and D_0 is the discriminant of the field $Q(\sqrt{k_1}, \sqrt{k_2})$. Furthermore, the classical estimate of fundamental units in real quadratic fields (cf. Schur [11]) gives

$$\overline{|a_j|} \leqslant k_j^{\sqrt{k_j}} \qquad (j = 1, 2),$$

and it is readily verified that

$$|D_0|^{1/2} \leqslant 64 k_1 k_2 |k_1 - k_2|.$$

Thus we obtain $a_0 < c_7 d^{(1/16)\delta} \log d$, whence $\log H = d^{(1/8)\delta}$ satisfies the above condition.

It is clear that the inequalities are consistent only if $d < c_8$, and this completes the proof of Theorem 2.

We conclude by expressing our thanks to Professor Heilbronn for pointing out a mistake in an earlier draft of this paper.

References

[1] A. G. Anferteva and N. G. Čudakov, *Minima of the norm function in imaginary quadratic fields*, Dokl. Akad. Nauk SSSR **183** (1968), pp. 255–256; = Soviet Math. Dokl. 9 (1968), pp. 1324–1344; see also Mat. Sb. 82 (1970), pp. 55–66.

[2] A. Baker, *A remark on the class number of quadratic fields*, Bull. London Math. Soc. 1 (1969), pp. 98–102.

[3] D. A. Burgess, *On character sums and L-series II*, Proc. London Math. Soc. 13 (1963), pp. 524–536.

[4] L. E. Dickson, *Introduction to the theory of numbers*, Chicago 1929.

[5] H. Heilbronn, *On the class number in imaginary quadratic fields*, Quarterly J. Math. Oxford 5 (1934), pp. 150–160; see Lemma XIV, p. 158.

[6] A. E. Ingham, *The distribution of prime numbers*, Cambridge 1932.

[7] E. Landau, *Vorlesungen über Zahlentheorie I*, Leipzig 1927.

[8] Yu. V. Linnik and A. I. Vinogradov, *Hyperelliptic curves and the least prime quadratic residue*, Doklady Akad. Nauk SSSR 166 (1966), pp. 259–261; = Soviet Math. Dokl. 7 (1966), pp. 612–614.

[9] G. B. Mathews, *Theory of numbers*, Cambridge 1892.

[10] A. Schinzel, *On two theorems of Gelfond and some of their applications*, Acta Arith. 13 (1967), pp. 177–236.

[11] I. Schur, *Einige Bemerkungen zu der vorstehenden Arbeit des Herrn G. Pólya: Über die Verteilung der quadratischen Reste und Nichtreste*, Göttingen Nachrichten (1918), pp. 30–36.

[12] C. L. Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arith. 1 (1935), pp. 83–86; = Ges. Abhandlungen I, pp. 406–409.

Received on 25. 1. 1970

ACTA ARITHMETICA
XVIII (1971)

Subgroups of the modular group generated by parabolic elements of constant amplitude

by

R. A. Rankin (Glasgow)

Dedicated to the memory of Harold Davenport

1. Let $\Gamma(1) = \mathrm{SL}(2, \mathbf{Z})$ denote the modular group of unimodular 2×2 matrices with rational integral entries and put

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \tag{1}$$

For each positive integer n we write $\Delta(n)$ for the normal closure of U^n; it is the subgroup of $\Gamma(1)$ generated by the conjugate parabolic matrices $L^{-1} U^n L (L \in \Gamma(1))$. Further, $\Delta(n)$ is a subgroup of the principal congruence group

$$\Gamma(n) = \{T \in \Gamma(1)\colon\ T \equiv I (\mathrm{mod}\, n)\}.$$

The purpose of this paper is to provide a new proof of the well known

THEOREM. (i) *For* $1 \leqslant n \leqslant 5$, $\Delta(n) = \Gamma(n)$. (ii) *For* $n \geqslant 6$, *the index* $[\Gamma(n)\colon\ \Delta(n)] = \infty$.

If G is any subgroup of $\Gamma(1)$, we denote by $\hat{G}$ the corresponding inhomogeneous group. Thus $\hat{\Gamma}(1) = \mathrm{LF}(2, \mathbf{Z})$ and $\hat{G}$ is the image of G under the natural mapping from $\Gamma(1)$ to $\Gamma(1)/\Lambda$, where $\Lambda = \{I, -I\}$ is the centre of $\Gamma(1)$. A corresponding theorem holds for $\hat{\Delta}(n)$ and $\hat{\Gamma}(n)$.

As pointed out by Knopp [5], who gave an independent proof of part (ii) of the theorem, the results stated can be excavated by the persevering reader from the first volume of Klein and Fricke's monumental treatise ([4], pp. 354–360). An alternative proof of part (i) of the theorem has been given by Brenner ([1], pp. 215–217). The quickest and most elegant proof of the theorem is obtained by using the canonical presentation of a Fuchsian group; see Wohlfahrt [10]. For an application of properties of tesselation groups to prove part (ii), see Mennicke [6].

The proof of part (i) given below uses elementary properties of indefinite binary quadratic forms. It is, perhaps, worth pointing out that it has the merit of providing a constructive method of expressing an arbitrary element of $\Gamma(n)$ as a product of elements of $\Delta(n)$. The fact that $\Gamma(n)$ is a free group when $n > 2$ is used in the second part of the theorem, which follows from the stronger result that $\Delta(n)\,\Gamma'(n)$ has infinite index in $\Gamma(n)$; here $\Gamma'(n)$ is the derived group.

The proof which follows appeared originally in lecture notes [9] of a course of lectures on the modular group given in Madras in September, 1968.

2. Proof of the first part of the theorem. We first dispose of the case $n = 1$. Since $\Gamma(1)$ is generated by the two parabolic matrices U and

$$W = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = V^{-1}UV, \quad \text{where} \quad V = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \tag{2}$$

we deduce that $\Gamma(1) = \Delta(1)$.

Now let $n \geqslant 2$ and take any

$$S = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in \Gamma(n). \tag{3}$$

The idea of the proof is to multiply S on the left by an element of $\Delta(n)$ so as to obtain a matrix with a smaller trace and so, after a finite number of steps, reach a matrix T of trace ± 2. If this is possible, it will then suffice to prove that $T \in \Delta(n)$. For such a matrix T is necessarily of the form $\varepsilon L^{-1} U^q L$, where $\varepsilon = \pm 1$ and $q \in \mathbf{Z}$, the set of rational integers. Since $T \in \Gamma(n)$, it follows that $q \equiv 0 \pmod{n}$ and that $\varepsilon = 1$ when $n > 2$. Since

$$-I = W^2 \cdot U^{-1} W^2 U \cdot U^{-2} \in \Delta(2),$$

we deduce that $T \in \Delta(2)$ when $n = 2$.

Write $t_0 = \operatorname{tr} S$, so that $t_0 \equiv 2 \pmod{n}$. Since $\Gamma(n)$ contains no elliptic matrices when $n \geqslant 2$, we must have $|t_0| \geqslant 2$ and, in view of the foregoing remarks, we can assume that $|t_0| > 2$.

Take any

$$L = \begin{bmatrix} v & u \\ y & x \end{bmatrix} \in \Gamma(1), \tag{4}$$

so that x and y can be any coprime integers. Then, for $r \in \mathbf{Z}$,

$$L^{-1} U^{nr} L = \begin{bmatrix} 1 + nrxy & nrx^2 \\ -nry^2 & 1 - nrxy \end{bmatrix},$$

and we put

$$S_1 = L^{-1} U^{nr} L \cdot S, \tag{5}$$

so that, as is easily verified,

$$t_1 = \operatorname{tr} S_1 = t + nrQ_S(x, y), \tag{6}$$

where

$$Q_S(x, y) = \gamma x^2 + (\alpha - \delta)xy - \beta y^2. \tag{7}$$

This is an indefinite quadratic form of discriminant

$$(\alpha - \delta)^2 + 4\beta\gamma = t_0^2 - 4 > 0.$$

Now, if $Q(x, y)$ is any indefinite binary quadratic form of discriminant D, coprime integers x, y can be found such that

$$0 < |Q(x, y)| \leqslant \sqrt{D/5}; \tag{8}$$

moreover, the minimum positive value of $|Q(x, y)|$ is equal to $\sqrt{D/5}$ if and only if $Q(x, y)$ is equivalent to the form

$$q(x^2 + xy - y^2) \tag{9}$$

for some real q. If this is not the case, then coprime integers x, y can be found such that

$$0 < |Q(x, y)| \leqslant \sqrt{D/8}. \tag{10}$$

See, e.g., J. W. S. Cassels [2]. The form (9) is the first special form in the co-called Markoff chain of forms.

In this connexion, two quadratic forms are said to be equivalent when one can be transformed into the other by a change of variables

$$(x, y) \to (dx + by,\ cx + ay),$$

where $a, b, c, d \in \mathbf{Z}$,

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad \det T = \pm 1.$$

Now it is easily verified that

$$Q_{TST^{-1}}(x, y) = \det T \cdot Q_S(dx + by,\ cx + ay). \tag{11}$$

This is well known; see [3], for example. Hence, if Q_S is equivalent to a form Q, we may choose T so that

$$Q_{TST^{-1}}(x, y) = \det T \cdot Q(x, y). \tag{12}$$

We now return to (5) and (6). It follows from (8) that we can find coprime integers x, y to make

$$0 < |Q_S(x, y)| \leqslant \{(t_0^2 - 4)/5\}^{1/2},$$

and, by (6), we can then choose r to make

$$|t_1| \leqslant \tfrac{1}{2} n |Q_S(x, y)| \leqslant n\{(t_0^2 - 4)/20\}^{1/2}. \tag{13}$$

Now $n\{(t_0^2 - 4)/20\}^{1/2} < |t_0|$ for $n = 2, 3$ and 4, so that $|t_1| < |t_0|$. In this way, by left multiplication by elements of $\Delta(n)$, we can reduce the trace successively, when $n = 2, 3$ or 4, and so reach an element of trace ± 2; as indicated above, this shows that $S \in \Delta(n)$. Observe that, in these three cases, we only require the result (8), which can be proved quite simply.

For the remainder of this section we assume that $n = 5$ and form S_1 from S as before; see (5). There are now two possibilities: either (i) $Q_S(x, y)$ is equivalent to the form (9), for some real q, or (ii) coprime x and y can be chosen to make

(14) $$|t_1| \leqslant \tfrac{5}{2}\,|Q_S(x, y)| \leqslant \tfrac{5}{2}\,\{(t_0^2-4)/8\}^{1/2} < |t_0|.$$

In case (i) we shall show directly that $S \in \Delta(5)$. If case (ii) holds, then (14) applies and, if $|t_1| > 2$, we premultiply S_1 by $L_1^{-1}U^{5r_1}L_1$ $(L_1 \in \Gamma(1), r_1 \in \mathbf{Z})$ and proceed as before. At each stage there are two possibilities and, after a finite number of steps, we reach either a matrix associated with a quadratic form equivalent to (9) or a matrix of trace ± 2, and so deduce that $S \in \Delta(5)$.

It therefore remains to show that $S \in \Delta(5)$ when $Q_S(x, y)$ is equivalent to the form (9). When this is the case we can choose a matrix T as above (see (11) and (12)), so that

(15) $$Q_{TST^{-1}}(x, y) = q(x^2+xy-y^2),$$

where, clearly, $q \in \mathbf{Z}$. Write

$$S_0 = TST^{-1} = \begin{bmatrix}\alpha_0 & \beta_0\\ \gamma_0 & \delta_0\end{bmatrix},$$

so that $S_0 \in \Gamma(5)$; note that both $\Gamma(5)$ and $\Delta(5)$ are invariant under the (possibly outer) automorphism $S \to TST^{-1}$. It therefore suffices to show that $S_0 \in \Delta(5)$. By (15),

$$\gamma_0 = \beta_0 = \alpha_0 - \delta_0 = q,$$

so that we obtain the Pellian equation

(16) $$t_0^2 = (\alpha_0+\delta_0)^2 = 5q^2+4.$$

The complete set of solutions of (16) is given by

$$t_0 = \varepsilon(u_{2m+1}+u_{2m-1}), \qquad q = \varepsilon u_{2m} \qquad (m \in \mathbf{Z}),$$

where $u_0 = 0, u_1 = 1, u_2 = 1, \ldots$ is the Fibonacci sequence and $\varepsilon = \pm 1$. Note that

$$u_{-m} = (-1)^{m-1}u_m \quad \text{and} \quad u_m+u_{m+1} = u_{m+2} \qquad (m \in \mathbf{Z}).$$

It follows that, for some $m \in \mathbf{Z}$,

(17) $$S_0 = \varepsilon\begin{bmatrix}u_{2m+1} & u_{2m}\\ u_{2m} & u_{2m-1}\end{bmatrix} = \varepsilon A^m,$$

where

$$A = UW = \begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}.$$

Now

$$A^2 = \begin{bmatrix}5 & 3\\ 3 & 2\end{bmatrix}, \quad A^3 = \begin{bmatrix}13 & 8\\ 8 & 5\end{bmatrix}, \quad A^4 = \begin{bmatrix}34 & 21\\ 21 & 13\end{bmatrix}, \quad A^5 = \begin{bmatrix}89 & 55\\ 55 & 34\end{bmatrix},$$

so that $-A^5 \equiv I \pmod 5$. Since $S_0 \in \Gamma(5)$, we deduce from this and (17) that

$$S_0 = (-A^5)^r \quad \text{for some } r \in \mathbf{Z}.$$

Accordingly, we need only show that $-A^5 \in \Delta(5)$ to complete the proof. In fact

(18) $$-A^5 = U^5U_1^5U_2^5U_3^5U_4^5U^{*5},$$

where

$$U_m = L_m^{-1}UL_m, \quad L_m = VU^{-m}V^{-1}U^2 \quad (1 \leqslant m \leqslant 4)$$

and

$$U^* = L^{*-1}UL^*, \quad L^* = VU^2;$$

see (2). The result follows.

In conclusion, we note that the generators for $\Delta(n)$ and therefore $\Gamma(n)$, for $n \leqslant 5$, can be found by using the fact that in each matrix $L^{-1}U^nL$, L is a product of the elements U, W and their inverses. Thus, when $n = 2$, since

$$U^{-1}W^2U = -W^{-2}U^2 \quad \text{and} \quad WU^2W^{-1} = -W^2U^{-2},$$

we conclude that

$$\Gamma(2) = \Delta(2) = \langle -I, U^2, W^2\rangle.$$

Similarly,

$$\Gamma(3) = \Delta(3) = \langle U^3, P^{-1}U^3P, P^{-2}U^3P^2\rangle,$$

where $P = VU$; note that $P^3 = -I$.

3. Proof of the second part of the theorem. We prove that $[\hat{\Gamma}(n): \Delta(n)\hat{\Gamma}'(n)] = \infty$ for $n \geqslant 6$; from this the second part of the theorem follows since each of the three inhomogeneous groups is isomorphic to the corresponding homogeneous group under the natural homomorphism.

It is an easy consequence of the Kurosh subgroup theorem that every subgroup of $\hat{\Gamma}(1)$ of index $\mu > 3$ is free; see Newman [7], or [9], § 4. From this and the fact that $\hat{\Gamma}'(1)$ has index 6 in $\hat{\Gamma}(1)$ it is easily deduced from Schreier's theorem that the rank of every free subgroup of $\hat{\Gamma}(1)$ of index μ is $1+\frac{1}{6}\mu$; this result was first stated by A. W. Mason.

Consider the group $\hat{\Gamma}(n)$ for $n \geqslant 2$. This is a free group; we write μ for its index and put $\mu = ns$ so that s is the parabolic class number. If $R\hat{\Gamma}(n)$ is any left coset, the n cosets $U^kR\hat{\Gamma}(n)$ $(0 \leqslant k < n)$ are easily seen to be disjoint. It follows that there exist s elements $R_\nu \epsilon \hat{\Gamma}(1)$ $(0 \leqslant \nu < s)$ such that the elements

$$U^kR_\nu \quad (0 \leqslant k < n,\ 0 \leqslant \nu < s)$$

run through all the μ cosets of $\hat{\Gamma}(n)$ in $\hat{\Gamma}(1)$.

Now take any $L \epsilon \hat{\Gamma}(1)$ so that $L = U^kR_\nu S$ for some $k < n$, $\nu < s$ and $S \epsilon \hat{\Gamma}(n)$. Then

$$L^{-1}U^nL = S^{-1}T_\nu S,$$

where

$$T_\nu = R_\nu^{-1}U^nR_\nu \quad (0 \leqslant \nu < s).$$

Hence $\hat{\Delta}(n)$ is generated by the elements

$$S^{-1}T_\nu S \quad \text{for} \quad S \epsilon \hat{\Gamma}(n) \quad \text{and} \quad 0 \leqslant \nu < s.$$

Now suppose that $A_1, A_2, \ldots, A_r$, where

$$r = 1+\frac{\mu}{6} = 1+\frac{ns}{6}, \tag{19}$$

are r independent generators of $\hat{\Gamma}(n)$ and consider the map

$$e\colon\ \hat{\Gamma}(n) \to \boldsymbol{Z}^r$$

defined by

$$e(S) = \big(e_1(S), e_2(S), \ldots, e_r(S)\big),$$

where $e_i(S)$ is the exponent sum of the generator A_i in the word S; see [8]. The kernel of the mapping e is $\hat{\Gamma}'(n)$. Since

$$e(L^{-1}U^nL) = e(S^{-1}T_\nu S) = e(T_\nu),$$

it follows that e maps $\hat{\Delta}(n)$ into a linear subspace H of $\boldsymbol{Z}^r$ of dimension not exceeding s. Note that $e^{-1}(H) = \hat{\Delta}(n)\hat{\Gamma}'(n)$.

Since $s < r$ when $n \geqslant 6$, by (19), it follows that

$$[\hat{\Gamma}(n)\colon\ \hat{\Delta}(n)\hat{\Gamma}(n)] = [\boldsymbol{Z}^r\colon\ H] = \infty.$$

References

[1] J. L. Brenner, *The linear homogeneous group, III*, Ann. of Math. 71 (1960), pp. 210–223.

[2] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge 1953.

[3] H. Cohn, *Approach to Markoff's minimal forms through modular functions*, Ann. of Math. 61 (1965), pp. 1–12.

[4] F. Klein und R. Fricke, *Vorlesungen über die Theorie der elliptischen Modulfunktionen*, Band 1, Leipzig 1890.

[5] M. I. Knopp, *A note on subgroups of the modular group*, Proc. Amer. Math. Soc. 14 (1963), pp. 95–97.

[6] J. L. Mennicke, *Finite factor groups of the unimodular group*, Ann. of Math. 81 (1965), pp. 31–37.

[7] M. Newman, *Free subgroups and normal subgroups of the modular group*, Illinois J. Math. 8 (1964), pp. 262–265.

[8] R. A. Rankin, *Lattice subgroups of the unimodular group*, Inventiones Math. 2 (1967), pp. 215–221.

[9] — *The modular group and its subgroups*, Madras 1969.

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

UNIVERSITY OF GLASGOW
Glasgow, Scotland

Received on 2. 2. 1970