cases} 1 & \text{if } A = 2\mathbb{Z}[\mu], \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

Proof. (i) follows from (3), (4) and Theorem 3.4.
(ii) follows from (i) and the observation that

$$-I \in \overline{\Delta}(A) \Leftrightarrow 2 \in A.$$

THEOREM 3.6. *If* A *and* B *are non-zero ideals of* $\mathbb{Z}[\mu]$, *then*

$$\Delta(A) \cdot \Delta(B) = \Delta(\gcd(A, B)).$$

Proof. Let $C = \gcd(A, B)$ and $D = \mathrm{lcm}(A, B)$.
As $C \supseteq A$, $B \supseteq D$, we have $\Delta(C) \supseteq \Delta(A)$, $\Delta(B) \supseteq \Delta(D)$ so that

$$\Delta(C) \supseteq \Delta(A) \cdot \Delta(B) \supseteq \Delta(D). \tag{5}$$

Suppose that P is a prime ideal dividing A or B, and that P^r (resp. P^s) is the greatest power dividing A (resp. B).

Suppose first that $r > s$. Then $D = P^r D_1$, $c = P^s C_1$, with $P \nmid C_1, D_1$. By Theorem 3.3 $\Delta(B)$ has coset representatives of $P^s D_1$ over $P^r D_1$, so we can replace $\Delta(D)$ by $\Delta(P^r D_1)$ in (5).

A similar result holds if $r < s$, while, if $r = s$, C and D have the same power of P.

Considering all primes dividing A or B, we get the result.

Suppose that $G \leq \Delta$ and that there are non-zero ideals A and B of $\mathbb{Z}[\mu]$ such that

$$\Delta(A) \leq G, \quad \Delta(B) \leq G.$$

Then, by Theorem 3.6,

$$\Delta(\gcd(A, B)) \leq G.$$

It follows that the set of ideals C with $\Delta(C) \leq G$ has a greatest element (with respect to set inclusion).

DEFINITIONS. A subgroup G of Δ is a *congruence subgroup* if, for some non-zero ideal A of $\mathbb{Z}[\mu]$,

$$\Delta(A) \leq G.$$

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

The *level* of the congruence subgroup G is the largest ideal A for which the inclusion is satisfied.

REMARK. By Theorem 3.3, we see that the description "principal congruence subgroup of level A" is consistent with the above definitions.

Since the level is maximal, we have the following result.

LEMMA 3.7. *If $G \leq \Delta$ is a congruence subgroup of level A and A' strictly contains A, then G does not contain coset representatives of $\Delta(A')$ over $\Delta(A)$.*

An alternative definition would be that $G \leq \Delta$ is congruence if, for some non-zero ideal A of $\mathbb{Z}[\mu]$, and for some subgroup H of $PSL_2[\mathbb{Z}[\mu]/\Delta)$,

$$G = \psi_A^{-1}(H). \tag{6}$$

The equivalence to the above definition is obvious. However from this point of view, the definition of level is rather obscure.

We shall see in Section 5 that not all subgroups of Δ are congruence subgroups. For the moment, we justify the remark in Section 1 that congruence subgroups of Δ do not in general lift to congruence subgroups of Γ.

EXAMPLE. Let $P = 2\mathbb{Z}[\mu]$. By Theorem 2.1(iii), P is prime and $N(P) = 2^3$. By Corollary 3.5(ii),

$$|\Delta : \Delta(P)| = 504.$$

Now,

$$\Delta \cong \Gamma / G$$

where G denotes the normal closure of $(tp)^7$ in Γ, so that each lifted subgroup has level 7 (in the sense of Γ). From Wohlfahrt's Theorem, [**13**], this will be congruence only if it contains $\Gamma(7)$. Since $|\Gamma : \Gamma(7)| = 168$, $\Delta(P)$ does *not* lift to a congruence subgroup.

4. The level-index inequality. In this section, we obtain an analogue for Δ of results in [**12**].

NOTATION. (i) For $\alpha \in \mathbb{Z}[\mu]$, define E_k, $V_k (k = 1, 2, 3)$ by

$$E_1(\alpha) = \begin{pmatrix} 1+\alpha & \alpha \\ -\alpha & 1-\alpha \end{pmatrix} = I + \alpha V_1,$$

$$E_2(\alpha) = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \qquad = I + \alpha V_2,$$

$$E_3(\alpha) = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \qquad = I + \alpha V_3.$$

(ii) Let P be a prime ideal of $\mathbb{Z}[\mu]$. Suppose that π generates P over P^2, and that $\{\alpha_1 = 1, \ldots, \alpha_s\}$ is a basis for $\mathbb{Z}[\mu]/P$ over its prime subfield. For $r \geq 0$, we put

$$U(P, r) = \{E_k(\pi^r \alpha_l) : k = 1, 2, 3; l = 1, \ldots, s\}.$$

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

REMARK. The matrices $E_k(\alpha)$ $(\alpha \neq 0)$ are parabolic. They play a rôle in the present theory much as in that of Γ, though, of course, Δ has no parabolic elements.

LEMMA 4.1. *Let A be a non-zero ideal of $\mathbb{Z}[\mu]$, and $\alpha, \beta, \gamma \in t$. Then, for $k, l, m \in \{1, 2, 3\}$, $n \in \mathbb{N}$,*

(i) $E_k(\alpha)^n = E_k(n\alpha)$,

(ii) $E_k(\alpha)E_l(\beta)E_m(\gamma) \equiv I + \alpha V_k + \beta V_l + \gamma V_m (\mathrm{mod}\, A^2)$

These are trivial exercises.

LEMMA 4.2. *Let P be a prime ideal of $\mathbb{Z}[\mu]$ and $r \geq 1$. Then $\overline{\Delta}(P^r)$ is generated over $\overline{\Delta}(P^{r+1})$ by cosets corresponding to matrices of $U(P, r)$.*

Proof. This follows from the proof of Theorem 3.4 (giving the structure of $\overline{\Delta}(P^r)/\overline{\Delta}(P^{r+1})$), and Lemma 3.1, once we note that $(P^r)^2$ is a multiple of P^{r+1}.

LEMMA 4.3. *Let P be a prime ideal in $\mathbb{Z}[\mu]$. Then $\overline{\Delta}$ is generated over $\overline{\Delta}(P)$ by cosets corresponding to matrices of $U(P, 0)$.*

Proof. After Theorem 3.3, it is enough to show that the elements of $U(P, 0)$ generate $SL_2(\mathbb{Z}[\mu]/P)$.

Suppose that $X = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}[\mu]/P)$. By considering $E_2(1)X$ if necessary, we can assume that $a + c \neq 0 (\mathrm{mod}\, P)$.

Now $Y = E_1(\alpha)X$ has (1, 1)-entry $a + \alpha(a + c)$. By chosing α appropriately Y has (1, 1)-entry congruent to 1 modulo P. Then $Z = E_3(\beta)Y$ has (2, 1)-entry congruent to 0 modulo P, where $B = -Y_{2,1}$. As $\det(Z) \equiv 1 (\mathrm{mod}\, P)$, Z has the form

$$Z = \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix},$$

i.e. $Z = E_2(\gamma)$.

THEOREM 4.4. *If $G \leq \Delta$ is a congruence subgroup of level A and index n, then*

$$n^3 \geq N(A).$$

Proof. Suppose that A has the prime decomposition

$$A = (2 - \mu)^{r(0)} \prod_{k=1}^{R} P_k^{r(k)},$$

where $r(0) \geq 0, r(k) > 0 (k = 1, \ldots, R)$. Let $p_0 = 7, P_0 = (2 - \mu)$, and, for $k = 1, \ldots, R$, let p_k be the rational prime contained in P_k. Note that, because of the possibility of a rational prime splitting in $\mathbb{Z}[\mu]$, the p_k need not be distinct. We choose a subset S of $\{P_k\}$ as follows:

Case 1: $P_k = p_k\mathbb{Z}[\mu]$. Then $P_k \in S$. Here $N(P_k) = p_k^3$.

Case 2: $p_k\mathbb{Z}[\mu] = T_1T_2T_3$. For $l = 1, 2, 3$, let $s(l)$ be the greatest power of T_l dividing A. Take l with $s(l)$ maximal. Then $T_l \in S$. The other conjugates of P_k do *not* belong to S. Here $N(T_l) = p_k$, and $N(T_1T_2T_3) = p_k^3$.

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

We have

$$N(A) = 7^{r(0)} \prod_{k=1}^{R} N(P_k)^{r(k)} \leq 7^{r(0)} \prod_{p_k \in S} P_k^{3r(k)} = L(A), \text{ say,}$$

with equality only when conjugate P_k all occur to equal powers. We now show that $n^3 \geq L(A)$.

For $P_k \in S$, we observe that G is not of level A/P_k, so that 3.7, 4.2 and 4.3 show that we can find $X_k \in \overline{\Delta}$ with

(i) $X_k \in \overline{\Delta}(A/P_k^{r(k)})$,

(ii) X_k corresponds modulo $P_k^{r(k)}$ to an element of the set $U(P_k, r(k) - 1)$,

(iii) $\{\pm X_k\} \notin G$.

As $P_k \neq (2 - \mu)$, P_k is unramified and we may take p as the element "π" in the definition of $U(P_k)$. Then X_k corresponds to $E_l(P_k^{r(k)-1}\alpha)$, with $l \in \{1, 2, 3\}$, $\alpha \neq 0 (\mathrm{mod}\, P_k)$. Let $Y_k \in \overline{\Delta}(A/P_k^{r(k)})$ correspond to $E_l(\alpha)$ in $U(P_k, 0)$. Then $\{\pm Y_k\}^m \in G$ if and only if $P_k^{r(k)}$ divides m.

If $r(0) = 0$, we put $Y_0 = I$ and $s = 0$.

If $r(0) > 0$, then we can choose an X_0 as above. Here however, P_0 generates P_0^4 over P_0, so p_0 is not a suitable "π". Choose u such that $u \equiv r(0) - 1 (\mathrm{mod}\, 3)$ and $u \in \{0, 1, 2\}$, and let $s = (r(0) + 2 - u)/3$. Then we can choose $Y_0 \in \overline{\Delta}(A/P_0^{r(0)})$ such that $\{\pm Y_0\}^m \in G \Leftrightarrow p_0^s$ divides m. Note that $s \geq r/0$, corresponding to an element of $U(P_0, u)$.

Finally, let $Y = Y_0 \prod_{P_k \in S} Y_k$. Since P_0 and the P_k (for $P_k \in S$) are co-prime, we see that $\{\pm Y\}^m \in G$ if and only if p_0^s and each $p_k^{r(k)}$ ($P_k \in S$) divides m. Thus G has index at least $p_0^s \prod_{P_k \in S} p_k^{r(k)} = L(A)^{1/3}$.

DEFINITIONS. For $n \in \mathbb{N}$, let $M(n)$ (resp. $M_c(n)$) be the number of subgroups (resp. congruence subgroups) of index n in Δ.

As in [**1**], the level-index inequality allows us to obtain an upper bound for $M_c(n)$.

THEOREM 4.5. *For $n \in \mathbb{N}$,*

$$M_c(n) \leq n^3 \cdot 3^{3 \log_2 n} \cdot n^{81 \log_2 n}.$$

Proof. Suppose that n is a positive integer, and that G is a congruence subgroup of level A and index n.

From (b), we see that G corresponds in a unique way to a subgroup H of $\Delta/\Delta(A)$. From 3.5(ii), the latter has order at most $N(A)^3$. Then, by 4.4, the order is at most n^9. From [**12**, Lemma 3.2] there are at most $n^{9 \log_2 n}$ possible H (though many of these have the wrong index).

Now we must consider the number of possible ideals A. Again by 4.4, we have $N(A) \leq n^3$. Let m be any integer in the range $2, \ldots, n^3$ (the case $n = 1$ is trivial). The integer m has at most $\log_2 m$ prime factors. At worst, each corresponds to a split ideal of $\mathbb{Z}[\mu]$, so there are three possible primes in $\mathbb{Z}[\mu]$ with each factor as norm. Hence the number of ideals A with norm m is at most $3^{\log_2 m} \leq 3^{3 \log_2 n}$. There are fewer than n^3 possible m, so the total number is at most

$$n^3 3^{3 \log_2 n}.$$

The result now follows.

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

THEOREM 4.6. *For $n \geq 335$,*

$$M(n) \geq 2^{(n-335)/42}.$$

Proof. This proof relies heavily on the ideas of [**11**]. In particular, we need the notions of *coset diagram*, (l)-polygon ($l = 1, 2, 3$), (1)-*composition, (half-)open diagram* and *specification*. We also require the *genus formula*, i.e. for a subgroup of index n in Δ, there are non-negative integers p, e, f and g such that

$$n = 84(p-1) + 21e + 28f + 36g. \tag{7}$$

Suppose that $n \geq 294$. Let f_0 (resp. g_0) be the least non-negative residue of n modulo 3 (resp. 7). Then $f_0 \leq 2$, $g_0 \leq 6$. Let $p_0 = 0$. Then we have

$$\begin{aligned} n' = n - 84(p_0 - 1) - 28f_0 - 36g_0 &\geq n + 84 - 56 - 216 \\ &\geq 64. \end{aligned}$$

Also, $n' \equiv 0 (\mathrm{mod}\, 3)$ and $n' \equiv 0 (\mathrm{mod}\, 7)$, so $21 | n'$. It follows that there is an integer $e_0 \geq 4$ such that (u, p_0, e_0, f_0, g_0) satisfies (7).

It follows that there is a half-open diagram E for Δ with n points. Recall (from [**10**]) that a half-open diagram either

(a) has at least one (1)-polygon, or

(b) is the result of (1)-composition.

There is a 42-point diagram E with three (1)-polygons. In case (a), we can (1)-compose E and X to get a diagram X' with at least two (1)-polygons. In case (b), we can "undo" the (1)-composition in X, then compose each of the resultant (1)-polygons with a (1)-polygon of E to get a diagram X' with at least one (1)-polygon (from E).

Hence, if $n \geq 294$, we can find an n-point diagram for Δ which has at least one (1)-polygon.

For $n = 294, \ldots, 335$, we choose a diagram $H(n)$ of the above type having n points. In each diagram, we designate a vertex and one (1)-polygon. In E, we choose two of the (1)-polygons.

We also have (from [**11**]) a 42-point diagram U which has two (1)-polygons and one (3)-polygon.

Suppose that $n \geq 335$. We can find an *integer* m such that $m \equiv n (\mathrm{mod}\, 42)$ and $294 \leq n - 42m \leq 335$.

We take the diagram $H(n - 42m)$ and *choose* a sequence of m diagrams from $\{E, U\}$. Using (1)-composition we join the sequence to the designated (1)-polygon of $H(n - 42m)$. Taking the designated vertex of $H(n - 42m)$ as that of the new diagram, we get an n-point diagram. It is clear that different sequences will lead to different subgroups of index n in Δ. Since $m \geq (n - 335)/42$, the result follows.

COROLLARY 4.7. *As $n \to \infty$,*

$$\frac{M_c(n)}{M(n)} \to 0.$$

This is clear from the estimates of 4.5, 4.6.

COROLLARY 4.8. *There exists an integer n_0 such that, for $n \geq n_0$, Δ has a non-congruence subgroup of index n.*

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

A (large) estimate for n_0 could be obtained from 4.5 and 4.6.

In the next section, we obtain *explicit* non-congruence subgroups.

5. Non-congruence subgroups.

THEOREM 5.1. *Suppose that* $A = \prod_{k=1}^{R} P_k^{r(k)}$ *is a non-zero ideal of* $\mathbb{Z}[\mu]$. *Then the composition factors of* $\Delta/\Delta(A)$ *consist of*

(a) *the groups* $PSL_2(\mathbb{Z}[\mu]/P_k)$ $(k = 1, \ldots, R)$,

(b) *cyclic groups of order* p_k, *where* p_k *is the rational prime contained in* P_k,

(c) *cyclic groups of order* 2.

Proof. We observe that we have a normal series

$$\frac{\Delta}{\Delta(A)} = \frac{G_0}{\Delta(A)} \triangleright \cdots \triangleright \frac{G_l}{\Delta(A)} \cdots \triangleright \frac{G_R}{\Delta(A)} \triangleright \langle 1 \rangle$$

where $G_l = \cap_{k=1}^{l} \Delta(P_k^{r(k)})$ $(l = 1, \ldots, R)$

Now, for $l = 1, \ldots, R$,

$$\frac{G_{l-1}}{G_l} \cong \frac{G_{l-1}}{G_{l-1} \cap \Delta\left(P_l^{r(l)}\right)}, \cong \frac{G_{l-1} \cdot \Delta\left(P_l^{r(l)}\right)}{\Delta\left(P_l^{r(l)}\right)} = \frac{\Delta}{\Delta\left(P_l^{r(l)}\right)}.$$

The final equality occurs for $l > 1$ by Theorem 3.6 since $G_{l-1} \supset \Delta(\prod_{k=1}^{l-1} P_k^{r(k)})$. It is obvious for $l = 1$. Using the ideas of 3.4, we can refine the factor $\Delta/\Delta(P_l^{r(l)})$ into $PSL_2(\mathbb{Z}[\mu]/P_l)$ and a number of p_l-cycles (if $r(l) > 1$).

There remains the factor $G_R/\Delta(A)$. Any element of G_R has the form $\{\pm X\}$, where, for $k = 1, \ldots, R$,

$$\phi_{pl}(X) \equiv \pm I \left(\mathrm{mod}\, P_l^{r(l)}\right).$$

Then $(\pm X^2) \in \Delta(A)$, so each element of $G_R/\Delta(A)$ has order 1 or 2. Hence $G_R/\Delta(A)$ is a product of 2-cycles.

DEFINITION. If $G \leq \Delta$, then $C(G)$, the *core* of G, is the greatest Δ-normal subgroup contained in G.

REMARK. If $G \leq \Delta$ is a congruence subgroup of level A, then $C(G) \supseteq \Delta(A)$ since the former is maximal.

COROLLARY 5.2. *For* $n \geq 168$, *there is a non-congruence subgroup of index* n *in* Δ.

Proof. Let $n \geq 168$. From [**2**], there is a subgroup G of index n in Δ such that $\Delta/C(G) \cong A_n$ (the alternating group).

Suppose that G is congruence of level A. By the above remark, $\Delta(A) \subseteq C(G)$, and we have the normal series

$$\frac{\Delta}{\Delta(A)} \triangleright \frac{C(G)}{\Delta(A)} \triangleright \langle 1 \rangle.$$

Since $\Delta/C(G)$ is simple it is a composition factor, contradicting 5.1.

Conder gives also the list of smaller n for which such a G exists (and, by the above proof, each G is non-congruence). The smallest index is 15. In fact, there are non-congruence groups of lower index already in the literature [**9**], as we now see.

https://doi.org/10.1017/S0017089506003223 Published online by Cambridge University Press

EXAMPLE. (Sinkov's subgroups).

Here it is convenient to use a permutation representation of subgroups. Suppose that $G \leq \Delta$ has index n and cosets $G = GX_1, \ldots, GX_n$. For $X \in \overline{\Delta}$, we write $\hat{X}$ for the permutation of coset suffices induced by post-multiplication by $\{\pm X\}$ we recall that

(i) $\{\pm X\} \in G \Leftrightarrow \hat{X}$ fixes 1,

(ii) $\langle \hat{U}_2, \hat{U}_3 \rangle \cong \Delta / C(G)$.

(ii) $\langle \pm X \rangle \in C(G) \Leftrightarrow \hat{X}$ fixes $(1), (2), \ldots, (n)$.

We note that, as $\{\pm U_2\}\{\pm U_3\}$ generate Δ, it is sufficient to give $\hat{U}_2$ and $\hat{U}_3$.

One of the subgroups in [**9**] corresponds to the permutations

$$\hat{U}_2 = (1, 2)(3, 14)(4, 5)(6)(7, 8)(9, 11)(10)(12, 13),$$

$$\hat{U}_3 = (1)(2, 3, 4)(5, 6, 7)(8, 9, 10)(11, 12, 14)(13).$$

Call this subgroup G. It is easy to check that $\langle \hat{U}_2, \hat{U}_3 \rangle$ is imprimitive, with blocks

$$\{1, 13\}, \{2, 12\}, \{3, 14\}, \{4, 11\}, \{5, 9\}, \{6, 10\}, \{7, 8\}.$$

Thus G has an overgroup H obtained from the action of $\hat{u}_2$ and $\hat{u}_3$ on the blocks. Clearly, H has index 7 in Δ. Using the method of [**11**], we can show that Δ has exactly 2 conjugacy classes of subgroups of index 7, and both consists of congruence subgroups of level $(2 - \mu)$. (They correspond to subgroups of $PSL_2(GF(7))$ isomorphic to S_4.) Since $\Delta / \Delta(2 - \mu))$ is simple, $C(H) = \Delta((2 - \mu))$.

An easy calculation, using (iii) above, shows that, if $\hat{Z} = \hat{U}_3^2 \hat{U}_2 \hat{U}_3 \hat{U}$ then $\hat{Z}$ has order 8 in $\Delta / C(G)$, but 4 in $\Delta / C(H)$. Thus $C(G) \not\supseteq C(H)$. In fact, $C(G)$ has index 1344 (see [**9**]).

Suppose that G is congruence of level A. From the argument of 4.4, since $(2 - \mu) | A$, we must have $A | 14\mathbb{Z}[\mu]$. Now we observe that $\Delta((2 - \mu))$ is generated over $\Delta(A)$ by elements of odd order. Let $\{\pm X\}$ be such an element. Then $\{\pm X\} \in \Delta((2 - \mu)) \subseteq H$ and $|H : G| = 2$, so $\{\pm X^2\} \in G$. Since $\{\pm X^2\}$ has the same order over $\Delta(A)$. Thus we would have $\Delta((2 - \mu)) \subset G$, and hence in $C(G)$. Since $\Delta((2 - \mu))$ has index 168 while $C(G)$ has index 1344, this gives a contradiction.

Thus Sinkov's subgroups are non-congruence.

6. Postscript. Other $(2, 3, n)$-groups can be obtained from quaternion algebras as in § 2. The key to our results is O'Meara's result (Theorem 3.2) which need the triangle group to be the *entire* set of unimodular matrices in the maximal order. As stated in [**7**], this occurs only for $n = 7, 9, 11$, so only for these can we expect to apply the methods of this paper.

We hope to show elsewhere that it is possible to prove Theorem 3.3 for a wider range of triangle groups (viz. those with n prime to 30), but this requires a new approach. Once an analogue of 3.3 is available, much of the present theory will go through.

REFERENCES

1. J. Cohen, On covering Klein's curve and generating projective groups, *The geometric vein* (Springer-Verlag, 1981), 511–18.

2. M. Conder, Generators for alternating and symmetric groups, *J. London Math. Soc* (2), **22** (1980), 75–86.

3. M. Eichler, *Lectures on modular correspondences*, Tata Institute of Fundamental Research, 1955–56.

4. R. Fricke, Über den arithmetischen Character der Verzweigungen (2, 3, 7) und (2, 4, 7) gehörigen Dreiecksfunktionen, *Math. Ann.* **41** (1893), 443–468.

5. A. M. Macbeath, Generators of the fractional linear groups, *Number Theory, Proc Sympos. Pure Math.* **12**, 1967 (Amer. Math. Soc., 1969), 14–32.

6. O. T. O'Meara, *Introduction to quadratic forms.* Die Grundlehren der mathematischen Wissenschaffen **117** (Springer-Verlag, 1963).

7. G. Shimura, *Automorphic forms and number theory*, Lecture Notes in Mathematics No. 54 (Springer-Verlag, 1968).

8. G. Shimura, Construction of class fields and zeta functions of algebraic curves, *Ann. of Math.* **85** (1967), 58–159.

9. A. Sinkov, On the group-defining relations (2, 3, 7; *p*), *Ann. of Math.* **38** (1937), 577–584.

10. W. W. Stothers, Subgroups of the modular group, *Proc. Cambridge Phil. Soc.* **75** (1974), 139–153.

11. W. W. Stothers, Subgroups of the (2, 3, 7) triangle group, *Manuscripta Math.* **20** (1977), 323–334.

12. W. W. Stothers, Level and index in the modular group, *Proc. Roy. Soc. Edinburgh* **99A** (1984), 115–126.

13. K. Wohlfahrt, An extension of F. Klein's level concept, *Illinois J. Math.* **8** (1964), 529–535.

14. O. Zariski and P. Samuel, *Commutative Algebra* (Vol. 1) (van Nostrand, 1960).