J. Austral. Math. Soc. (*Series A*) **47** (1989), 350–367

GEOMETRY OF NEUMANN SUBGROUPS

RAVI S. KULKARNI

(Received 25 February 1986)

Communicated by H. Lausch

Abstract

A Neumann subgroup of the classical modular group is by definition a complement of a maximal parabolic subgroup. Recently Neumann subgroups have been studied in a series of papers by Brenner and Lyndon. There is a natural extension of the notion of a Neumann subgroup in the context of any finitely generated Fuchsian group Γ acting on the hyperbolic plane $\mathbf{H}$ such that $\Gamma\backslash\mathbf{H}$ is homeomorphic to an open disk. Using a new geometric method we extend the work of Brenner and Lyndon in this more general context.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*) (1985 *Revision*): 20 E 28.

1. Introduction

(1.1) This note essentially consists of some remarks on a series of recent papers by Brenner and Lyndon concerned with the Neumann subgroups of the classical modular group; cf. [1]–[3]. We first recall their definition. Let $\Gamma = \langle x, y | x^2 = y^3 = e\rangle$ act as the modular group on the upper half plane $\mathbf{H}$ in the standard way. Then the subgroup $P = \langle z\rangle$, where $z = xy$, is a maximal parabolic subgroup of Γ and all such subgroups are conjugate. A subgroup Φ of Γ is said to be *non-parabolic* if it contains no parabolic element. If Φ is a complement of P in Γ, that is (1) $P \cap \Phi = \{e\}$, and (2) $P \cdot \Phi = \Gamma$, then Φ is called a *Neumann subgroup*; cf. [1]. A Neumann subgroup is maximal among non-parabolic subgroups; cf. [1, (2.8)].

This work was partially supported by the Max-Planck-Institut für Mathematik, Bonn, Germany and an NSF grant.

(1.2) In connection with these subgroups, Brenner and Lyndon were led to study *transitive triples* (Ω, A, B) (cf. [2]) where Ω is a countable set, A and B are permutations of Ω of orders 2 and 3 respectively such that the group $\langle C \rangle$, where $C = AB$, is transitive on Ω. If (Ω, A, B) is a transitive triple then Γ, as in (1.1), acts on Ω in the obvious way so that the subgroup P is transitive on Ω. In particular $\Omega \approx \Gamma/\Phi$ for a suitable subgroup Φ whose conjugacy class is well-defined. Since $P \approx \mathbf{Z}$ it is clear that either P acts simply transitively on Ω in which case Ω is an infinite set, or else P acts ineffectively on Ω in which case Ω is a finite set. In the first case Φ is a Neumann subgroup. In the second case $(\Gamma : \Phi) < \infty$ and $\Phi \backslash \mathbf{H}$ has only one cusp. Such a subgroup was called *cycloidal* by Petersson [9]. Thus the study of transitive triples amounts to a simultaneous study of Neumann and cycloidal subgroups of Γ. For the earlier work on Neumann subgroups see [8], [6], [13], and also [7, pages 119–122].

(1.3) A principal result in [1] which extends Theorem 2 of [13] is a structure-and-realization theorem for Neumann subgroups. Similar and more general results were proved by Stothers [10]–[12]. The proof in [1] is based on a correspondence between transitive triples and Eulerian paths in cuboid graphs, that is, the graphs with vertex-valences at most 3. For the triples associated with torsion-free Neumann-or-cycloidal subgroups the correspondence is one-to-one, but in general to make the correspondence one-to-one one would need to put an extra structure on the cuboid graphs. The same method is used in [3] to produce maximal-among-non-parabolic subgroups which are not Neumann.

(1.4) In this note we extend this work to

$$\Gamma = \prod_{i=1}^{*n} \Gamma_i, \quad \Gamma_i = \langle x_i \rangle \approx \mathbf{Z}_{m_i}, \qquad 2 \leq m_i, \ n < \infty. \tag{1.4.1}$$

Except for $\Gamma \approx \mathbf{Z}_2 * \mathbf{Z}_2$, these groups can be realized as discrete subgroups of the orientation-preserving isometries of the hyperbolic plane $\mathbf{H}$ such that $\Gamma \backslash \mathbf{H}$ has finite area, and x_i acts as a rotation through angle $2\pi/m_i$ around its fixed point. Then the element $u = x_1 x_2 \cdots x_n$ is parabolic.

(1.5) The above remarks are meant only for motivation. In the following, hyperbolic geometry will not be used explicitly. We start with Γ as in (1.4.1). Let $u = x_1 \cdots x_n$. The conjugates of u^k, $k \neq 0$ are called the *parabolic elements* of Γ. Let $P = \langle u \rangle$. A subgroup of Γ is called *parabolic* if all of its non-identity elements are parabolic. Clearly the maximal parabolic subgroups are precisely the conjugates of P. A subgroup of Γ is called *non-parabolic* if it contains no parabolic element. A complement Φ of P in Γ is called a *Neumann subgroup*. Thus, for a Neumann subgroup Φ one has (i) $P \cap \Phi = \{e\}$, and (ii) $P \cdot \Phi = \Gamma$. The latter implies (ii)$'$ $|P \backslash \Gamma / \Phi| = 1$.

Conversely, if (ii)′ holds and $(\Gamma: \Phi) = \infty$ then Φ is a Neumann subgroup. If (ii) holds and $(\Gamma: \Phi) < \infty$ then as in [5], Φ is called a *1-cycloidal subgroup.* (In the correspondence between subgroups of the Fuchsian groups and holomorphic maps among Riemann surfaces, the 1-cycloidal subgroups precisely correspond to meromorphic functions on closed Riemann surfaces with a single pole; these functions may be considered as generalizations of polynomial maps; cf. [5].) One sees (cf. (2.1)) that a Neumann subgroup is maximal among non-parabolic subgroups.

(1.6) Let Γ be as in (1.4.1). For $\Phi \le \Gamma$, in [4] we attached a diagram X_Φ and its thickening $\mathbf{X}_\Phi$ with canonical projections $X_\Phi \to X_\Gamma$, $\mathbf{X}_\Phi \to \mathbf{X}_\Gamma$. Here $\mathbf{X}_\Phi$ is an orientable surface with non-empty boundary $\partial\mathbf{X}_\Phi$. One may think of $\mathbf{X}_\Gamma$ as "$\Gamma\backslash\mathbf{H}$ with the cusp cut off". This makes the "cuspidal infinity" more tangible—for example, one gets the following useful characterizations: $\Phi \le \Gamma$ is Neumann (respectively 1-cycloidal) if and only if $\partial\mathbf{X}_\Phi$ is connected and non-compact (respectively connected and compact). Pinching each circle in X_Φ to a point one obtains a graph Y_Φ whose structure suggests the notion of an $(m_1, \dots, m_n)$*-semiregular* graph; cf. (2.4). If Φ is a Neumann subgroup then the image of $\partial\mathbf{X}_\Phi$ in Y_Φ is a special type of Eulerian path which we simply call *admissible.* This provides a natural explanation of the initially intriguing Brenner-Lyndon correspondence between Neumann and 1-cycloidal subgroups of the modular group and the Eulerian paths in cuboid graphs. A natural extension of their results is as follows: *the conjugacy classes of Neumann (respectively 1-cycloidal) subgroups of Γ are in one-to-one correspondence with the admissible Eulerian paths in $(m_1, \dots, m_n)$-semiregular graphs.*

(1.7) For Γ as in (1.4.1) and $\Phi \le \Gamma$ we have by Kurosh's theorem,

$$\Phi \approx F_r * \prod_{i=1}^{*n} \left(\prod_{j \in J_i}^{*} \Phi_{ij} \right), \tag{1.7.1}$$

where F_r denotes the free group of rank r and $\Phi_{ij} \cong \mathbf{Z}_d$, $d|m_i$, are conjugates to subgroups of Γ_i. In (1.7.1) we assume that $\Phi ij \not\cong \{e\}$ with the understanding that if J_i is empty then $\prod^*_{j\in J_i} \Phi_{ij} = \{e\}$. Let

$$r_i(d) = \#\{\Phi_{ij} = \mathbf{Z}_{m_i/d}\}, \qquad d|m_i, d < m_i. \tag{1.7.2}$$

The numbers $r_i(d)$ may be possibly infinite. In Section 4 we prove a structure-and-realization theorem for Neumann subgroups. For example, *if at most one m_i is even, then Φ as in* (1.7.1) *is realizable as a Neumann subgroup if and only if either* (1) $r = \infty$ *or* (2) r *is an even integer and* $r_i(1) = \infty$ *for at least* $n - 1$ *values of* i. If there are two even m_i's there is a curious new family of Neumann subgroups (cf. (2.11)) of which there is no analogue in the case of the modular group. This family makes the full structure theorem a bit complicated, but the underlying geometric idea is simple. For details see Section 4.

(1.8) Finally in Section 5 we give some geometric constructions of subgroups which are maximal, or maximal with respect to some additional properties such as Neumann, 1-cycloidal, non-parabolic but non-Neumann,....

I wish to thank W. W. Stothers for drawing my attention to [2].

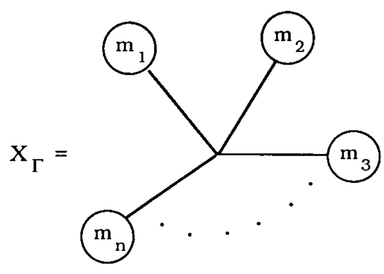
2. Preliminaries

(2.0) Throughout this section let Γ, Γ_i, x_i, u be as in (1.4) and (1.5). We use the terminology introduced there.

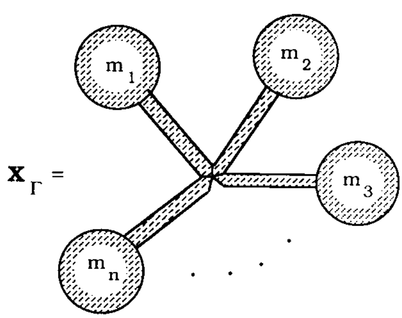
(2.1) **Proposition.** *A Neumann subgroup is maximal among non-parabolic subgroups.*

Proof. Let Φ be a Neumann subgroup of Γ, and $P = \langle u \rangle$. So P acts simply transitively on Γ/Φ. The isotropy subgroup of P at $a\Phi$ is $P \cap a\Phi a^{-1} = \{e\}$. So $a^{-1}Pa \cap \Phi = \{e\}$, that is, Φ is a non-parabolic subgroup. If Ψ is a subgroup of Γ which properly contains Φ then P acts transitively but not simply transitively on Γ/Ψ. But since $P \approx \mathbf{Z}$, this means that the P-action on Γ/Ψ is ineffective and $|\Gamma/\Psi| < \infty$. Hence $P \cap \Psi \neq \{e\}$, that is, Ψ contains parabolic elements. So Φ is maximal among non-parabolic subgroups.

(2.2) As in [4], let X_Γ be a diagram for Γ and $\mathbf{X}_\Gamma$ its thickening



(2.2.1)

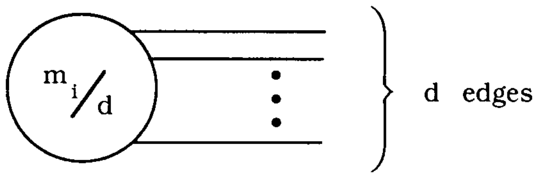


https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

In the thickening each circle with $m < \infty$ is replaced by a disk and each segment is replaced by a rectangular sheet. These pieces are attached as shown in the figure so as to form a compact orientable surface with boundary. The case of a circle with $m = \infty$ does not occur in this paper. In that case a circle would be replaced by an annulus.

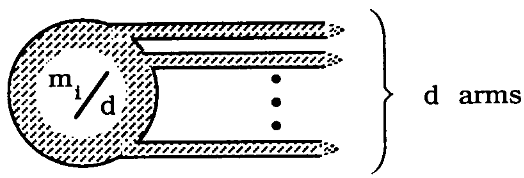
A building block of type i has the form

(2.2.2)



and is denoted by $B_i(d)$. A diagram X_Φ is built out of such $B_i(d)$'s and there is a canonical projection $X_\Phi \to X_\Gamma$. The thickening of $B_i(d)$ is

(2.2.3) $\mathbf{B}_i(d) =$



The thickening $\mathbf{X}_\Phi$ of X_Φ is built out of $\mathbf{B}_i(d)$'s. Notice that $\mathbf{X}_\Phi$ is an orientable surface with boundary $\partial\mathbf{X}_\Phi$. There is a canonical projection $p\colon \mathbf{X}_\Phi \to \mathbf{X}_\Gamma$ and also a "thinning" map $\mathbf{X}_\Phi \to X_\Phi$ (the restriction $p|_{\operatorname{int}\mathbf{X}_\Phi}\colon \operatorname{int}\mathbf{X}_\Phi \to \operatorname{int}\mathbf{X}_\Gamma$ is a branched covering of surfaces; if Γ is realized as an orientation-preserving, properly discontinuous group of homeomorphisms of $\mathbf{R}^2$ then $p|_{\operatorname{int}\mathbf{X}_\Phi}$ is equivalent to the canonical map $\Phi\backslash\mathbf{R}^2 \to \Gamma\backslash\mathbf{R}^2$). The shape of $\mathbf{B}_i(d)$ may be described as "a closed disk with d arms." Each of the dotted edges at the end of an arm is its *half-outlet*; together they form an *outlet*. In $\mathbf{X}_\Phi$ the outlets come in groups of n. So we may use the obvious and suggestive terminology of an *angle formed by the half-outlets*. For example the interior angle formed by the half-outlets of an arm is $2\pi/n$. In $\mathbf{X}_\Phi$ one half-outlet of an arm of a $\mathbf{B}_i(*)$ is joined at half-outlet of arm of $\mathbf{B}_{i+1}(*)$, where the subscript i is counted $\bmod n$. In the sequel, it will be important to keep in mind that

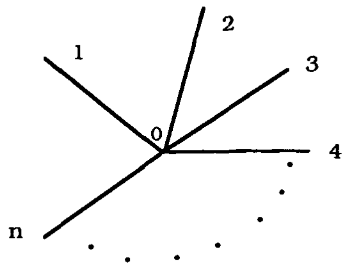
(2.2.4) $$\partial\mathbf{B}_i(d) = \{\mathbf{B}_i(d) \cap \partial\mathbf{X}_\Phi\} \cup \{\text{the outlets}\}.$$

(2.3) Pinching each circle in X_Φ to a point one gets a graph Y_Φ. Again one has a canonical projection denoted by $p\colon Y_\Phi \to Y_\Gamma$. Notice that the *terminal vertices*, that is, the vertices of valence 1, of Y_Φ are precisely the images of m_i

https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

in X_Φ. The vertices adjacent to terminal vertices will be called *sub-terminal vertices*. Now

(2.3.1) Y_Γ =



has $n+1$ vertices—the image of m_i is numbered i, and the "base-vertex" is numbered 0. So the vertices of Y_Φ are divided into $n+1$ disjoint subsets:

(2.3.2) $$\alpha_i = \{v | p(v) \text{ has number } i\}.$$

The structure of Y_Φ motivates the following

(2.4) Definition. Let n and $m_1, \ldots, m_n$ be positive integers, each at least 2. An $(m_1, \ldots, m_n)$-semiregular graph is a connected graph G whose vertices are divided into $n+1$ disjoint subsets α_i, $i = 0, 1, \ldots, n$, such that

(a) $v \in \alpha_i$ implies valence $v = n$ (respectively a divisor of m_i) if $i = 0$ (respectively if $i \geq 1$),

(b) each edge of G has one end in α_0 and the other in α_i, $i \geq 1$,

(c) given $v \in \alpha_0$ and $i \geq 1$, there is a unique edge joining v to a vertex in α_i.

Clearly Y_Φ, as in (2.3), is an $(m_1, \ldots, m_n)$-semiregular graph.

If $n = 2$ (respectively some m_i is even), then the vertices in α_0 (respectively certain vertices in α_i) have valence 2, and if convenient may well be not counted as vertices. Thus for example, not counting the vertices in α_0, an (m_1, m_2)-semiregular graph is a bipartite graph. Again if G is a $(2, k)$-semiregular graph such that all vertices in α_1 (respectively α_2) have valence 2 (respectively k), then not counting the vertices either in α_0 or in α_1, one has a k-regular graph in the usual sense. Thus if $\Gamma = \mathbb{Z}_2 * \mathbb{Z}_k$, and $\Phi \leq \Gamma$ is *torsion-free* then Y_Φ may be considered as a k-regular graph. In particular, corresponding to torsion-free subgroups of the modular group one gets cubic graphs.

(2.5) Remark. Let G be an $(m_1, \ldots, m_n)$-semiregular graph. Then the edges at a vertex in α_0 come equipped with a natural cyclic order. Now suppose at each vertex $v \in \alpha_i$, $i \geq 1$, we specify some cyclic order among the edges incident with v. Then we may replace each $v \in \alpha_i$, $i \geq 1$, by a circle and attach the v-ends of the edges incident at v to the circle consistent with the prescribed cyclic order, and obtain a diagram X. This diagram has a canonical projection $p: X \to X_\Gamma$; see the proof of [4, Theorem 1]. X may be considered as X_Φ for a subgroup Φ whose conjugacy class is well defined. Thus G is isomorphic to Y_Φ for some $\Phi \leq \Gamma$.

(2.6) Taking a base point in $\partial \mathbf{X}_\Gamma$ we may represent $u = x_1 \cdots x_n$ by the oriented boundary $\partial \mathbf{X}_\Gamma$. Now $p^{-1}(\partial \mathbf{X}_\Gamma) = \partial \mathbf{X}_\Phi$, so the components of $\mathbf{X}_\Phi$ are in one-to-one correspondence with the double cosets $P \backslash \Gamma / \Phi$. If C is a component of $\partial \mathbf{X}_\phi$ and $P|_C: C \to \partial \mathbf{X}_\Gamma$ has degree d (possibly infinite), then d is the number of points in the corresponding P-orbit in Γ/Φ. In particular C is non-compact if and only if $d = \infty$, which is if and only if the P-action on the corresponding orbit is effective. Clearly one gets

PROPOSITION. (1) *Φ is a non-parabolic subgroup if and only if $\partial \mathbf{X}_\Phi$ has no compact component.*

(2) *Φ is a Neumann (respectively 1-cycloidal) subgroup if and only if $\partial \mathbf{X}_\Phi$ is connected and non-compact (respectively connected and compact).*

(2.7) We recall some elementary facts from the topology of surfaces. Let M be any connected surface possibly with non-empty boundary. A connected, compact subsurface S of M is said to be *tight* if $M - \operatorname{int} S$ has no compact component. Notice that if S is a compact subsurface then $M - S$ has only finitely many components. So if S is compact and connected then $S_1 = S \cup \{\text{compact components of } M - \operatorname{int} S\}$ is a tight subsurface. It is now clear that M admits an *exhaustion* by tight subsurfaces, that is, a sequence S_i, $i = 1, 2, \ldots,$ of tight subsurfaces such that $S_i \subseteq \operatorname{int} S_{i+1}$ and $M = \bigcup_i S_i$.

Now suppose that the fundamental groups of M based at $*$ is finitely generated. So there exist finitely many based loops C_i such that $\pi_1(M, *) = \langle [C_i] \rangle$, where $[C_i]$ denotes the homotopy class of C_i. One says that an arc-connected subset A of M *carries* π_1 if the canonical map $\pi_1(A) \to \pi_1(M)$ is surjective. Clearly any arc-connected subset A containing $\bigcup C_i$ carries π_1. Now let S be a tight subsurface which contains $\bigcup C_i$. In this case in fact the canonical map $\pi_1(S) \to \pi_1(M)$ is an isomorphism and it is easy to see that each component of $M - S$ is either a cylinder or a disk. If $\partial M \neq \emptyset$ these cylinders or disks may also have non-empty boundary.

(2.8) We apply the considerations in (2.7) to $\mathbf{X}_\Phi$. Let S be a tight subsurface of $\mathbf{X}_\Phi$. Then for each building block $\mathbf{B}_i(d)$ we see that a component of $S \cap \mathbf{B}_i(d)$ is also tight. Let S_1 be the union of S and all $\mathbf{B}_i(d)$'s which intersect S in a subset with non-empty interior. Then S_1 has the additional property

$$\partial S_1 = (S_1 \cap \partial \mathbf{X}_\Phi) \cup A, \tag{2.8.1}$$

where A is the union of the half-outlets on some arms of the building blocks. Now suppose Φ is as in (1.7.1). Then its free part F_r may be identified with

$\pi_1(Y_\Phi)$ or $\pi_1(\mathbf{X}_\Phi)$ (see the discussion in [5, Section 2]. Suppose that $r < \infty$. So there exist tight subsurfaces of $\mathbf{X}_\Phi$ which carry π_1; cf. (2.7). We call a tight subsurface *characteristic* if it carries π_1 and has the additional property stated in (2.8.1).

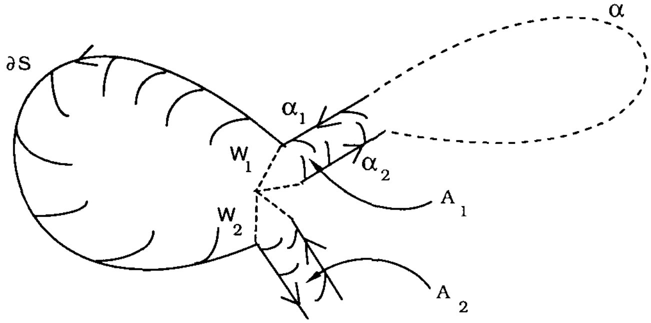
From the above discussion it is clear that *if* $r < \infty$, $\mathbf{X}_\Phi$ *admits an exhaustion by characteristic subsurfaces.*

(2.9) PROPOSITION. *Let* Φ *as in* (1.7.1) *be a Neumann subgroup of* Γ *and* $r < \infty$. *Let* S *be a characteristic subsurface of* $\mathbf{X}_\Phi$. *Then* ∂S *is connected and contains exactly one pair of half-outlets making an exterior angle* $2\pi/n$; *cf.* (2.2). *Moreover* $\operatorname{int}(\mathbf{X}_\Phi - S)$ *is homeomorphic to an open disk and* $\partial(\mathbf{X}_\Phi - S)$ *has two components, each homeomorphic to an open interval.*

PROOF. Since S is characteristic, we have $\partial S = \{S \cap \partial \mathbf{X}_\Phi\} \cup A$, where A is a union of half-outlets. Since ∂S is compact, $\partial \mathbf{X}_\Phi$ is connected and noncompact (cf. (2.6)) we see that each component of ∂S must intersect A as well as $\partial \mathbf{X}_\Phi$. Notice that the half-outlets in A come in pairs—each pair forms a connected arc, and different pairs form disjoint arcs.

First we claim that ∂S is connected. Suppose C_1, C_2 are two disjoint components of ∂S. Since C_1, C_2 contain points of $\partial \mathbf{X}_\Phi$, and $\partial \mathbf{X}_\Phi$ is connected, there is an arc $\varepsilon \subseteq \partial \mathbf{X}_\Phi$ joining a point p_1 in C_1 to a point p_2 in C_2. But since S is connected there is an arc $\beta \subseteq S$ joining p_1 to p_2 and passing though a base-point $*$. But then $\alpha \cup \beta$ forms a based loop whose homotopy class is clearly not contained in $\pi_1(S, *)$. This would contradict that S carries π_1. So ∂S is connected.

(2.9.1)



Next suppose, if possible, that there are two pairs of half-outlets, each pair forming an arc. Then $\partial S - A$ has two components which must be connected by an arc $\subseteq \partial \mathbf{X}_\Phi$, and we get a contradiction exactly as above.

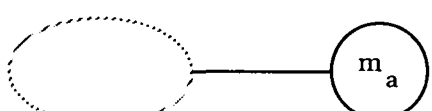
Next suppose that the pair w_1, w_2 of half-outlets makes an exterior angle strictly greater than $2\pi/n$ (see the figure in (2.9.1)). Then the arms A_1, A_2 in

$\mathbf{X}_\Phi - \operatorname{int} S$ incident with w_1 and w_2 are distinct. Again since $\partial \mathbf{X}_\Phi$ is connected, the components α_1, α_2 of $A_1 \cap \partial \mathbf{X}_\Phi$ are joined by an arc $\alpha \subseteq \partial \mathbf{X}_\Phi$. Clearly $\alpha \cap S = \varnothing$. Now $\partial A_1 \cup \alpha$ forms a Jordan curve outside S. Since S carries π_1 this Jordan curve must bound a disk. But then $\mathbf{X}_\Phi - \operatorname{int} S$ would have a compact component and S would not be tight. This contradiction shows that the exterior angle formed by w_1, w_2 must be $2\pi/n$, and so $w_1 \cup w_2$ is an outlet of an arm lying outside $\operatorname{int} S$. This arm connects $\mathbf{X}_\Phi - \operatorname{int} S$ to S. In particular $\mathbf{X}_\Phi - \operatorname{int} S$ has only one component. From the remarks in (2.7) it is now clear that $\operatorname{int}(\mathbf{X}_\Phi - \operatorname{int} S)$ is homeomorphic to a disk and $\partial(\mathbf{X}_\Phi - \operatorname{int} S)$ has two components each sharing one endpoint of $\partial S - A$.

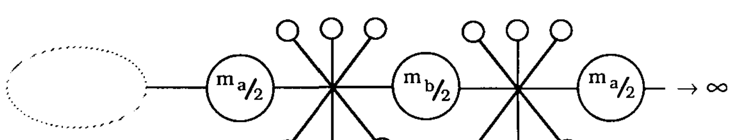
(2.10) The above proposition may be used to get an intuitive understanding of a Neumann subgroup Φ with $r < \infty$. Let $S_1 \subset S_2 \subset \cdots$ be an exhaustion of $\mathbf{X}_\Phi$ by characteristic subsurfaces. Each $S_{k+1} - \operatorname{int} S_k$ is homeomorphic to a closed disk. Also each S_k has exactly one pair of half-outlets with exterior angle $2\pi/n$. Inserting an appropriate $\mathbf{B}_i(1)$ in this outlet we obtain a new diagram $\widetilde{S}_k \approx \mathbf{X}_{\Phi_k}$ where Φ_k is a 1-cycloidal subgroup. Thus we get a sequence Φ_k, $k = 1, 2, \ldots$, of 1-cycloidal groups so that $\mathbf{X}_{\Phi_k}$ contains some $\mathbf{B}_i(1)$ and $\mathbf{X}_{\Phi_{k+1}}$ is obtained from $\mathbf{X}_{\Phi_k}$ by removing some $\mathbf{B}_i(1)$, and inserting some $\mathbf{B}_i(d)$, $d > 1$, together with some outer building blocks so that the union of the newly inserted building blocks is a subset homeomorphic to a closed disk. We express this by saying that Φ *is obtained by unfolding a sequence of 1-cycloidal subgroups* Φ_k.

(2.11) We shall now describe a special "unfolding" of a single 1-cycloidal subgroup. It will be important in the structure theory of Neumann subgroups in Section 4. Suppose we have two m_i's, say m_a, m_b, which are even integers. Let Φ_0 be a 1-cycloidal subgroup so that X_{Φ_0} contains either $B_a(1)$ or $B_b(1)$, say the first. Then we can obtain a Neumann subgroup Φ as follows, which is best described by its diagram X_Φ.

Suppose

(2.11.1) $X_{\Phi_0} =$ 

Let

(2.11.2) $X_\Phi =$ 

https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

Here all the unlabelled building blocks in the newly inserted portion are $B_i(1)$'s, $i \neq a, b$. We shall say that Φ *is a simple* (m_a, m_b)*-unfolding of a* 1*-cycloidal subgroup* Φ_0.

(2.12) REMARK. Let Φ as (2.7.1) be a Neumann subgroup with $r = \infty$. Then $\mathbf{X}_\Phi$ contains no characteristic subsurface. But it is not difficult to see that still $\mathbf{X}_\Phi$ admits an exhaustion S_k, $k = 1, 2 \ldots$, by tight subsurfaces which satisfy the property stated in (2.8.1) and such that ∂S_k may contain several pairs of half-outlets. Filling these pairs by suitable $\mathbf{B}_i(1)$'s we obtain $\widetilde{S}_k \approx \mathbf{X}_{\Phi_k}$ where Φ_k is a 1-cycloidal subgroup. In this sense Φ can still be considered as an "unfolding of a sequence of 1-cycloidal subgroups."

3. Eulerian paths

Let G be a graph. Each edge of G can be directed in two ways and so corresponds to two *directed edges*, each of which is the *inverse* of the other. A path in G is *reduced* if it contains no consecutive pair of inverse edges. *An Eulerian path* in G is a path which contains each directed edge once and which is reduced except at the terminal vertices.

(3.2) Let G be a $(m_1, \ldots, m_n)$-semiregular graph; cf. (2.4). An *admissible* path in G is a path in which the vertices occur in the following consecutive order:

(3.2.1) $$\cdots v_1 w_1 v_2 w_2 \cdots, \qquad v_i \in \alpha_0,\ w_i \in \alpha_{k+i},$$

where k is some fixed integer and $\alpha_{k+i} = a_j$, where j is the unique positive integer such that $1 \leq j \leq n$, $k + i = j(n)$.

(3.3) THEOREM. *Let* Γ *be as* (1.4.1). *Then the conjugacy of Neumann* (*respectively* 1*-cycloidal*) *subgroups of* Γ *are in one-to-one correspondence with the admissible Eulerian paths in infinite* (*respectively finite*) $(m_1, \ldots, m_n)$*-semiregular graphs.*

PROOF. Let Φ be a Neumann (respectively 1-cycloidal) subgroup Γ. Then Y_Φ is an $(m_1, \ldots, m_n)$-semiregular graph. Since Φ is Neumann (respectively 1-cycloidal), Y_Φ is infinite (respectively finite). Now orient $\mathbf{X}_\Phi$; this also orients $\partial \mathbf{X}_\Phi$. If A is an arm of a building block of $\mathbf{X}_\Phi$ then $A \cap \partial \mathbf{X}_\Phi$ consists of two edges, which, under the canonical projection $\mathbf{X}_\Phi \to X_\Phi \to Y_\Phi$, project onto a pair of mutually inverse directed edges. It follows that the image of $\partial \mathbf{X}_\Phi$ in Y_Φ is an admissible Eulerian path.

https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

Conversely, let G be an infinite (respectively finite) $(m_1, \dots, m_n)$-semi-regular graph, and E an admissible Eulerian path in G. Let $v \in \alpha_i$, $i \geq 1$. Introduce a cyclic order among the (undirected) edges incident with v as follows: an edge f cyclically follows e if and only if in E the directed edge e ending in v follows the directed edge f beginning at v. By the remark in (2.5) we can construct an infinite (respectively finite) diagram X which corresponds to a conjugacy class of a subgroup Φ. But the existence of E also shows that $\partial \mathbf{X}$ is connected and is non-compact (respectively compact) so Φ is Neumann (respectively 1-cycloidal).

It is easy to see that this establishes the one-to-one correspondence asserted in the theorem.

4. A structure theorem

(4.0) Throughout this section Γ is as in (1.4.1) and Φ is as in (1.7.1) and we use the notation used there. If $\Gamma \approx \mathbf{Z}_2 * \mathbf{Z}_2$ it is easy to see that the two conjugacy classes of subgroups isomorphic to $\mathbf{Z}_2$ precisely consist of all the Neumann subgroups in Γ. Henceforth we shall assume that $\Gamma \not\approx \mathbf{Z}_2 * \mathbf{Z}_2$.

(4.1) Proposition. *If $r = \infty$ then Φ is realizable as a Neumann subgroup.*

Proof. The details of this proof are similar to (and simpler than) those of [5, Theorem (1.5)], which deals with the case of 1-cycloidal subgroups. So we shall be brief. First of all, the diophantine condition (see [5, (3.2)]), needed there is no longer necessary since the "difficulties can be thrown off to infinity." Recall that for $d | m_i$, $d < m_i$

$$r_i(d) = \#\{\phi_{ij} | \Phi_{ij} \simeq \mathbf{Z}_{m_i/d}\},$$

which may be infinite. We set $r_i(m_i) = \infty$. Choose $r_i(d)$ copies of $B_i(d)$'s; cf. (2.2.2). The objective is to construct a diagram X with these building blocks so that $\mathbf{X}$ has infinite genus and $\partial \mathbf{X}$ is connected and non-compact. Using all $\mathbf{B}_1(*)$'s and some of the $\mathbf{B}_2(*)$'s construct a complex $\mathbf{H}$ homeomorphic to the closed upper half space so that $\partial \mathbf{H}$ contains infinitely many pairs of half-outlets. (*Note.* That $\mathbf{H}$ contains infinitely many pairs of half-outlets is obvious for $n \geq 3$. For $n = 2$ this would fail exactly when $m_1 = 2 = m_2$. We have explicitly excluded this case in (4.0).) Now attach the remaining building blocks appropriately at these half-outlets so as to get $\mathbf{X}$ with the required properties.

https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

(4.2) PROPOSITION. *If Φ with $r < \infty$ is realizable as a Neumann subgroup then r is an even integer.*

PROOF. Indeed, $F_r \approx \pi_1(\mathbf{X}_\Phi)$. If S is a characteristic subsurface, we observed in (2.7), (2.8) that $\pi_1(S) \approx \pi_1(\mathbf{X}_\Phi)$ and S is a compact orientable surface with one boundary component. So $r = 2g$ where g is a genus of S.

(4.3) PROPOSITION. *Let Φ with $r = 2g < \infty$ be realizable as a Neumann subgroup. Then either*

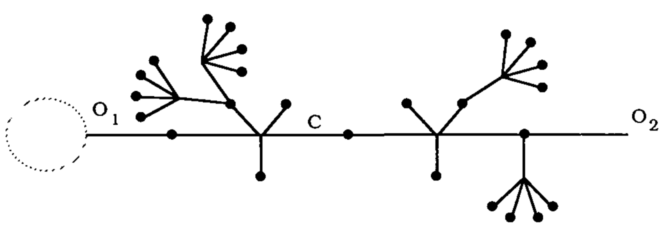
(1) $r_i(1) = \infty$ *for* $\geq n-1$ *values of* i, *or,*
(2) (A) $r_i(1) = \infty$, $r_i(d) < \infty$ *for* $n-2$ *values of* $i \neq a, b$, *say,*
(B) $r_i(2) = \infty$, $r_i(d) < \infty$, $d \neq 2$ *for* $i = a, b$,
(C) Φ *is a simple* (m_a, m_b)*-unfolding of a* 1*-cycloidal subgroup*; *cf.* (2.11).

PROOF. Let $S_1 \subset S_2 \subset \cdots$ be an exhaustion of $\mathbf{X}_\Phi$ by characteristic subsurfaces. Let $D_k = S_{k+1} - \operatorname{int} S_k$, $k = 1, 2, \ldots$. As observed in (2.10), D_k is a closed disk and

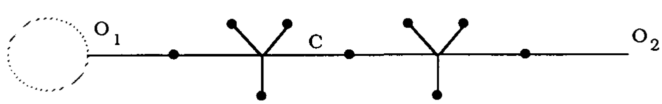
(4.3.1) $\partial D_k = \{D_k \cap \partial \mathbf{X}_\Phi\} \cup \{$the two pairs of half-outlets in $\partial S_k \cup \partial S_{k+1}\}$.

The projection of D_k in Y_Φ has the following two possible forms.

(4.3.2)



(4.3.3)



Here O_1, O_2 are the projections of the pairs of half-outlets in ∂D_k and C is the shortest path joining O_1 to O_2. (Since D_k is homeomorphic to a closed disk, C is unique.) The large dark vertices are in $\bigcup \alpha_i$, $i \geq 1$, and the small ones are in α_0. The two forms are distinguished by the following fact. In (4.3.3) all vertices in α_0 lie on C—hence each is subterminal (cf. (2.3)) and is incident with $n-2$ terminal vertices. In (4.3.2) there are some vertices in α_0 which do not lie on C, and so there are some subterminal ones among

https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

them which are incident with $n-1$ terminal vertices. Now each terminal vertex is an image of a $B_i(1)$ and hence contributes to $r_i(1)$. So it follows that $r_i(1) = \infty$ for at least $n-2$ values of i. Suppose if possible that there actually exist two distinct values a, b of i such that $r_a(1) < \infty$, $r_b(1) < \infty$. Then the infinitely many building blocks $\mathbf{B}_a(1)$'s and $\mathbf{B}_b(1)$'s are contained in some characteristic subsurfaces S_{k_0}. But then for $k > k_0$, D_k is necessarily of the form (4.3.3) and the building blocks with two arms in D_k are necessarily $\mathbf{B}_a(2)$'s and $\mathbf{B}_b(2)$'s. Since X_{k_0} is compact it follows that $r_i(d) < \infty$ for $d \neq 1$, $i \neq a$, b, as well. Finally the discussion in (2.11) shows that in this case Φ must be an (m_a, m_b)-unfolding of a suitable 1-cycloidal subgroup.

(4.4) PROPOSITION. *Let $r = 2g < \infty$, and suppose $r_i(1) = \infty$ for at least $n-1$ values of i. Then Φ is realizable as a Neumann subgroup.*

PROOF. Suppose $r_i(1) = \infty$ for $i \neq 1$. The objective is to construct a diagram X with $r_i(d)$ copies of $B_i(d)$'s, $d < m_i$, and any (possibly infinite) number of copies of $B_i(m_i)$'s so that the thickening diagram $\mathbf{X}$ is an orientable surface of genus g with $\partial\mathbf{X}$ connected and noncompact. Now using finitely many $B_i(d)$'s we can clearly construct a complex S whose thickening $\mathbf{S}$ is a compact, orientable surface of genus g such that $\partial\mathbf{S}$ is connected. (*Note*: if $(m_1, m_2) \neq (2.2)$ or if $g = 0$ we can do this using only $B_i(d)$'s, $i \leq 2$. Otherwise we shall need to use $B_i(d)$'s, $i \leq 3$. Here again we are using the assumption that $\Gamma \not\approx \mathbb{Z}_2 * \mathbb{Z}_2$.) Now using all the remaining $B_1(d)$'s and $B_i(f)$'s, $i \geq 2$, $f \neq 1$, construct a connected complex V whose thickening $\mathbf{V}$ is an orientable surface of genus g such that $\partial\mathbf{V}$ is connected, and contains infinitely many pairs of half-outlets where the infinitely many $\mathbf{B}_i(1)$'s, $i \geq 2$, can be inserted to form $\mathbf{X}$. Clearly $\partial\mathbf{X}$ is connected and $\mathbf{X} = \mathbf{X}_\Psi$ where Ψ is a Neumann subgroup isomorphic to Φ.

(4.5) Combining (2.11), (4.1)–(4.4), we get the following

STRUCTURE THEOREM. *Let Γ be as in* (1.4.1), *$\Gamma \not\approx \mathbb{Z}_2 * \mathbb{Z}_2$ and Φ be given as an abstract group as in* (1.7.1). *Then Φ is realizable as a Neumann subgroup of Γ if and only if one of the following three conditions holds*:

(1) $r = \infty$;
(2) (A) *r is an even non-negative integer,*
    (B) *$r_i(1) = \infty$ for $\geq n-1$ values of i;*
(3) (A) *r is an even non-negative integer,*
    (B) *$r_i(1) = \infty$, $r_i(d) < \infty$, $d \neq 1$, for $n-2$ values of $i \neq a$, b say,*
    (C) *$r_i(2) = \infty$, $r_i(d) < \infty$, $d \neq 2$ for $i = a, b$,*
    (D) *there exists Φ_0, realizable as a* 1*-cycloidal subgroup such that Φ is a simple (m_a, m_b)-unfolding of Φ_0.*

https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

(4.6) REMARK. Suppose Φ is as in (1.7.1) and (3) (A)–(C) are satisfied. Let Ψ_0 be the finite free product of F_r and $\Phi_{ij} \approx \mathbf{Z}_{m_i/d}$, $i \neq a, b$ and $d \neq 1$, or $i = a, b, d \neq 2$. If Φ is realizable as a Neumann subgroup then Φ_0 referred to in (3) (D) is isomorphic to $\Psi_0 * \Theta_0$ where Θ_0 is a finite free product of groups conjugate to Γ_i, $i \neq a$, b, or conjugate to the subgroups of Γ_a (respectively Γ_b) isomorphic to $\mathbf{Z}_{a/2}$ (respectively $\mathbf{Z}_{b/2}$). Moreover Φ_0 must contain at least one factor isomorphic to Γ_a or Γ_b. From the way X_{Φ_0} would be constructed (cf. (2.11)) it is clear that there are only *finitely many* possibilities for Θ_0, and hence, also only finitely many possibilities for Φ_0. Now [5, Theorem (1.5)] gives an effective procedure for deciding whether any of these Φ_0 can be realized as a 1-cycloidal subgroup. Thus one has an effective procedure for deciding realizability of Φ as a Neumann subgroup.

(4.7) REMARK. The condition (3) (D) is *not* a consequence of (3)(A)–(C). For example, take $\Gamma = \mathbf{Z}_4 * \mathbf{Z}_4$ and $\Phi = \prod^* \mathbf{Z}_2$ (infinite product). Write Φ as

$$\prod_{i=1}^{*2}\left(\prod_{j\in J_i}^{*}\Phi_{ij}\right), \quad \Phi_{ij} \approx \mathbf{Z}_2, \quad |J_i| = \infty.$$

It is easy to see that (3)(A)–(C) hold, but Φ is not realizable as a Neumann subgroup.

(4.8) REMARK. We should point out two possible interpretations for the phrase "Φ as in (1.7.1) is realizable as...". If m_i's are pairwise coprime then there is a *unique* value of i for a finite factor of Φ to be conjugate to a subgroup of Γ_i. If two or more m_i's have common factors then there may be a choice for a finite factor of Φ to be interpreted as a particular Φ_{ij}. In our statement of the structure theorem we have tacitly assumed that these choices have already been made. Thus if Φ is only given as an abstract group there may be a bit more freedom first to put it in the form (1.7.1) and then realize as a ....

(4.9) REMARK. The condition (3)(C) of course requires that m_a and m_b are *even* integers. So if there is at most one M_i which is an even integer then the condition (3) is not applicable.

5. Maximal subgroups

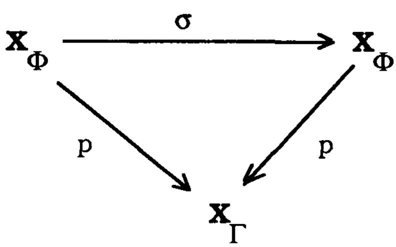
(5.0) In [3] and [13] there are constructions of subgroups of the classical modular group which are maximal among nonparabolic subgroups, and which are different from the ones discovered by Neumann [8], or which are not

https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

Neumann subgroups in the sense of (1.1). These constructions are rather elaborate and require a very careful analysis. In terms of the diagrams X_Φ's one can give such constructions more readily, and in fact one may construct maximal, or maximal and Neumann, or maximal and 1-cycloidal, or maximal and non-parabolic but not Neumann... subgroups.

(5.1) Let Γ be as in (1.4.1) and $\Phi \leq \Gamma$. A *symmetry* of $\mathbf{X}_\Phi$ is simply a branched-covering-transformation of $p: \mathbf{X}_\Phi \to \mathbf{X}_\Gamma$, that is, a homeomorphism $\sigma: \mathbf{X}_\Phi \to \mathbf{X}_\Phi$ such that

(5.1.1)



commutes. Then σ preserves orientation and carries building blocks into building blocks.

Notice that in an unbranched covering space a non-identity covering transformation has no fixed points. But in a branched covering it is not necessarily so.

We say that $\mathbf{X}_\Phi$ *has no fixed-point-free symmetry* if every non-identity symmetry of $\mathbf{X}_\Phi$ has a fixed point.

Notice also that a symmetry $\sigma: \mathbf{X}_\Phi \to \mathbf{X}_\Phi$ induces maps (again denoted by) $\sigma: X_\Phi \to X_\Phi$ and $\sigma: Y_\Phi \to Y_\Phi$, and these maps commute with the thinning map and the canonical projection $X_\Phi \to Y_\Phi$.

(5.2) Orient $\mathbf{X}_\Phi$; this also orients $\partial \mathbf{X}_\Phi$. Let C be a component of $\partial \mathbf{X}_\Phi$. The *pattern along* C is simply the finite or doubly infinite sequence of $\mathbf{B}_i(d)$'s one meets along C while walking in the "positive" direction. Notice that a block $\mathbf{B}_i(d)$ with $d > 1$ (see the picture in (2.3)), is counted k times in the patterns along C if C contains k "circular arcs" on $\mathbf{B}_i(d)$, that is, the components $\partial \mathbf{B}_i(d) - \partial\{\cup \text{arms}\}$.

The pattern is finite if and only if C is compact and in that case the number of terms in the pattern is a multiple of n. We say that the pattern along C is *not periodic* if either (1) C is noncompact and the pattern has no finite period or (2) C is compact, the pattern contains αn elements, $\alpha \in \mathbf{Z}_{>0}$, and (in the cyclic order) the pattern has no period strictly less than αn.

(5.3) Let $B = B_i(d)$ be a building block of X_Φ. The *neighbors* of B are the building blocks at the end of the paths containing two edges emanating from B. So, in all B has $d(n-1)$ neighbors.

(5.4) THEOREM. *Let Γ be as in* (1.4.1) *where all m_i's are primes. Let $\Phi \leq \Gamma$ be as in* (1.7.1). *Assume that* (1) *each $B \approx B_i(m_i)$ in X_Φ has a $B_j(m_j)$ for each $j \neq i$ as a neighbor, and* (2) *either* (a) $r = 0$ *and $\mathbf{X}_\Phi$ has no fixed-point-free symmetry or* (b) *the patterns along different components of $\partial\mathbf{X}_\Phi$ are pairwise distinct and none is periodic. Then Φ is maximal.*

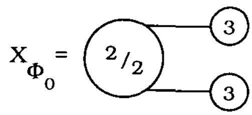
PROOF. Suppose $\Phi \leq \Psi \leq \Gamma$, and consider the branched covering $q: \mathbf{X}_\Phi \to \mathbf{X}_\Psi$. Suppose $\mathbf{X}_\Psi$ contains a branch point. Since m_i's are assumed to be primes this means that there is a building block $B \subseteq X_\Phi$ such that $B \approx B_i(m_i)$ and $q(B) \approx B_i(1)$. But then (1) implies that $q(X_\Phi) = X_\Gamma$, that is, $\Psi = \Gamma$.

Now suppose $\Psi \neq \Gamma$. Hence q is *unbranched.* Under the condition (2a), $\mathbf{X}_\Phi$ is simply connected. But then q is the universal (in particular regular) covering of $\mathbf{X}_\Psi$. Since we assumed that $\mathbf{X}_\Phi$ has no fixed-point-free symmetry it follows that degree $q = 1$, and $\Phi = \Psi$. Under the condition (2b) we see that $q|_{\partial\mathbf{X}_\Phi}$ is a homeomorphism. Also clearly $q^{-1}(\partial\mathbf{X}_\Psi) = \partial\mathbf{X}_\Phi$. So again degree q = degree $q|_{\partial\mathbf{X}_\Phi} = 1$ and $\Phi = \Psi$. Hence Φ is maximal.

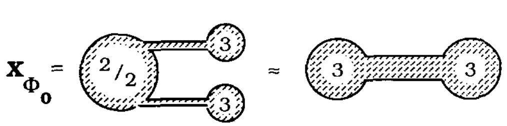
(5.5) REMARKS. (1) Clearly there are many varieties of sufficient sets of conditions for maximality in terms of X_Φ's. For instance one may assume that all but finitely many building blocks of X_Φ have the property stated in (1) and then "mess up" the diagram near these finitely many blocks.

(2) If $n \geq 2$ or two m_i's ≥ 3 the conditions in (5.4) are easy to ensure. The condition that $\mathbf{X}_\Phi$ has no fixed-point symmetry is ensured if we have a compact subsurface $S \subseteq X_\Phi$ satisfying the condition (2.8.1) such that S is homeomorphic to a closed disk and the pattern of the building blocks in S does not repeat in X_Φ, or at least the "distances" among its repetitions do not repeat. Then any symmetry of $\mathbf{X}_\Phi$ would leave S invariant and would have a fixed point by Brouwer's theorem.

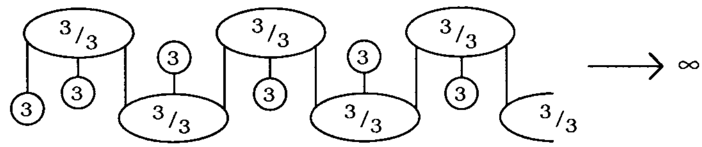
(3) If $n = 2$ and some $m_i = 2$ then the direct application of (5.4) produces only finitely many examples, all of finite index. But excluding the degenerate case $\Gamma \approx \mathbf{Z}_2 * \mathbf{Z}_2$ one may first pass to an appropriate 1-cycloidal subgroup in Γ and then apply the above considerations. For example let $\Gamma = \mathbf{Z}_2 * \mathbf{Z}_3$, and let $\Phi_0 \subset \Gamma$, $\Phi_0 \approx \mathbf{Z}_3 * \mathbf{Z}_3$ whose diagram is



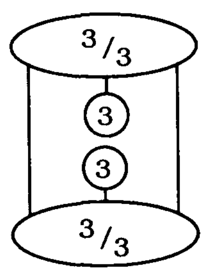
(5.5.1)



Consider $\Phi \subset \Phi_0$ whose diagram is

(5.5.2) $X_\Phi =$ 

Clearly Φ is a Neumann subgroup of Φ_0, and in fact maximal in Φ. Also clearly Φ is a Neumann subgroup of Γ. As a subgroup Γ, the diagram of Φ is obtained from (5.5.2) by sticking in (2/2) on each edge. If we do this sticking and then replace one —(2/2)—(3) by —(2) we obtain a new Neumann subgroup of Γ which is clearly not contained in Φ_0. It would be also maximal in Γ. Making the pattern in (5.52) doubly infinite in the obvious way one obtains a subgroup $\Phi_1 \subset \Phi_0$ for which ∂X_{Φ_1} contains two components both noncompact. This Φ_1 is not Neumann and it is maximal among nonparabolic subgroups in Φ_0 and also in Γ but it is not maximal. For clearly $\Phi_1 \leq \Psi_1 \leq \Phi_0$, where

(5.5.3) $X_{\Psi_1} =$ 

so Φ_1 is not maximal. On the other hand if $\Phi_1 < \Psi < \Phi_0$, then $q: X_{\Phi_1} \to X_\Psi$ must be unbranched; see the argument in (5.4). Now X_{Φ_1} is simply connected so q must be a regular covering. One sees that the only symmetries of X_{Φ_1} are the obvious "horizontal" translations, and so X_Ψ is compact, that is, $(\Phi_0 : \Psi) < \infty$. So Ψ contains parabolic elements. Thus Φ_1 is maximal among non-parabolic subgroups. On the other hand one may start with a doubly infinite version of (5.5.2) where the attachment of (3)— non-periodic. Then one would obtain a maximal-and-non-parabolic subgroup of Φ_0 which is not Neumann. By sticking in —(2) somewhere (as described above) one would obtain such subgroups also in Γ.

https://doi.org/10.1017/S1446788700033085 Published online by Cambridge University Press

References

[1] J. L. Brenner and R. C. Lyndon, 'Nonparabolic subgroups of the modular group', *J. Algebra* **77** (1982), 311–322.

[2] J. L. Brenner and R. C. Lyndon, 'Permutations and cubic graphs', *Pacific J. Math.* **104** (1983), 285–315.

[3] J. L. Brenner and R. C. Lyndon, 'Maximal nonparabolic subgroups of the modular group', *Math. Ann.* **263** (1983), 1–11.

[4] R. S. Kulkarni, 'An extension of a theorem of Kurosh and applications to Fuchsian groups', *Michigan Math. J.* **30** (1983), 259–272.

[5] R. S. Kulkarni, 'Geometry of free products', *Math. Z.* **193** (1986), 613–624.

[6] W. Magnus, 'Rational representations of Fuchsian groups and nonparabolic subgroups of the modular group', *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **9** (1973), 179–189.

[7] W. Magnus, *Noneuclidean tesselations and their groups*, (Academic Press, New York, 1974).

[8] B. H. Neumann, 'Über ein gruppen-theoretisch-arithmetisches problem', *Sitzungsber. Preuss. Akad. Wiss. Math.-Phys. Kl.* **10** (1933).

[9] H. Petersson, 'Über einen einfachen typus von untergurppen der modulgruppe', *Arch. Math.* **4** (1953), 308–315.

[10–12] W. W. Stothers, 'Subgroups of infinite index in the modular group I–III', *Glasgow Math. J.* **20** (1979), 103–114, *ibid.* **22** (1981), 101–118, *ibid.* **22** (1981), 119–131.

[13] C. Tretkoff, 'Nonparabolic subgroups of the modular group', *Glasgow Math. J.* **16** (1975), 90–102.

Department of Mathematics
City University of New York
33 W. 42nd Street
New York, New York 10036-8099
U.S.A.