SUBGROUPS OF INFINITE INDEX IN THE MODULAR GROUP II

by W. W. STOTHERS

(Received 8 July, 1979)

Let H be a subgroup of Γ , the modular group. Let h be the number of orbits of $\mathbb{Q} \cup \{\infty\}$ under the action of H. In each orbit, the stabilizers are H-conjugate. Let U be the mapping $z \mapsto z+1$. Each stabilizer is Γ -conjugate to $\langle U^c \rangle$ for some non-negative integer c. The integer c is the cusp-width of the orbit. Let h_0 be the number of orbits with non-trivial stabilizer, i.e. with c > 0. The sequence $(c(1), \ldots, c(h_0))$ of non-zero cusp-widths is the cusp-split of H. Clearly, $h_0 \leq h$, and $h_{\infty} = h - h_0$ is the number of orbits with trivial stabilizer.

We associate with H the list $(r, s, t_1, h_0, h, c(1), \ldots, c(h_0))$, the specification of H. The definitions of r, s, and t_1 are in [4]. If H has index u, not necessarily finite, and $h_{\infty} = 0$, then $h_0 = h > 0$ and

$$u = c(1) + \ldots + c(h_0).$$
 (1)

When u is finite, some positive power of each Γ -conjugate of U belongs to H so that $h_{\infty} = 0$. Further, r, s, and t_1 are non-negative integers with t_1 even and

$$u = 3r + 4s + 6(h_0 - t_1 - 2).$$
⁽²⁾

Given a list (u, r, s, t_1, h_0) satisfying (2) with $u, h_0 > 0$, and t_1 even, there is a corresponding subgroup of Γ , see [4] or [3].

Here we consider a list with $h_{\infty} = 0$ and ask whether it is the specification of a subgroup of infinite index. We obtain a necessary and sufficient condition which is the analogue of (1) and (2). In a further paper, we propose to consider cases with $h_{\infty} > 0$.

We shall use the various D-diagrams of [4] and the L-diagrams of [6]. We remind the reader that the red loops (edges) in [4] correspond to E_1 -loops (pairs of edges) in [6], and that the green edges in [4] are the P_1 -edges in [6], but are oppositely directed. It follows that the L-diagrams can be obtained from the D_1 -diagrams. The existence results are generalisations of the final theorems in [6].

We are indebted to the referee for the observation that our L-diagrams are graphs underlying the triangular maps discussed in [1], [2]. Our results and proofs can be restated in the language of maps.

To restrict attention to lists which satisfy known necessary conditions, we say that $(r, s, t_1, h_0, h, c(1), \ldots, c(h_0))$ is a specification provided

- (i) each of r, s, t_1 , h_0 , h is a non-negative integer or is ∞ ,
- (ii) for each *i*, c(i) is a positive integer,
- (iii) if $h_{\infty} = 0$, then $h_0 = \infty$ (see (1)),
- (iv) if $h_{\infty} = 1$ and $t_1 + h_0 < \infty$, then t_1 is even (see Theorem 4.4 of [4]),
- (v) $r+s+t_1+h_0+h=\infty$ (see Lemma 4.2 of [4]).

Glasgow Math. J. 22 (1981) 101-118.

In a specification, let h(k) be the number of values of i for which $c(i) \le k$.

1. L-diagrams and subdiagrams. Given a pseudograph P, we write $\mathcal{V}(P)$ (resp. $\mathcal{E}(P)$) for the set of vertices (resp. edges) of P. We use the same notation for an L-diagram, referring to the underlying pseudograph. For $V \in \mathcal{V}(P)$, v(V) denotes the degree of the vertex V. Since we assume that $h_{\infty} = 0$, we always have $v(V) < \infty$.

Let L be an L-diagram. A subdiagram E consists of a vertex set $\mathcal{V}(E) \subseteq \mathcal{V}(L)$, with the edge set containing all edges of L with both ends in $\mathcal{V}(E)$. We avoid contexts which involve free edges. Observe that a subdiagram is determined by its vertex set, i.e. by a subset of $\mathcal{V}(L)$.

The edges round each vertex of a subdiagram E are ordered by the ordering in the parent L-diagram. Hence E is a K-diagram. However, E need not be an L-diagram, e.g. if $\mathcal{V}(E)$ consists of 2 vertices which are joined by a single edge in the L-diagram but have no loops.

Suppose that E is a subdiagram of L. Then the completion of E, denoted \overline{E} , is the subdiagram with $\mathcal{V}(\overline{E}) = \mathcal{V}(E) \cup \mathcal{A}(E)$ where $\mathcal{A}(E)$ is the set of vertices in L which are adjacent to vertices of E. The boundary of E, denoted $\mathcal{B}(E)$, is the set of vertices of E which are adjacent to vertices of L - E. Note that each vertex of $\mathcal{B}(E)$ is adjacent to a vertex of $\mathcal{A}(E)$, and conversely. The centre of E, $\mathcal{C}(E)$, is $\mathcal{V}(E) - \mathcal{B}(E)$, the set of vertices of E all of whose L-neighbours belong to $\mathcal{V}(E)$.



Figure 1.

Figure 1 shows an L-diagram. Here, as later, we omit the labelling of the edges at each vertex since we draw our diagrams so that the ordering is anti-clockwise in each case. We define subdiagrams E and F as follows:

$$\begin{aligned} &\mathcal{V}(E) = \{W, X, Y, Z\}, \qquad &\mathcal{B}(E) = \mathcal{V}(E), \qquad &\mathcal{C}(E) = \emptyset, \qquad E = L, \\ &\mathcal{V}(F) = \mathcal{V}(E) \cup \{U\}, \qquad &\mathcal{B}(F) = \mathcal{V}(E), \qquad &\mathcal{C}(F) = \{U\}, \qquad \bar{F} = L. \end{aligned}$$

We recall that, in an L-diagram L, ABC is a triangle if AB, BC, CA in $\mathscr{C}(L)$ are such that BC follows AB at B, CA follows BC at C and AB follows CA at A. Note that a triangle may involve one, two or three different vertices. If a triangle involves less than three vertices, then at least one loop is involved (see Figure 5 of [6]).

Since an L-diagram has the triangle property, each edge is in precisely two triangles. (We are ignoring free edges.) Suppose that E is a subdiagram of L. A paired (resp. *unpaired*, *loose*) edge of E is an edge in two (resp. one, no) triangles of L lying entirely in E. Observe that an unpaired or loose edge must have both end points in $\mathscr{B}(E)$ since it belongs to at least one triangle with an edge not in E. In Figure 1, XY is an unpaired edge of E since it belongs to WXY and UXY, but $U \notin \mathscr{V}(E)$. It is a paired edge of F. Also, XZ is a loose edge of E and an unpaired edge of F.

In considering L-diagrams, it is convenient to avoid vertices which have loops. We state results for diagrams which have no loops, but it should be clear that they can be applied to regions of a general L-diagram where there are no loops.

LEMMA 1.1. (i) Suppose that L is an L-diagram without loops. If F is a non-empty subdiagram of L and $E = \overline{F}$, then any vertex P in $\mathcal{B}(E)$ is incident with at least two unpaired edges of E.

(ii) Suppose, further, that $P \in \mathcal{B}(E)$ and that $v(P) \leq 6$. Then P is on exactly two unpaired edges of E.

Proof. (i) Observe that, by the definition of completion, $\mathcal{V}(F) \subseteq \mathcal{C}(E)$. Hence there is a vertex $Q \in \mathcal{C}(E)$ with $QP \in \mathcal{C}(E)$ and, as $P \notin \mathcal{C}(E)$, a vertex $X \in \mathcal{V}(L-E)$ with $PX \in L-E$. In one direction round P, let the sequence of edges be $PA_0 (=PQ), PA_1, \ldots, PA_n (=PX)$. Suppose that PA_{i+1} is not in $\mathcal{C}(E)$. Then the triangle $PA_{i-1}A_i$ belongs to E, but PA_iA_{i+1} does not. Hence PA_i is unpaired. Note that $i \neq n$ as $X \notin \mathcal{V}(E)$.

Similarly, in the other direction round P, there is a first unpaired edge PB, say. As before, $PB \neq PX$.

If $PA_i = PB$, then each must be PQ since PA_i lies between PQ and PX in one direction, and PB between them in the other. (Neither is PX.) Further, as $Q \in \mathscr{C}(E)$, PQ is paired by the remarks following the definition of paired edge. The contradiction shows that $PA_i \neq PB$, so that P lies on at least two unpaired edges.

(ii) Now suppose that $v(P) \le 6$. As in (i), there are unpaired edges PA and PB closest to PQ on either side. If PR is in the arc AQB round P, then PR is paired since PA, PB are nearest. Also, each of PA, PB is in one triangle with an edge within this arc.

Suppose that $C \in \mathcal{V}(E)$ with PC not paired in E. Then PC is not adjacent to PA at P, or PAC is a triangle and PA is paired. Similarly, PC is not adjacent to PB at P. Hence there are edges PX between PA and PC, and PY between PB and PC. Then we have distinct edges PQ, PA, PX, PC, PY, PB at P. As $v(P) \leq 6$, these are all the edges at P. As PA and PB are unpaired, neither X nor Y is in $\mathcal{V}(E)$, so that PC is a loose edge. Since we have accounted for all the edges at P, we see that P is on exactly two unpaired edges. Note that we have not proved that P does lie on a loose edge, even when $v(P) \leq 6$.

This proof has a useful consequence. Let $E = \overline{F}$, for F a subdiagram of a loop-free

diagram L. When $v(P) \le 6$ for $P \in \mathcal{V}(E)$, the two unpaired edges at P divide the other edges into two sectors. In one the edges belong to E and are paired (the sector including Q in the above notation). In the other, they are loose or do not belong to E. Thus, if $R \in \mathcal{V}(\overline{E}) - \mathcal{V}(E)$ has $v(R) \le 6$, there are two possibilities.

(a) Every L-edge incident with R belongs to \overline{E} (e.g. vertex U in Figure 1). Suppose that the edges round R are, in order, RA_0, \ldots, RA_{n-1} . Then $A_{i-1}A_i$ is unpaired in E and each A_i belongs to $\mathfrak{B}(E)$. Since $A_{i-1}A_i$ is not a loop, $A_{i-1} \neq A_i$. We may have $A_{i-1} = A_{i+1}$ for some values of *i*. If we also have $A_iA_{i-1} = A_iA_{i+1}$, then this edge preceeds and follows RA_i in the ordering round A_i , so that $v(A_i) = 2$. Since A_i is joined to a vertex in $\mathscr{C}(E)$, we get a contradiction. Thus we have $A_iA_{i-1} \neq A_iA_{i+1}$. By the remarks above, the edges in the sector opposite A_iR between A_iA_{i-1} and A_iA_{i+1} all belong to E.

(b) There is an L-edge RA with $A \notin \mathcal{V}(\overline{E})$. Then $R \in \mathcal{B}(\overline{E})$. Suppose that R belongs to u triangles in \overline{E} . Each such triangle must have one edge in E. Then these triangles must involve at least u+1 edges in $\overline{E}-E$. (Otherwise they involve u edges, each in two of the triangles so that they are adjacent in pairs and clearly exhaust the edges round R in the L-diagram. Then we would have case (a), contradicting the assumption that $A \notin \mathcal{V}(\overline{E})$.)

LEMMA 1.2. In the situation of Lemma 1.1, assume that $|\mathscr{B}(E)| = q > 0$, that there are p edges joining E to L-E and that $v(P) \le 6$ for $P \in \mathscr{B}(E)$. Then $|\mathscr{B}(\overline{E})| \le p - q$.

Proof. Let J be the subdiagram with $\mathcal{V}(J) = \mathcal{C}(\bar{E})$, i.e. the vertices of J are those of E together with those of type (a). Then $\bar{J} = \bar{E}$ since the vertices of J - E are joined only to vertices of E so that the vertices added in the completion all belong to \bar{E} . We note that the completion of J involves no vertices of type (a), so that $\mathcal{V}(\bar{J}) - \mathcal{V}(J) = \mathcal{B}(\bar{J})$.

Suppose that $\overline{E} - E$ has k vertices of type (a), and that the sum of their degrees is m. Then, clearly, $m \ge k$. Also $\mathcal{B}(\overline{J}) = q - k$ and the number of edges from J to L - J is p - m. (For the latter, recall that each vertex of type (a) in \overline{E} is joined only to vertices of E and hence of J.)

By Lemma 1.1(ii), J has q - k unpaired edges. Suppose that these belong to triangles involving n vertices of $\overline{J} - J$. Since each vertex in $\mathfrak{B}(\overline{J})$ is of type (b), these triangles involve at least q - k + n of the edges leaving J (by the analysis of type (b) vertices). Each of the remaining edges leaving J may give a "new" vertex of \overline{J} . Then, since $m \ge k$, and $\overline{J} = \overline{E}$,

$$|\mathscr{B}(\bar{E})| = |\mathscr{B}(\bar{J})| = (p-m) - (q-k+n) + n$$
$$= (p-q) + (k-m)$$
$$\leq p-q.$$

2. The generalised genus formula. Formula (2) is known as the genus formula since it can be used to compute $\frac{1}{2}t_1$, the genus of the Riemann surface corresponding to the subgroup. Note that t_1 must be even for a subgroup of finite index and compare (iv) in the introductory remarks.

Now suppose that H is a subgroup of infinite index with $h_{\infty} = 0$. Since the D₁-diagram for H is connected, see [4], we can construct an increasing sequence of subpseudographs

 $(S_i: i = 1, 2, ...)$ in the following way. Let S_1 be the green polygon containing the origin and all red edges joining vertices of this polygon. Once S_i is defined, S_{i+1} is obtained by adding all green polygons linked to S_i by red edges, and all red edges joining vertices of the new structure. Observe that this is analogous to the successive completion of the L-diagram since green polygons and red edges in the D₁-diagram become vertices and the edges in the L-diagram.

Assume now that, in addition, H has $h_0 - h(6)$ finite, i.e. that there are only finitely many green polygons with more than six sides. Then there is an integer n such that S_n includes all of these. By considering S_i for $i \ge n$, we may assume that the boundary consists of "small" (i.e. at most six-sided) polygons.

Assume, further, that r, s, and h(1) are finite, and that n has been chosen so that there are no red, green or (implied) blue loops outside S_n . This simplifies the corresponding part of the L-diagram.

Let E_i be the subdiagram (of the L-diagram for H) which corresponds to S_i . Then $E_{i+1} = \overline{E}_i$.

For the moment, we also assume that, for i > n, $L - E_i$ has no loops. For i > n, we construct a finite L-diagram L_i as follows. We delete all loose edges of E_i . By Lemma 1.1, the unpaired edges of the remaining pseudograph form a subpseudograph which is regular of degree two, i.e. which consists of disjoint cycles. We suppose the entire pseudograph redrawn so that each of these cycles is a convex polygon with all other vertices and edges outside the polygons. Let k(i) be the number of cycles. We add new vertices $X_1, \ldots, X_{k(i)}$, with X_i inside the *j*th polygon, and join each vertex of the *j*th polygon to X_i . We can assume that the (redrawn) pseudograph E_i has the edges round each vertex ordered in the usual way. Now, each vertex in $\mathscr{B}(E_i)$ is outside E_n and hence of degree at most six in E_i . By the remark after Lemma 1.1, the deletion of any loose edges at such a vertex leaves (in E_i) only paired edges, with both natural neighbours in E_i , and the two unpaired, each with one of its neighbours (one of the paired edges). We order the edges in L_i to maintain the anticlockwise convention at each vertex, including the X_i . Then it is clear that L_i is a K-diagram with the triangle property. To see that it is an L- diagram, we must check that it is connected. Since E_1 has one vertex, it is connected. Now, $E_i \subseteq \overline{E}_{i-1}$, so that each E_i is connected by induction on i. For $i \ge n+1$, the region beyond E_i is free of loops. Then Lemma 1.1 applies, so that each vertex of $E_i - E_{i-1}$ is joined to a vertex of E_{i-1} by an unpaired edge. Hence, in this region, the removal of loose edges does not disconnect the diagram. The new vertices, the X_i , are each joined to a vertex of E_i . Thus L_i is connected.

Following [2], we define a T_2 for the D_2 -diagram for H. We use the S_i . From S_1 we take all but one of the green edges and no red edges. Suppose that we have dealt with S_i . We expand the spanning tree to S_{i+1} by taking one red edge to each new polygon and all but one of the green edges from each. This subpseudograph spans S_{i+1} . In the limit, we get a T_2 for the D_2 -diagram. We fix this and choose a graph G of the type considered in [4].

Consider now the corresponding subdiagrams E_i . A red edge from a polygon of S_i to one in $S_{i+1} - S_i$ corresponds to an edge from $\mathcal{V}(E_i)$ to $\mathcal{V}(E_{i+1}) - \mathcal{V}(E_i)$. Since $E_{i+1} = \overline{E}_i$, we have $\mathcal{V}(E_i) = \mathscr{C}(E_{i+1})$, so that such an edge is paired in E_{i+1} .

Let H_i be the subgroup of Γ defined by the L-diagram L_i . A T_2 for the D₂-diagram for H_i is obtained from the spanning tree for S_i described above. We add one red edge to each of the polygons corresponding to the X_i , and all but one of the green edges of these polygons. This gives a spanning tree and thus a T_2 .

For the next lemma, we require some notation: for each i > n,

 $M(i) = |\mathscr{C}(E_i)|,$

 $N(i) = \sum v(P)$, where the sum is over $P \in \mathscr{C}(E_i)$,

$$q(i) = |\mathscr{B}(E_i)|,$$

p(i) = number of edges joining E_i to $L - E_i$,

- n(i) = number of triangles in E_i ,
- n'(i) = number of triangles in L_i ,
- t(i) = number of red edges in $S_i T_2$ which do not correspond to directed edges of G.

Since G is chosen once for all, and since $S_i \ge S_j$ for $i \ge j$, t(i) is non-decreasing. Recall from [2] that the edges enumerated by the t(i) correspond to hyperbolic generators in the presentation of H derived from the chosen T_2 and G. If t_1 is finite, then $t(i) = t_1$ when i is large enough. If t_1 is infinite, then t(i) tends to infinity with i.

Observe that, in the construction of L_i from E_i , only loose edges are deleted, so no triangles are lost. Also, E_i has q(i) unpaired edges (Lemma 1.1), so that q(i) new triangles are created. Hence,

$$n'(i) = n(i) + q(i).$$

LEMMA 2.1. For i > n + 1,

$$N(i) + 2q(i) - p(i) \ge 3r + 4s + 6(M(i) + t(i) - 1).$$

Proof. At present we are assuming that r+s+h(1) is finite, that $h_0-h(6)$ is finite, and that the L-diagram for H has finitely many loops. We have chosen n so that the D-diagram (L-diagram) for H has no loops outside S_n (E_n). Further, in the L-diagram, $v(P) \le 6$ whenever P is outside E_n . From these, it follows that the D-diagram for H_i has r red loops and s blue loops for i > n.

From [6], the (finite) index of H_i is equal to the sum of the degrees of the vertices of L_i . The vertices in $\mathscr{C}(E_i)$ have total degree N(i). There are q(i) vertices in $\mathscr{B}(E_i)$, and each has degree at most six in the L-diagram for H and hence in each E_i . The L-edges in L_i contribute at most 6q(i) - p(i) to the total degree of these vertices in L_i . The new edges (to the X_i) contribute q(i) to their degree. The X_i have total degree q(i) by the construction. Hence H_i has index less than or equal to N(i) + q(i) - p(i).

The number of vertices in L_i is M(i)+q(i)+k(i) (from $\mathscr{C}(E_i)$, $\mathscr{B}(E_i)$ and from the X_i). This is the parabolic class number, the h_0 , of H_i .

Using the T_2 for H_i constructed above and the relationship between D- and L-diagrams, there are at least n(i)+t(i)+q(i)-k(i) red edges not in the T_2 . Now, L_i has n'(i) triangles and is finite, so the corresponding G has n'(i)-1 directed edges. Thus, H_i has at least

$$(n(i) + t(i) + q(i) - k(i)) - (n'(i) - 1) = t(i) - k(i) + 1$$

hyperbolic generators. Using the genus formula (2) for the finite index subgroup H_i , we have

$$(N(i) + 8q(i) - p(i)) \ge 3r + 4s + 6(M(i) + q(i) + k(i) + t(i) - k(i) + 1 - 2)$$

= 3r + 4s + 6(M(i) + q(i) + t(i) - 1).

The result follows at once.

Define G(i) (for i > n) by

$$G(i) = N(i) - 6M(i) - 3r - 4s - 6t(i) + 6.$$
(3)

COROLLARY 2.2. For i > n+1, $G(i) \ge 0$.

Proof. We note that, in our new notation, Lemma 1.2 can be rewritten

$$q(i+1) \leq q(i) - p(i).$$

Suppose that, for some i > n+1, G(i) < 0. Then, from the lemma,

$$G(i) = p(i) - 2q(i)$$

so that

$$q(i) > p(i) - q(i) \ge q(i+1).$$

Now, t(i) is non-decreasing. As all polygons outside S_n have at most six sides, N(i)-6M(i) is non-increasing. Also, when i > n, r and s are fixed. Hence G(i) is non-increasing for i > n. Thus, on our hypothesis that G(i) is once negative, it is negative thereafter. Then q(i) is strictly decreasing. As q(i) is, by definition, a non-negative integer, we have a contradiction. Hence $G(i) \ge 0$ for i > n + 1.

We now describe how a diagram may be altered to allow us to apply the corollary in more general circumstances. We must be able to remove all but finitely many red, blue and green loops from a D-diagram, and to remove loops outside a finite region of an L-diagram. We retain the assumption that the D-diagram under discussion has finitely many green polygons with more than six sides. We may also assume that we operate outside S_n or E_n for a suitably chosen n.

For the removal of blue and red loops, see the sketches in [5]. These do not have the green edges, but, since green edges are determined by red and blue edges, they can be added to justify the assertions which follow.

In a diagram with more than one vertex, a blue loop occurs at a vertex of a green polygon with at least two sides. The removal reduces the size of this polygon by one, but does not affect any other green polygon. It produces a new red loop.

In an infinite diagram, indeed in a diagram with more than six vertices, a blue loop is at a vertex joined to a blue triangle and a triangle has at most two such vertices attached. This, like some assertions below concerning the context of coloured loops, follows from the fact that the red and blue edges connect a diagram. With an infinite diagram, the removal of all of the blue loops, even if there are infinitely many, leaves an infinite diagram. Further, the removal can only reduce the sizes of green polygons, so the new diagram has $h_{\infty} = 0$, and $h_0 = \infty$. The value of $h_0 - h(6)$ cannot increase. Now assume that this has been done.

In an infinite diagram, indeed in a diagram with more than three vertices, a red loop occurs at a vertex of a blue triangle, and each triangle has at most two red loops attached. The removal of red loops is achieved in one of two ways, depending on whether the associated triangle has one or two red loops. In the present context, we can assume that we operate outside S_n so that each red loop is at a vertex of a green polygon with at most six sides. A sketch shows that a triangle with two red loops is joined to another triangle with no red loops (or we should have a heptagon at least) and without a red edge joining the remaining vertices (or we have a six-vertex diagram). Under these circumstances, the procedure in [5] cannot lead to a new red loop, nor can the presence of a triangle with two loops affect the removal of a triangle with a single loop.

A triangle with a single red loop does not have a red edge joining the other vertices, except in a three-vertex diagram. It follows that the removal of such a loop does not produce a further loop. A sketch shows that a chain of six or more triangles each with a red loop must involve a green polygon with more than six sides. Such a feature cannot occur outside S_n . A finite chain of triangles can be removed in one step in the same way as a single triangle.

Either of the procedures in [5] affects at most two green polygons. Each affected polygon is reduced in size, but cannot disappear. We can remove all red loops in a single operation. As no green polygons are lost, we are left with an infinite diagram. The new diagram has $h_{\infty} = 0$ and $h_0 = \infty$. Since we operated outside S_n , $h_0 - h(6)$ is unchanged.

Suppose that the D-diagram has no red or blue loops outside S_n . Apart from a one-vertex diagram, a green loop must occur at a vertex V of a blue triangle with a red edge joining V to another vertex of the triangle. For connectedness, the third vertex of this triangle is joined to a further blue triangle. The other vertices of the triangle including V belong to the same green polygon, see Figure 2. We refer to this as the polygon associated with the green loop. Provided there are no red loops involved, the second blue triangle has distinct red edges at each vertex (or we would have a six-vertex diagram).



Figure 2.

We can remove both triangles and repair the broken red edges much as for a triangle with one red loop. This removes the green loop, reduces the size of the associated green polygon by four and that of a further green polygon by one. The last polygon is that attached to the associated polygon. We see that the associated polygon has at least five sides. Outside S_n , such a polygon is a pentagon or a hexagon. In the latter case, the removal leaves a digon. In the former, we get a new green loop. Figure 2 shows that this occurs when the original diagram D was the result of composition (see [4]) of an infinite diagram D_0 with a six-vertex diagram C. This C has two blue triangles, two red edges and two red loops and one green loop; there is only one such diagram. The composition must occur at a triangle of D_0 with two red loops. Figure 3 confirms that such a triangle has its vertices in a green polygon with at least six sides (as D_0 is infinite). Since composition does not alter the sizes of green polygons, D_0 has no large polygons outside a finite region S. Thus the two loop triangle is in a hexagon or lies within S. The case of a hexagon arises from the composition of an infinite diagram D_1 with the six-vertex diagram C_1 , where C_1 has two blue triangles and four red loops. In such a case, we can undo this composition to get a triangle with two red loops. Since the original D was connected, a finite sequence of decompositions leads into S_n . The entire finite chain of hexagons begins at a triangle in S_n and ends in a green pentagon associated with a green loop. Since each triangle in S_n is the start of at most one such chain, and there are finitely many triangles in S_n , there are finitely many chains. Now, since each chain has finite length, we may as well increase n to ensure that all chains lie in S_n .



Figure 3.

We may now assume that each green loop outside S_n is associated with a hexagon. There may be infinitely many such loops, each associated with a different hexagon. The removal of all of them leaves an infinite diagram with $h_{\infty} = 0$, $h_0 = \infty$ and $h_0 - h(6)$ unchanged.

Now suppose that we have a D-diagram with no loops outside S_n . The L-diagram may have loops outside E_n . Recall (from [4]) that a loop in an L-diagram corresponds to a red loop or a red edge which joins two vertices in a single green polygon. Outside E_n , all are of the latter type.

In [4], we described a procedure which removes a red edge (a cut). In the case of a red edge corresponding to a loop in the L-diagram, Lemma 3.2 or Lemma 3.7 of [4] applies. The result of a cut is a diagram or a pair of diagrams with two green polygons in place of the original one involving the red edge. As the original had at most six vertices, one of the new ones has at most three. A vertex with a red loop which has a green loop is a one-vertex diagram (the red and green edges connect a diagram) and so has a blue loop also. A cut does not affect blue edges and so cannot produce a blue loop. Since we are

outside S_n , the split cannot give a green loop, i.e. both new polygons have at least two vertices. A green digon with a red loop at one vertex occurs only in a three-vertex diagram since the green digon of this type forces a red edge between two vertices of a triangle with a red loop. This diagram has a green loop. This was not affected by the cut since affected polygons have a red loop. Again since we are outside S_n , the split cannot give a digon. Thus the smallest green polygon which can arise has three vertices. Since we began with at most six vertices, we must have had a green hexagon and this must yield two green triangle must be the result of the composition of a diagram (perhaps null) with the diagram consisting of a blue triangle with three red loops. Now consider the reversal of the cut, i.e. the join (see [4]) of the red loops of the triangles. This gives a green hexagon of the type we met when dealing with green loops.

Much as in the case of green loops, we can have chains of such hexagons, but now these can be of two types.

(c) A chain is finite. Then both ends involve polygons with at least seven vertices, i.e. within S_n . Since S_n is finite, there can be only finitely many such chains. As each is of finite length, we can increase *n* to ensure that the entire chains all lie in S_n .

(d) A chain extends infinitely in one direction. We refer to such a chain as a 6-chain. Each 6-chain terminates in S_n , so there are finitely many. Undoing the composition where a 6-chain enters S_n , we are left with a 6-chain ending in two red loops, C(6), and another diagram D. The latter may be infinite or finite. We treat these separately, but first we observe that, in C(6), each red edge belongs either to the T_2 or to the G (once T_2 has been chosen, there is no choice for the corresponding G).

(i) D is infinite. Then Corollary 3.10(ii) of [4] applies, showing that the removal of each copy of C(6) reduces t_1 by one, the relevant edge being inside S_n . Thus t(i) is reduced by one for i > n. Since only hexagons are removed, N(i) - 6M(i) is not changed. The process does not alter s but increases r by two. Thus the G(i) for i > n are unchanged. The removal of the last of the finitely many 6-chains may lead to the following case.

(ii) D is finite. Now Corollary 3.10(i) of [4] applies. By the remarks in (i) regarding red edges of a 6-chain, t(i) = t(n) for i > n, and t(n) is equal to the t_1 of the original diagram and of D. Applying (2) to the subgroup corresponding to D, we obtain

$$N(n) = 3(r+2) + 4s + 6(M(n) + t(n) - 2)$$

= 3r + 4s + 6(M(n) + t(n) - 1).

Hence G(n) = 0. Since t(i) = t(n) (i > n) and all green polygons outside S_n have six sides, G(i) = 0 for $i \ge n$.

LEMMA 2.3. Let H be a subgroup of infinite index in Γ with $h_{\infty} = 0$ and $r + s + h(1) + (h_0 - h(6))$ finite. Then $t_1 + h(5)$ is finite.

Proof. Let D be the D-diagram for H. By the hypotheses, D has finitely many loops and finitely many green polygons with more than six vertices.

111

Suppose first that D consists of the composition of a finite diagram D_0 with a (necessarily finite) number of 6-chains. The remarks under (d) above show that the red edges enumerated by t_1 and all large green polygons lie in D_0 . Since D_0 is finite, t_1 and h(5) are finite, as required.

Now suppose that the removal of all 6-chains leaves an infinite diagram D^* . As in (d)(i), the G(i) for D^* are equal to those for D when i > n. Also, t_1 has been decreased by a finite quantity (viz. the number of chains removed) and r increased by twice this amount. Finally, s and h(5) are unchanged. Let L be the L-diagram corresponding to D^* . Since we have removed all composition chains, or taken them inside S_n (see (a)-(d) above), L has no loops outside E_n . All vertices outside E_n have degree at most six, so that

$$e(i) = N(i) - 6M(i) - 6t(i)$$

is non-increasing (see the proof of Corollary 2.2). Suppose that e(i) decreases infinitely often. Then, for *i* sufficiently large, D^* has G(i) negative, contradicting Corollary 2.2. Hence e(i) is constant for $i \ge m$. It follows that E_m includes all vertices of degree less than six and all edges enumerated by t_1 . Thus D^* , and hence D, has $t_1 + h(5)$ finite.

THEOREM 2.4. Suppose that H is a subgroup of infinite index in Γ with $h_{\infty} = 0$ and $h_0 - h(6)$ finite. Then $r + s + t_1 + h(5)$ is finite and

$$\sum_{j=1}^{\infty} (c(j)-6) \ge 3r+4s+6t_1-6.$$
(4)

(Note that, since h(5) is finite, there is an integer m with c(j) = 6 for $j \ge m$. Thus the sum on the left of (4) is over a finite range.)

Proof. We begin by noting that the procedures for removing loops from a D-diagram cannot increase $h_0 - h(6)$ and do not affect t_1 .

Suppose that $r+s+h_0-h(6)$ is finite but that h(1) is infinite. Since $h_0-h(6)$ is finite, all but finitely many of the green loops are associated with green pentagons or hexagons. By the earlier analysis of green loops, only finitely many of them are associated with pentagons. Of the rest, only a finite number lie in S_n the region containing all red and blue loops and all the large green polygons. If we remove the infinite collection associated with hexagons and outside S_n , then we have a diagram with r+s+h(1) finite but h(2) and hence h(5) infinite. This contradicts Lemma 2.3.

Now suppose that $s + h_0 - h(6)$ is finite but that r + h(1) is infinite. By removing all red loops outside a region S containing all blue loops and all large green polygons, we obtain a diagram with $r + s + h_0 - h(6)$ finite but with h(5) infinite (each red loop outside S is in a green polygon with at most six sides, its removal leaves a polygon with at most five sides). For the new diagram, either h(1) is finite, contradicting Lemma 2.3, or h(1) is infinite, when we get a contradiction as in the previous paragraph.

Finally, suppose that s is infinite. The removal of all blue loops gives a diagram with $h_0 - h(6)$ still finite, but with r infinite (the removal of a blue loop leaves a red loop). Then we get a contradiction as in the previous paragraph.

It follows that we must have $r+s+h(1)+h_0-h(6)$ finite. Then Lemma 2.3 shows that $t_1+h(5)$ is finite.

To show that (4) holds, we consider first the case where the removal of all 6-chains leaves a finite diagram. By the analysis under (d)(i), (ii) above, there is an integer n such that the diagram has G(i) = 0 and $t(i) = t_1$ for i > n. As h(5) is finite, there is an integer m such that S_m contains all green polygons with fewer than five sides. Let $k = \max\{m, n+1\}$. Then all green polygons outside S_k have six sides. Thus

$$\sum_{j=1}^{\infty} (c(j) - 6) = N(k) - 6M(k).$$

Also, G(k) = 0. The result (4) follows readily.

112

Now suppose that the removal of 6-chains leaves an infinite diagram. As in (d)(i), the removal does not affect G(i) for large *i*. After the removal of these chains, we can choose *n* so that S_n (for the new diagram) contains all loops and all green polygons with other than six sides, and all finite composition chains. The corresponding L-diagram has no loops outside E_n . For i > n,

$$G(i) - 6M(i) = N(n) - 6M(n) = \sum_{j=1}^{\infty} (c(j) - 6).$$

Then (4) follows from Corollary 2.2.

We observe that formulae (1) and (2) can be rewritten as

$$\sum_{j=1}^{h_0} (c(j)-6) = 3r + 4s + 6t_1 - 12.$$
(5)

These formulae are frequently used to calculate the genus $(g = \frac{1}{2}t_1)$ of the Riemann surface corresponding to the subgroup. We refer to (2) or (5), and their infinite analogue (4) as the genus formulae. The discrepancy of six on the right hand sides of (4) and (5) may be interpreted as a consequence of the fact that a finite index subgroup has $h_0 - 1$ parabolic generators and one of infinite index has h_0 (in a "standard" presentation, see [4]).

In Theorem 2.4, our hypotheses were sufficient to ensure that the sum in (4) was finite. We say that (4) is satisfied whenever the left hand side is infinite. Note that, when this happens, there are infinitely many of the c(j) greater than six.

3. Existence results. To prove a converse of Theorem 2.4, we describe how suitable L-diagrams can be constructed. We are now free to avoid unpleasant features, e.g. loops and multiple edges, whenever possible. In fact we will construct basic L-diagrams and explain how the corresponding D-diagrams may be altered to introduce red, blue and green loops, etc.

Recall that, if we have an L-diagram with no loops, then the green polygon corresponding to each vertex involves just one vertex of each blue triangle that it meets. Let L be such an L-diagram, and let D be the corresponding D-diagram.

If L has adjacent vertices of degrees 4 and 5, then D has a pair of triangles joined by a red edge with one end in a 4-gon and the other in a 5-gon. We can add a new (blue) triangle in the middle of this red edge. The remaining vertex of the new triangle has a red loop so that we have a new D-diagram. The orientation of the triangle can be chosen to give two green 6-gons; one will have a red loop at one vertex.

If, instead, L has adjacent vertices of degrees 3 and 5, then we can add a triangle to get a D-diagram with a 6-gon and a 5-gon with a red loop at one vertex. We can then add a further vertex with a blue loop in place of the red loop. We obtain two 6-gons, one with a blue loop at one vertex.

Finally suppose that L has adjacent vertices of degree 2 and 5. Adding a triangle as before, we get a 6-gon and a 4-gon with a red loop at one vertex. We replace the red loop with a red edge leading to a further blue triangle, this one having a red edge joining the other vertices. We then have two 6-gons and a green loop (see Figure 2).

These three constructions are simply the reversals of the procedures for removing loops discussed earlier.

Our existence theorem is a generalisation of Theorem 2.2 of [6]. In the proof of that result, we constructed an L-diagram with its vertices lying at a point P and on circles C(k), k = 1, 2, ..., of radius k round P. Here we sometimes need to add vertices between C(k) and C(k+1) to form the boundary of E_k , the kth stage of the construction. See Figures 4 and 5. For the moment, we restrict ourselves to the basic construction of [6].

Let (c(j): j = 1, 2, ...) be a sequence of integers with $c(j) \ge 6$. The vertex set of E_1 consists of P and of c(1) points on C(1). The edge set consists of the arcs of C(1) and edges joining P to each of the points on C(1).

Suppose that we have constructed E_i such that

(i) vertices inside C(i) have degrees $c(1), \ldots, c(n_i)$,

(ii) vertices on C(i) are numbered $n_i + 1, ..., n_i + m_i$, vertex k having degree a(k) = 3 or 4.

If X and Y are adjacent vertices on C(i), we add vertex v(X, Y) on C(i+1) between PX and PY. The new vertex is joined to X and to Y. Let Z be the other vertex on C(i)adjacent to Y. We join Y to v(X, Y), to v(Y, Z) and to c(k) - a(k) - 2 vertices on C(i+1)between these. This gives E_{i+1} with the properties (i) and (ii).

LEMMA 3.1. In the above notation, the boundary of E_i has m_i vertices of degree 4 and $\sum_{j=1}^{n_i} (c(j)-6)+6$ of degree 3.

A proof by induction is straight-forward. We observe that if c(j) = 6 for large j then the number of boundary vertices of degree 3 will remain constant after some point.

We now explain how we can relax the condition that each c(j) be greater than or equal to six.

We can allow some c(j) equal to five. In the above construction, we can take a boundary vertex Y of degree 3 and treat it like one of degree 4, i.e. add no vertices between v(X, Y) and v(Y, Z).

To introduce a vertex of degree 4 inside E_{i+1} (so it will be a vertex of this degree in E_j for j > i), we begin with a vertex B of degree 3 on the boundary of E_i . Let A and C be the vertices adjacent to B on the boundary. Instead of v(A, B) and v(B, C), we add a single vertex X on C(i+1). X is joined to A, B and C and to its neighbours on C(i+1). To the left and right of B we proceed as usual. On the boundary of E_{i+1} , the vertex X has degree 5. In the continuation, we must treat X like B, i.e. we join it to just one new vertex. Thus each subsequent E_j will have a vertex of degree 5. Adding more than one vertex of degree 4 requires care as we see in the next paragraph.

Suppose that E_i has two boundary vertices of degree 5 separated by l of degree 4. Figure 4 shows the situation with U and V of degree 5 and l = 1. It is quite easy to see that, in general, E_{i+1} will have two vertices of degree 5 just l-1 apart. Subsequent continuation will lead to adjacent vertices of degree 5, e.g. B and C in Figure 4. This cannot be continued in the usual way without introducing a vertex of degree 6. The figure shows how we continue to E_{i+2} . The new vertex Y is in E_{i+2} but not on its boundary. The boundary vertices X and Z have degree 5, so that this technique must be employed at each stage.





Rather than have a large number of figures, we now give a verbal description of the construction; we use this type of description later. Given successive vertices A, B, C, F with v(A) = v(F) = 4, v(B) = v(C) = 5, add vertices X, Y, Z with edges AY, BY, CY, FY, AX, FZ, XY, YZ. Note that the new vertices are listed in the same order as they are to appear, and the new edges are to be added outwards from E_i . This gives the graph underlying E_{i+1} . We get the subdiagram (of an L-diagram) by numbering the edges anti-clockwise at each vertex.

115

It is an easy exercise in the continuation techniques to show that a diagram with three vertices of degree 5 separated only by vertices of degree 4 yields a finite diagram on continuation. We therefore avoid creating such a situation.

It is also easy to see that, if the boundary of E_i has an adjacent pair of vertices of degree 5 separated by vertices of degree 4 from a vertex of degree 3, then, for some k, E_{i+k} has successive vertices A, B, C with v(A) = v(B) = 5, v(C) = 3. This stage is shown in Figure 5, together with the continuation. Once again, the vertex Y does not lie on the new boundary. The new boundary has just one vertex of degree 5, viz. X.



Finally, unless a further vertex of degree 5 intervenes, a pair of boundary vertices of degrees 3 and 5 remain a constant distance apart in each further stage. Thus a diagram with alternating vertices of these degrees separated by vertices of degree 4 leads to an infinite diagram. In fact, a calculation like that for Lemma 3.1 shows that the number of new vertices at each stage is fixed.

For later purposes, we may wish to amend E_i to obtain a more convenient disposition of the vertices of degree 3. Suppose that E_i has vertices B and C of degree 3 separated by k of degree 4. Between C(i) and C(i+1), we add vertices of the type required for the usual continuation " E_{i+1} " in the arc BC above the k vertices. A sketch shows that this new boundary has B and C of degree 4 and two vertices of degree 3 separated by k-1 of degree 4. See the portion of Figure 6 up to the line WY. Still remaining within C(i+1), we can repeat this process until we have an adjacent pair of vertices of degree 3. We refer to such a pair as a special pair. In Figure 6, two stages are needed, since the original E_i had k = 2. The new boundary is ... AWXYZF..., with X, Y as the special pair.

Suppose now that we have a boundary with four vertices of degree 3 separated by k_1 , k_2 and k_3 of degree 4. If $k_1 > k_3$, then we can proceed as in the previous paragraph to decrease the distance between the last pair to $k_3 - 1$. Note that k_1 is not affected. By repeating the process a suitable number of times, we get a boundary with $k_1 = k_3$. We can now apply the process to the central pair, increasing both k_1 and k_3 by one each time to get a special pair with further vertices of degree 3 an equal distance on each side. Of course, an analogous process can be applied if $k_1 < k_3$.





Suppose that we have a specification with $h_{\infty} = 0$ and $h_0 - h(6)$ and $r + s + t_1 + h(5)$ finite, and such that (4) of Theorem 2.4 is satisfied. We re-order the sequence (c(i)) so that those greater than six occupy the initial places, and there follow K sixes (a specific K could be determined from the constructions which follow).

We construct the first few E_i as usual, proceeding until all vertices of degree greater than six have been included, and the boundary consists of vertices which will have degree six in the full diagram. By Lemma 3.1, the boundary of E_i will have $\sum_{c(i)>6} (c(i)-6)+6$ vertices of degree 3, the rest being of degree 4. By (4), we will have a vertex of degree 3 for each *i* with c(i) = 5, two for each with c(i) = 4, etc., and at least $3r + 4s + 6t_1$ others.

We orientate the boundary of E_i for later convenience, and assume that the vertices of degree 3 have been suitably arranged.

(a) To obtain a vertex of degree 5, we take one of degree 3 on the boundary and treat it as described above.

(b) To obtain a vertex of degree 4, we begin with a special pair. Suppose that the pair is B, C, and that they are between A and F, the latter being of degree 4. We add a new vertex X and edges BX, CX and FX. The new boundary includes A, B, X and F, with degrees 4, 4, 3 and 5 respectively.

(c) To obtain a vertex of degree 3, we take a vertex of degree 3 followed by a special pair. With the special pair in A, B, C, F as in (b), we add the edge BF. In the new graph, the degrees of B and F are 4 and 5 respectively, C is no longer on the boundary. We now have vertices of degree 3 and 5 in this order, not necessarily adjacent.

(d) To obtain a vertex of degree 2, we take a vertex of degree 3, a special pair and another of order 3. As above, we may assume that the vertices of degree 3 are equidistant from the special pair. If the pair are B and C, we add a new vertex X and edges BX, XC and BC. The boundary vertices are unchanged, but B and C now have degree 5. If we

continue the diagram between the remaining vertices of degree 3, then we obtain a boundary with vertices of degree 4 in the new part.

(e) To obtain a green loop, we can use adjacent vertices of degrees 2 and 5. We take a vertex of degree 3, followed by four others arranged as for case (d). Firstly, we assume that the four occur in the sequence A_1, \ldots, A_6 , with A_1, A_3, A_4 and A_6 of degree 3 and A_2 , A_5 of degree 4. We add a new vertex X and edges $A_3X, A_4X, A_3A_4, A_3A_5, A_2A_5, A_2A_6, A_1A_6$. The new boundary has adjacent vertices A_1, A_6 of degrees 4, 5 respectively. With the first vertex of degree 3, we have a boundary pair of degree 3 and 5 in order, as in (c). This can be generalised to any of the situations in (d).

(f) To obtain a red loop, we use adjacent vertices of degrees 4 and 5. We take a vertex of degree 3 followed by a special pair B, C. We treat B as if it were of degree 4, and C as if it were of degree 5 in later construction. Thus, in effect, we have a pair of degrees 3, 5.

(g) To obtain a blue loop, we require adjacent vertices of degree 3 and 5. As in (e), we consider first adjacent vertices A_1, \ldots, A_6 . We add new vertices X and Y, and edges $A_3A_5, A_3X, A_5X, A_2Y, A_1Y, XY, XA_6$. The new boundary has vertices A_1, Y, X and A_6 , of degrees 4, 3, 5 and 4. Once again, this is readily generalised to all arrangements in (d).

(h) To obtain the desired value of t_1 requires a more elaborate construction. Above, we assumed that an L-diagram was being built outwards from the boundary at each stage. By inverting with respect to the boundary at the *i*th stage, we can easily see that we could equally well arrange that the subsequent stages all lie within a polygon.

Suppose that we have six vertices of degree 3 on the boundary. We may assume that these are A_1, \ldots, A_6 , and that A_1, A_2 and A_5, A_6 are special pairs, and that all intermediate vertices have degree 4. We add an edge A_2A_5 , and treat A_2 and A_5 as of degree 5 in a new boundary obtained by ignoring those between A_2 and A_5 . We add a new vertex X and edges A_iX , i = 1, 2, 5, 6. The relevant part of the boundary is now A_1, X, A_6 , and all have degree 4. This leaves another boundary component consisting of A_2 and A_5 (of degree 5) and those between, including A_3 and A_4 . As explained above, we can continue indefinitely from such a boundary using vertices of degree 6, though now we must work inwards.

As we assume that (4) is satisfied, we have $6t_1$ vertices of degree 3 available for such constructions, so we can introduce t_1-1 additional boundary components and continue each indefinitely. An application of Lemma 3.1, a count of the number of vertices of degree 3 and 5 added and removed in the constructions and an application of (2) show that the final diagram will have the required t_1 .

As $r + s + t_1 + h(5)$ is finite, all these can be carried out in a finite number of stages.

We conclude this section by considering specifications with $h_{\infty} = 0$ and $h_0 - h(6)$ infinite. We note that we then have infinitely many c(i) greater than six.

The construction of a suitable D-diagram can be undertaken as above, with obvious changes to use vertices of degree at least six in place of those of degree six. This simplifies most of the construction since a boundary vertex of degree 5 can be assumed to become a vertex of degree at least 7 in later stages. We may have to introduce an infinite number of

special features, but this is possible since a few stages consisting of vertices of degree 7 or more will give enough boundary vertices of degree 3 for further constructions of type (a) to (h). We have proved our final result.

THEOREM 3.2. If a specification has $h_{\infty} = 0$ and satisfies (4), then it is the specification of a subgroup of Γ .

REFERENCES

1. G. A. Jones and D. Singerman, Theory of maps on orientable surfaces, Proc. London Math. Soc. (3) 37 (1978), 273-307.

2. G. A. Jones, Triangular maps and non-congruence subgroups of the modular group, Bull. London Math. Soc. 11 (1979), 117-123.

3. M. H. Millington, Subgroups of the classical modular group, J. London Math. Soc. (2) 1 (1969), 351-357.

4. W. W. Stothers, Subgroups of infinite index in the modular group, Glasgow Math. J. 19 (1978), 33-43.

5. W. W. Stothers, The number of subgroups of given index in the modular group, *Proc. Roy.* Soc. Edinburgh Sect. A 78 (1977), 105–112.

6. W. W. Stothers, Diagrams associated with subgroups of Fuchsian groups, Glasgow Math. J. 20 (1979), 103-114.

UNIVERSITY OF GLASGOW GLASGOW G12 8QW