DOUBLY GENERATED FUCHSIAN GROUPS

A. W. Knapp

In Memory of H. Mirkil

INTRODUCTION

Our main theorem concerns geodesic polygons in hyperbolic geometry; it gives a necessary condition for certain finite sets of real two-by-two matrices to generate discrete groups. We suppose that we are given a compact polygon with a finite, even number of sides, and we suppose that the sides are matched in pairs by linear fractional transformations satisfying an orientation condition. If the group generated by these transformations is discrete, then the theorem is that the translates by the group of this polygon cover most of the points in the hyperbolic plane the same number of times.

This result is stated more precisely and proved in Section 1. The rest of the paper illustrates the theorem by applying it to the following question: When is a doubly generated group of analytic automorphisms of the upper half-plane discrete? The theorem of Section 1 will be used to answer the question for the case where neither generator is hyperbolic. The nature of the criterion we give is that one need only find whether the trace of the product of a well-determined power of one generator by a power of the other generator appears in a list. For example, if both generators are elliptic, the list contains two classes of major cases corresponding to classical presentations of triangle groups, and it contains five classes of exceptional cases corresponding to some geometric configurations discussed at the beginning of Section 3 (see Theorem 2.3).

The polygon theorem is given in Section 1, the criterion for groups with two eliptic generators is stated in Section 2 and proved in Section 3, and doubly generated groups at least one of whose generators is parabolic are treated in Section 4. I wish to thank B. Maskit for suggesting the problem solved by Theorem 2.3 and for his advice connected with these results, and I wish to thank the referee for some improvements in the exposition.

1. POLYGON THEOREM

Throughout this section, U denotes the upper half-plane with the hyperbolic geometry. The hyperbolic area of a set X is m(X), the boundary of X is ∂X, and the image of X under a set D of transformations of U is D(X).

THEOREM 1.1. *Let* P *be an open, simple, finite-sided geodesic polygon in* U *whose closure* $\overline{P}$ *in* U *is compact. Suppose that all the sides of* P *are matched in disjoint pairs by analytic automorphisms of* U. *Say,* $L_a(s_a) = s_{a'}$. *As an orientation condition on the transformations* L_a, *suppose that for each* x *within the side* s_a, *each sufficiently small disc* N *centered at* x *satisfies the inclusion condition*

$$N \subseteq s_a \cup P \cup L_a^{-1}(P).$$

Received August 2, 1967.
Supported by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR Grant No. 335-63 at M.I.T.

Finally, suppose that the group D *generated by the transformations* L_a *is discrete.*

(i) *Then the set* $U - D(\partial P)$ *is open, and every point in it is equivalent to members of* P *under exactly* t *elements of* D, *where* t *is a finite integer independent of the point.*

(ii) *Moreover, if* Q *is a standard open fundamental polygon for* D, *then* $m(P) = tm(Q)$.

Before giving the proof, we make a few remarks. Informally, the orientation condition is that if the s_a are oriented consistently, then L_a carries s_a into $s_{a'}$ with reversed orientation. The nontrivial part of the conclusion of the theorem is the second half of (i). It cannot be changed to say that each point of $U - D(\partial P)$ is equivalent under D to t elements of P, unless we count fixed points of elliptic transformations according to their multiplicities. The proof of the theorem will suggest (but it does not establish definitely) that P actually decomposes into the union of t disjoint fundamental sets for D; however, we shall not need to know whether this stronger assertion is true.

It is easy to see that $U - D(\partial P)$ is open, in other words, that $D(\partial P)$ is closed. The set ∂P is compact, and the discreteness of D implies that if K is any compact set in U, then $d(\partial P) \cap K$ is empty for all but finitely many d in D. Hence $K \cap D(\partial P)$ is compact, and $D(\partial P)$ must be closed.

If x is in U, we denote by t(x) the number of members of D that map x into P. Before proving Theorem 1.1, we shall establish three lemmas about t(x) for x in P.

LEMMA 1.2. *For each* x *in* U, t(x) *is finite.*

Proof. The lemma is an immediate consequence of the compactness of $\overline{P}$ and the discontinuity of discrete groups of automorphisms of U.

LEMMA 1.3. *The function* t(x) *is constant on each component of the open set* $P - D(\partial P)$.

Proof. Let x be in $P - D(\partial P)$, let C be the component of x in $P - D(\partial P)$, and let d be an element of D mapping x into P. Since ∂P disconnects P from the complement of $\overline{P}$, since C is connected and d is continuous, and since d(C) does not meet ∂P, we conclude that d(C) lies entirely in P or entirely outside of $\overline{P}$. But d(x) is in P, and thus d(C) is in P. We have therefore proved that $t(y) \geq t(x)$ for all y in C. By symmetry, equality must hold.

LEMMA 1.4. *If* x *is a point of* ∂P *having a neighborhood in* $D(\partial P)$ *that is merely an arc of a side of* P, *and if* x *is not the image under* D *of a vertex of* P, *then, for each sufficiently small disc* N *about* x, $N \cap P$ *lies in a single component of* $P - D(\partial P)$. *If* N *is sufficiently small, let* x' *be the unique point of* ∂P *with which* x *is matched by a generator* L_a *of* D, *and let* N' *be the image of* N *under* L_a. *Then* $t(N \cap P) = t(N' \cap P)$.

Remarks. The function t is constant on $N \cap P$, by Lemma 1.3, and so $t(N \cap P)$ is a well-defined integer. Because $N' \cap P$ is not the image of $N \cap P$ under L_a, the obvious fact that t assumes the same value on points equivalent under D does not prove the lemma immediately.

Proof. N can be taken as any disc that is hyperbolically centered at x and is so small that its intersection with $D(\partial P)$ consists of a single diameter of N. The half-disc $N \cap P$ is connected and is entirely in $P - D(\partial P)$. Hence it lies in a single component of $P - D(\partial P)$.

Let $\bar{t}(x)$ be the number of elements of D that map x into $\bar{P}$. Then $\bar{t}(x) \geq t(N \cap P)$, because each element of D that maps a point of $N \cap P$ into P must map $N \cap P$ into P and hence must map the closure of $N \cap P$, which contains x, into $\bar{P}$. The point of the proof will be to compute the difference between $\bar{t}(x)$ and $t(N \cap P)$.

Suppose d is an element of D mapping x into P. If M is a sufficiently small subdisc of N, then $d(M \cap P)$ is in P, since P is open. Since $N \cap P$ is connected, $d(N \cap P)$ is in P. Thus the d's that contribute to t(x) also contribute to $t(N \cap P)$. The remaining d's that contribute to $\bar{t}(x)$ map x to a point of ∂P, evidently a point whose local behavior in $D(\partial P)$ is the same as that of x. These d's occur in pairs. In fact, such a d followed by the matching transformation of d(x) (which is uniquely defined, since d(x) is not a vertex of P) is the other member of the pair. By the orientation condition on the matching transformation, one of the two transformations in the pair sends $N \cap P$ outside P and the other sends $N \cap P$ inside P. The d's with $N \cap P$ mapped inside P contribute to $t(N \cap P)$, and the others do not. That is, $\bar{t}(x)$ is the sum of $t(N \cap P)$ and the number of pairs of d's sending x into ∂P. Since $\bar{t}(x')$ is then the sum of $t(N' \cap P)$ and the same number of pairs, and since $\bar{t}(x) = \bar{t}(x')$ for the equivalent points x and x', we obtain the relation $t(N \cap P) = t(N' \cap P)$.

Proof of Theorem 1.1. For the proof of (i), we first note that only finitely many points in ∂P are bad in the sense that they fail to satisfy the conditions of Lemma 1.4. In fact, the number of vertices of P is finite, and hence so is the number of points of ∂P that are images of vertices, under D. Consider the points of ∂P without neighborhoods in $D(\partial P)$ that are simply arcs of sides of P. If there are infinitely many such points, let x be one of their limit points. Each neighborhood of x contains a point different from x that belongs to at least two noncollinear sides of $D(\partial P)$, and hence infinitely many members of D map points of $\bar{P}$ into $\bar{P}$, contrary to the discontinuity of D. It follows that the number of points in $\bar{P} \cap D(\partial P)$ equivalent to the bad points of ∂P is finite.

Next, we show that t(x) is constant for x in $P - D(\partial P)$. In fact, let R be $\bar{P}$ minus the points of $D(\partial P)$ equivalent to bad points. The result of the preceding paragraph shows that R is connected. R consists of the disjoint union of $P - D(\partial P)$ and the set of points of $\bar{P} \cap D(\partial P)$ equivalent to points of Lemma 1.4. We define a function f on R as follows. On $P - D(\partial P)$, $f(x) = t(x)$. If x is in $R \cap D(\partial P)$, $f(x) = t(y)$ for all y in $P - D(\partial P)$ sufficiently close to x. The content of Lemma 1.4 is that f(x) is well-defined. The function f is obviously continuous and integer-valued. Since R is connected, f is constant. Thus t is constant on $P - D(\partial P)$.

To complete the proof of (i), it is enough to show that every point of $U - D(\partial P)$ is equivalent under D with a point of $P - D(\partial P)$. Thus, let y in $U - D(\partial P)$ be given, and take any x in $P - D(\partial P)$. By moving x slightly, we may assume that the geodesic arc g from y to x meets no images of vertices under D, since the set of images of vertices is countable. Since g is compact and the set of images of vertices is closed, g is bounded away from the images of vertices.

Find the point x_1 of $g \cap \partial P$ closest to y. By construction, x_1 is not a vertex, and hence it lies within a side s_1 of P. Let L_1 be the matching transformation sending s_1 into its matched side. By the orientation assumption on L_1, $L_1^{-1}(P)$ contains points farther along g toward y than x_1. Let x_2 be the point of $g \cap L_1^{-1}(\partial P)$ closest to y. Then x_2 is not the image of a vertex, and hence it lies within the image under L_1^{-1} of a side s_2 of P. Let L_2 send s_2 into its matched side. By the orientation assumption, $L_1^{-1} L_2^{-1}(P)$ contains points still farther along

g. Continue in this way, using $L_1^{-1} L_2^{-1} L_3^{-1}, \cdots$. In finitely many steps, the points x_j reach y. In fact, otherwise the preimages in $\overline{P}$ of the new pieces of polygons appearing on g at each stage would accumulate in $\overline{P}$. These preimages are geodesic arcs connecting sides of P, and their lengths tend to 0. Hence they are eventually close to vertices, in contradiction to the fact that g is bounded away from the images of vertices. Thus (i) is proved.

For (ii), recall that m is the measure in U. By complete additivity, $m(D(\partial P)) = 0$ and $m(D(\partial Q)) = 0$. Part (i) and the known covering property of standard fundamental regions thus imply that

$$\sum_{d \in D} \chi_P(d(x)) = t \quad \text{and} \quad \sum_{d \in D} \chi_Q(d(x)) = 1$$

almost everywhere. Therefore

$$\begin{aligned} tm(Q) &= \int_U t\chi_Q(x)\,dm(x) = \int_U \sum_d \chi_P(d(x))\chi_Q(x)\,dm(x) \\ &= \int_U \sum_d \chi_P(x)\chi_Q(d^{-1}(x))\,dm(x) = \int_U \sum_d \chi_P(x)\chi_Q(d(x))\,dm(x) \\ &= \int_U \chi_P(x)\,dm(x) = m(P). \end{aligned}$$

2. GROUPS WITH TWO ELLIPTIC GENERATORS

In this section we give a criterion for deciding when a group of real two-by-two unimodular matrices generated by two elliptic matrices is discrete. The proof of the result, which is an application of Theorem 1.1, appears in Section 3.

Let G be the group of all two-by-two real matrices of determinant one, let σ be the homomorphism sending the matrix $(a, b \mid c, d)$ into the linear fractional transformation $z \to (az + b)/(cz + d)$, and let G^* be the image of σ. The kernel of σ is the two-element center of G. If a subgroup of G is discrete, then its image under σ is discrete in G^*, and if a subgroup of G^* is discrete, then its preimage under σ is discrete in G. This enables us to pass back and forth freely between G and G^*.

If α and β are elliptic elements of G^* that generate a discrete group, α and β are certainly of finite order. If α and β are of finite order and commute, the group they generate is finite and hence discrete. We shall therefore assume that the generators α and β are of finite order and that their fixed points $FP(\alpha)$ and $FP(\beta)$ in the upper half-plane U are distinct. In this case we shall find that if the subgroup D of G^* generated by α and β is discrete and G^*/D is compact (this is the difficult case), then D must be a triangle group of Schwarz:

$$\{A, B: A^p = B^q = (AB)^r = 1\},$$

where $p^{-1} + q^{-1} + r^{-1} < 1$. Thus, before stating the theorem, we shall write down explicitly matrices whose images under σ generate such a group.

If (a, b | c, d) is an elliptic matrix, its conjugacy class within G is determined by the trace a + d and the sign of b (the trace alone is not enough). To verify this assertion, conjugate (cos θ, sin θ | - sin θ, cos θ) by the most general member of G and examine the result. The elliptic matrix (a, b | c, d) will be said to have *extreme negative trace* if $a + d = 2\cos(\pi - \pi/n)$ for an integer n, and (a, b | c, d) is *normalized* if it has extreme negative trace and if b is positive. If A is an elliptic matrix of finite order, then there exists an integer k relatively prime to the order of $\sigma(A)$ such that A^k has extreme negative trace. In this case, exactly one of the powers A^k and A^{-k} is normalized. Thus, in deciding whether a subgroup of G with two elliptic generators is discrete, we lose no generality by assuming that both generators are normalized, since the images under σ of the normalized generators generate the same subgroup of G^* as the images of the given generators.

Let A and B be normalized elliptic matrices with distinct fixed points. We wish to show that the image under σ of the group generated by A and B is generated geometrically from a polygon, in the sense of Theorem 1.1. From the proof we shall easily see the conditions on A and B in order that $\sigma(A)$ and $\sigma(B)$ be the standard generators of a triangle group, and we shall be in a position to apply Theorem 1.1 to obtain our criterion for discreteness of the group generated by $\sigma(A)$ and $\sigma(B)$.

To see that $\sigma(A)$ and $\sigma(B)$ generate a group geometrically, we shall conjugate them so that their fixed points are conveniently located in U. Namely, we can conjugate A and B by a single matrix chosen so that the fixed point of $\sigma(A)$ is at i and the fixed point of $\sigma(B)$ occurs at ui for some real number $u > 1$. Then the new A and B have the forms

$$(1)\quad A = \begin{pmatrix} \cos(\pi - \pi/\ell) & \sin(\pi - \pi/\ell) \\ -\sin(\pi - \pi/\ell) & \cos(\pi - \pi/\ell) \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \cos(\pi - \pi/m) & b \\ c & \cos(\pi - \pi/m) \end{pmatrix},$$

where ℓ and m are positive integers, $|c| < |b|$, $b > 0$, $c < 0$, and det B = 1. According as AB is elliptic, parabolic, or hyperbolic, $\sigma(A)$ and $\sigma(B)$ have the effect suggested in Figure 1, 2, or 3. We shall verify this assertion only in the elliptic case, the other cases being similar. (There is, however, a slight difference in the manner in which the sides are defined in Figure 3 and in Figures 1 and 2. In Figure

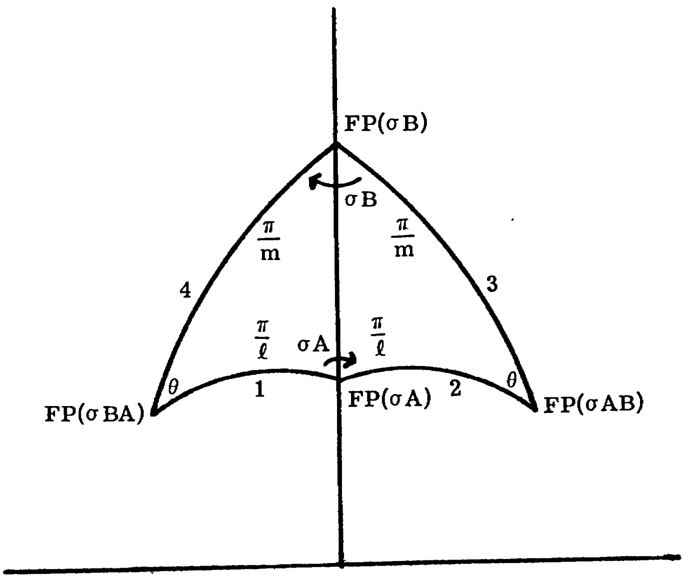


Figure 1

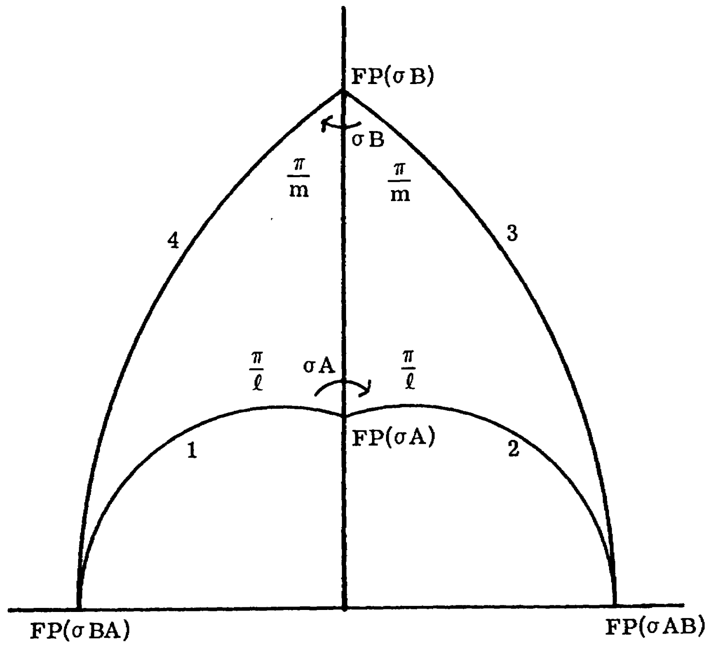


Figure 2

3, the sides are defined by the two fixed points and the angles at the imaginary axis, whereas in Figures 1 and 2 they are defined by the four fixed points.)

LEMMA 2.1. *Let* $M = (a, b \mid c, d)$ *be an elliptic matrix with*

$$a + d = 2 \cos \phi \quad \textit{and} \quad \operatorname{sign} b = \operatorname{sign} \sin \phi .$$

Then, at the fixed point of $\sigma(M)$ *in* U, $\sigma(M)$ *rotates directions counterclockwise through the angle* 2ϕ.

Proof. Neither the hypothesis nor the conclusion is affected by conjugation within G, and we may thus assume that $M = (\cos \phi, \sin \phi \mid -\sin \phi, \cos \phi)$. The fixed point

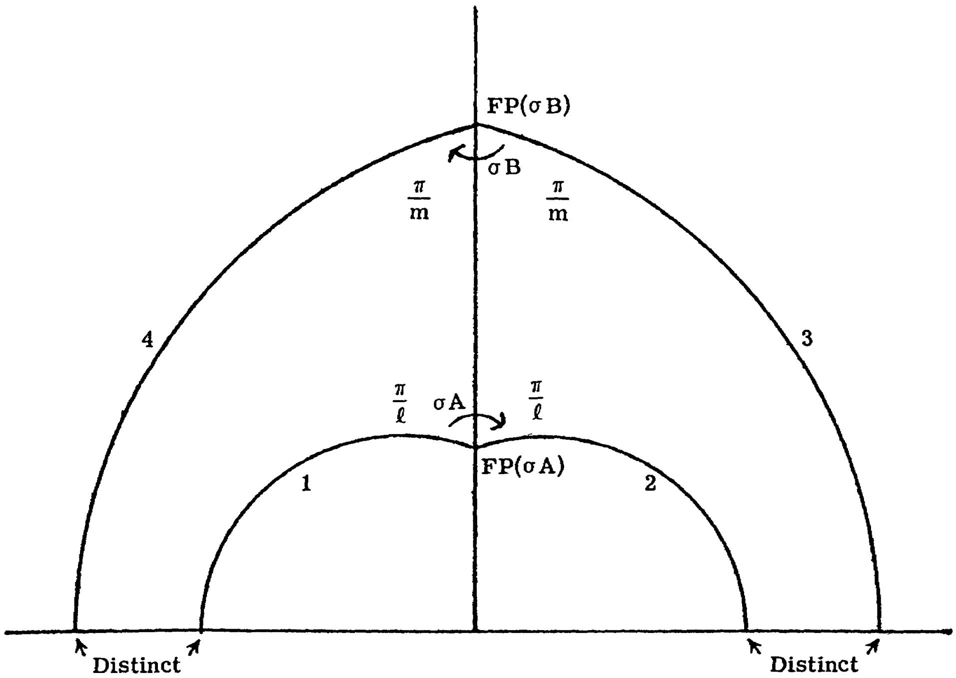


Figure 3

of M is i, and the rotation at i is counterclockwise through the angle $\arg[\sigma(M)'(i)]$, which is easily seen to be 2ϕ.

Return to the verification that $\sigma(A)$ and $\sigma(B)$ map the vertices and sides in Figure 1 as suggested by the picture. The sides in this picture are defined as the geodesic arcs connecting the indicated fixed points. Since $\sigma(A)$ carries $FP(\sigma BA)$ into $FP(\sigma AB)$ and $\sigma(B)$ carries $FP(\sigma AB)$ into $FP(\sigma BA)$, $\sigma(A)$ carries side 1 into side 2 and $\sigma(B)$ carries side 3 into side 4. Lemma 2.1 shows that $\sigma(A)$ rotates directions at i counterclockwise through $2\pi - 2\pi/\ell$ or clockwise through $2\pi/\ell$. Thus the indicated angle between side 1 and side 2 is $2\pi/\ell$. Similarly, the indicated angle between side 3 and side 4 is $2\pi/m$. The real part of $FP(\sigma AB)$ is

$$\frac{(c+b)\sin(\pi - \pi/\ell)}{2(c\cos(\pi - \pi/\ell) - \cos(\pi - \pi/m)\sin(\pi - \pi/\ell))}. \tag{2}$$

The numerator has the sign of b, which is positive, and both terms in the denominator are positive. Hence $FP(\sigma AB)$ lies in the right half-plane. To complete the verification, it is enough to show that $FP(\sigma BA)$ is the reflection in the imaginary axis of $FP(\sigma AB)$. In fact, the reflection in the imaginary axis of $FP(\sigma AB)$ is the fixed point of σ of

$$\begin{pmatrix}1&0\\0&-1\end{pmatrix}AB\begin{pmatrix}1&0\\0&-1\end{pmatrix} = \begin{pmatrix}1&0\\0&-1\end{pmatrix}A\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}1&0\\0&-1\end{pmatrix}B\begin{pmatrix}1&0\\0&-1\end{pmatrix} = A^{-1}B^{-1} = (BA)^{-1},$$

and hence it is the fixed point of σBA. The validity of Figure 1 is verified.

We shall need to know the angles θ in Figure 1. If the trace of AB is $2\cos(\pi+\phi)$ with $0 < \phi < \pi$, then $\theta = \phi$. In fact, the lower left entry of AB is half the denominator of expression (2), and that is positive. Hence the upper right entry of AB is negative (since AB is assumed to be elliptic) and has the same sign as $\sin(\pi+\phi)$. By Lemma 2.1, σAB rotates directions about $FP(\sigma AB)$ counterclockwise through $2(\pi+\phi)$ or counterclockwise through 2ϕ. An argument similar to that in the preceding paragraph then shows that $\theta = \phi$.

PROPOSITION 2.2. *If* A *and* B *have the form* (1) *and if* n *is any positive integer, then there are only finitely many choices of* b *and* c *such that* $(\sigma AB)^n = 1$. *If* b *and* c *are chosen so that*

$$bc = -\sin^2(\pi - \pi/m), \qquad |c| < \sin(\pi - \pi/m), \tag{3a}$$

$$2\cos(\pi - \pi/m)\cos(\pi - \pi/\ell) + (c-b)\sin(\pi - \pi/\ell) = 2\cos(\pi + \pi/n), \tag{3b}$$

then AB *has extreme negative trace* $2\cos(\pi + \pi/n)$, *and if* AB *has this extreme negative trace, then* $\sigma(A)$ *and* $\sigma(B)$ *generate the* (ℓ, m, n)*-triangle group of Schwarz.*

Proof. Equations (1) require that b and c satisfy (3a). If $(\sigma AB)^n = 1$, then the trace of AB must be $2\cos(\pi + s\pi/n)$ for an integer s with $0 < s < n$. This is the condition that (3b) hold, but with the right side replaced with $2\cos(\pi + s\pi/n)$. For each s, the resulting system consisting of (3a) and the modified (3b) has at most one solution, and the first statement of the proposition follows.

If the system is solved with $s = 1$, then the trace of AB is $2\cos(\pi + \pi/n)$, and so θ in Figure 1 is π/n. With $\theta = \pi/n$, Figure 1 becomes the standard picture showing how the (ℓ, m, n)-group is generated.

THEOREM 2.3. *Let* A *and* B *be normalized elliptic matrices with distinct fixed points. The group generated by* A *and* B *is discrete if and only if the traces of* AB, A, *and* B *satisfy one of the seven conditions:*

(I) $|\text{trace AB}| < 2$ *and* AB *has extreme negative trace,*

(II) $|\text{trace AB}| \geq 2$,

(III) trace A = trace B *and* trace AB = $2\cos(\pi + 2\pi/n)$ *with* $n \geq 3$ *and odd,*

(IV) trace A = 0, trace B = $2\cos(\pi - \pi/n)$, *and* trace AB = $2\cos(\pi + 2\pi/n)$ *with* $n \geq 3$ *and odd (or the same thing with* A *and* B *interchanged),*

(V) trace A = $2\cos(\pi - \pi/3)$, trace B = $2\cos(\pi - \pi/n)$, *and*

$$\text{trace AB} = 2\cos(\pi + 3\pi/n)$$

with $n \geq 7$ *and not divisible by* 3 *(or the same thing with* A *and* B *interchanged),*

(VI) trace A = trace B = $2\cos(\pi - \pi/n)$ *and* trace AB = $2\cos(\pi + 4\pi/n)$, *with* $n \geq 7$ *and odd,*

(VII) trace A = $2\cos(\pi - \pi/3)$, trace B = $2\cos(\pi - \pi/7)$, *and*

$$\text{trace AB} = 2\cos(\pi + 2\pi/7)$$

(or the same thing with A *and* B *interchanged).*

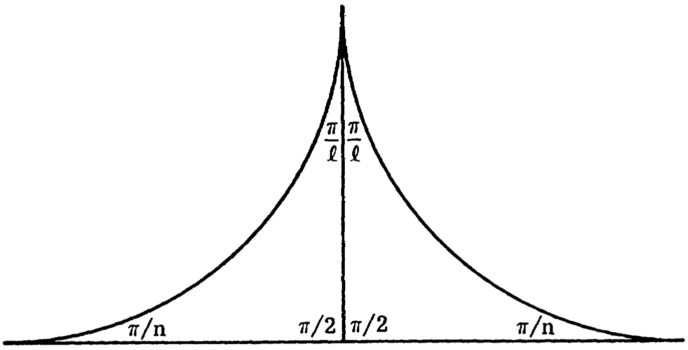
Remarks. In Theorem 3B of [3], Greenberg lists all triangle groups that are maximally contained in larger triangle groups. The three entries in his list correspond to our Cases III, IV, and V, but with each divisibility condition replaced by its denial: he requires n to be even in Cases III and IV and to be divisible by 3 in Case V. This correspondence is no accident, and the reason for it can be deduced from Figure 4. Case VI corresponds to no entry in Greenberg's list, because it is built out of Cases III and IV together.

3. PROOF OF THEOREM 2.3

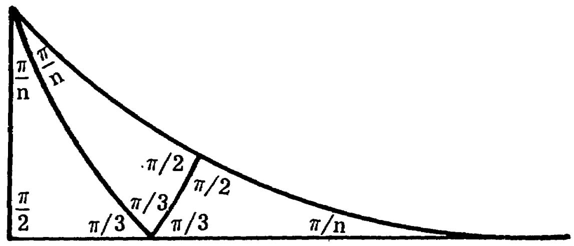
Some of the details of the proof of the theorem of the preceding section are repetitious, and we shall omit many of them, since our main interest lies in the geometric aspects of the proof.

We begin with the proof that Cases I to VII give discrete groups. In Case I, the group is discrete by Proposition 2.2. For Case II, once Figures 2 and 3 are verified, we again have standard pictures associated with two elementary classes of Fuchsian groups, and the group is discrete (by Poincaré's Theorem, for instance, see [4, pp. 221-227]).

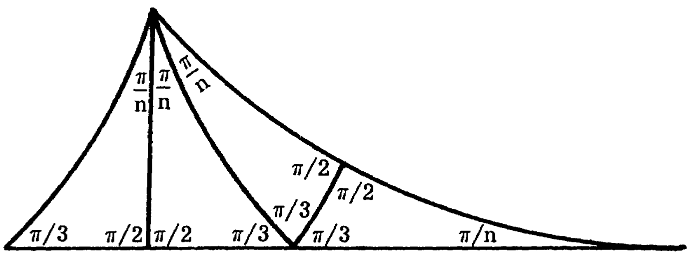
The remaining cases are intimately connected with special geometric configurations. In each case, the method of proving discreteness is suggested by a hyperbolic triangle with angles π/ℓ, π/m, and $s\pi/n$ and with the property that it can be partitioned into congruent subtriangles each of whose angles is π times the reciprocal of some integer (henceforth, we refer to such angles as *submultiples of* π). These configurations are given in Figure 4. Their relevancy will be apparent from the converse half of the proof.



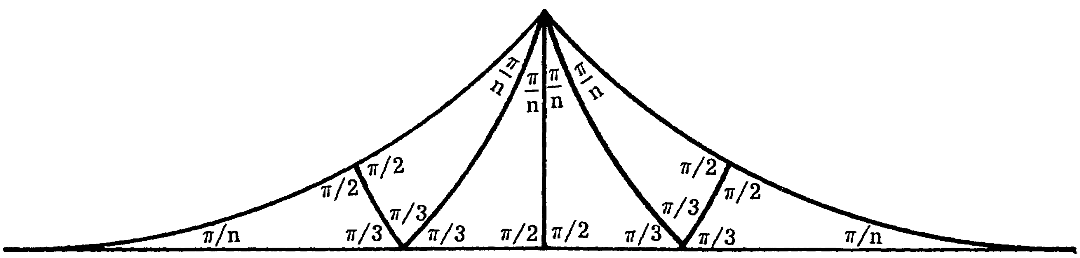
Case III



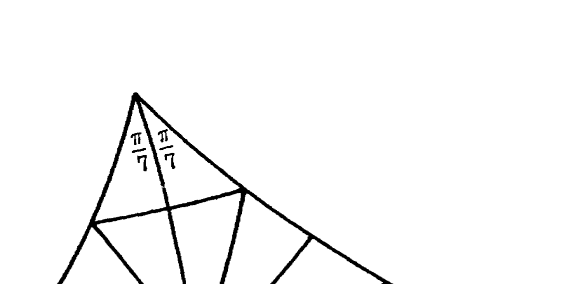
IV



V



VI



VII

Figure 4

In Case III, let trace A = trace B = $2\cos(\pi - \pi/\ell)$ and trace AB = $2\cos(\pi + 2\pi/n)$ (n odd). Put

$$(4) \qquad k = \begin{cases} \frac{1}{2}(n+1) & \text{if } n \equiv 3 \pmod 4, \\ -\frac{1}{2}(n-1) & \text{if } n \equiv 1 \pmod 4. \end{cases}$$

The group generated by $\sigma(A)$ and $\sigma(B)$ is the same as the group generated by $\sigma(A)$ and $\sigma A^{-1}(AB)^k$. We claim that A and $A^{-1}(AB)^k$ are normalized elliptic and that their product $(AB)^k$ has extreme negative trace. (Trace$[A^{-1}(AB)^k]$ will be 0.) If we can prove the claim, then discreteness follows from Case I. We shall have shown the image under σ of the group in question is a $(2, \ell, n)$-triangle group.

Choose a matrix C such that $\sigma(C)$ maps $FP(\sigma AB)$ into i and maps the direction from $FP(\sigma AB)$ toward i into the upper imaginary axis. Put $A' = C(AB)^{-1}C^{-1}$ and $B' = CAC^{-1}$. Since the upper right entry of AB is negative, since A' has its fixed point at i, and since B' has its fixed point on the upper imaginary axis, A' and B' have the forms

$$A' = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \qquad (\phi = \pi + 2\pi/n),$$

$$B' = \begin{pmatrix} \cos(\pi - \pi/\ell) & b' \\ c' & \cos(\pi - \pi/\ell) \end{pmatrix} \qquad (c' < 0,\ |b'| > |c'|,\ b'c' = -\sin^2(\pi - \pi/\ell)).$$

Since A'B' is conjugate to B^{-1}, it has the trace of B. That is,

$$(5) \qquad 2\cos(\pi - \pi/\ell) = 2\cos(\pi - \pi/\ell)\cos\phi - (c' - b')\sin\phi.$$

The trace of $A^{-1}(AB)^k$ is the trace of the inverse $(AB)^{-k}A$, and the sign of its upper right entry is the sign of the lower left entry of $(AB)^{-k}A$. It suffices to consider the conjugate matrix $(A')^k B'$. The trace is

$$(6) \qquad \text{trace}[(A')^k B'] = 2\cos(\pi - \pi/\ell)\cos k\phi - (c' - b')\sin k\phi.$$

Since $k\phi \equiv \pi + \pi/n \pmod{2\pi}$ for any odd n, we obtain the relations

$$(7) \qquad \cos k\phi = -\sin(\phi/2) \quad \text{and} \quad \sin k\phi = \cos(\phi/2).$$

Multiply both sides of (6) by $2\cos k\phi$, substitute from (7), and add the result to (5). The result is $-2\,\text{trace}[(A')^k B']\sin(\phi/2) = 0$. Since $n \geq 3$, $\sin(\phi/2) \neq 0$ and $\text{trace}(A')^k B' = 0$. The lower left entry of $(A')^k B'$ is $c'\cos k\phi + \cos(\pi - \pi/\ell)\sin k\phi$, and both terms are positive. Hence $A^{-1}(AB)^k$ is normalized. Finally, the trace of $(AB)^k$ is the trace of $(A')^k$, namely $2\cos k\phi$, which by (7) equals $2\cos(\pi + \pi/n)$. That is, $(AB)^k$ has extreme negative trace.

In Case IV, let us say

$$\text{trace } A = 0, \qquad \text{trace } B = 2\cos(\pi - \pi/n), \qquad \text{trace } AB = 2\cos(\pi + 2\pi/n).$$

Define k as in equation (4). Again, $\sigma(A)$ and $\sigma A^{-1}(AB)^k$ generate the same group as $\sigma(A)$ and $\sigma(B)$. A calculation similar to the one in Case III shows that A and $A^{-1}(AB)^k$ are normalized (with $\text{trace}[A^{-1}(AB)^k] = -1$) and that their product has extreme negative trace $2\cos(\pi + \pi/n)$. By Case I, the given group is discrete, and in fact, its image under σ is a (2, 3, n)-triangle group.

In Case V, let us say

$$\text{trace } A = 2\cos(\pi - \pi/3), \qquad \text{trace } B = 2\cos(\pi - \pi/n), \qquad \text{trace } AB = 2\cos(\pi + 3\pi/n).$$

Define

$$k = \begin{cases} n + \nu & \text{if } n = 3\nu - 1, \\ n - \nu & \text{if } n = 3\nu + 1. \end{cases}$$

Again, $\sigma(A)$ and $\sigma A^{-1}(AB)^k$ generate the same group as $\sigma(A)$ and $\sigma(B)$, and calculations show that A and $A^{-1}(AB)^k$ are normalized (with $\text{trace}[A^{-1}(AB)^k] = 0$) and that their product has extreme negative trace $2\cos(\pi + \pi/n)$. By Case I, the given group is discrete, and in fact its image under σ is a (2, 3, n)-group.

In Case VI, trace A = trace B = $2\cos(\pi - \pi/n)$ and trace AB = $2\cos(\pi + 4\pi/n)$. Put

$$k = \begin{cases} -\frac{1}{2}(n - 1) & \text{if } n \equiv 3 \pmod 4, \\ \frac{1}{2}(n + 1) & \text{if } n \equiv 1 \pmod 4. \end{cases}$$

Again $\sigma(A)$ and $\sigma A^{-1}(AB)^k$ generate the whole group, and again the argument in Case III shows that A and $A^{-1}(AB)^k$ are normalized (with $\text{trace}[A^{-1}(AB)^k] = 0$). Their product $(AB)^k$ has trace $2\cos(\pi + 2\pi/n)$, and n is odd. By Case IV, the given group is discrete, and in fact its image under σ is a (2, 3, n)-triangle group.

Finally, in Case VII let us say that

$$\text{trace } A = 2\cos(\pi - \pi/3), \qquad \text{trace } B = 2\cos(\pi - \pi/7), \qquad \text{trace } AB = 2\cos(\pi + 2\pi/7).$$

Here $\sigma(AB)^{-4}$ and $\sigma(A)$ generate the whole group, since $\sigma(AB)$ has order 7. Also, $(AB)^{-4}$ and A are normalized with $\text{trace}(AB)^{-4} = 2\cos(\pi - \pi/7)$, and $(AB)^{-4}A$ has trace $2\cos(\pi + 3\pi/7)$. The latter statement boils down to the identity

$$\cos \pi/7 + \cos 3\pi/7 + \cos 5\pi/7 = 1/2,$$

an assertion that follows from the fact that the sum of all the seventh roots of -1 is 0. Discreteness now follows from Case V, and the image under σ of the group in question is a (2, 3, 7)-group.

We turn to the proof that Cases I to VII are necessary for discreteness. Let A and B denote normalized elliptic matrices of the form (1), and let D be the subgroup of G that they generate. We suppose that D is discrete. By Case II, we may assume that AB is an elliptic matrix, and we may clearly suppose that AB has finite order. Let the trace of AB be $2\cos(\pi + s\pi/n)$, with $0 < s < n$ and with s and n relatively prime. By Case I, we may assume $s > 1$. Figure 1 applies, and the angles θ in that figure are $s\pi/n$. The proofs of the next two lemmas were supplied by B. Maskit.

LEMMA 3.1. $\sigma(D)$ *is a triangle group (with elliptic generators).*

Proof. By Theorem 1.1, $U/\sigma(D)$ is a compact Riemann surface, say of type (g, n), and so $\sigma(D)$ is isomorphic to one of the standard Fuchsian groups. If $(g, n) \neq (0, 3)$, there are uncountably many nonconjugate groups in G^* isomorphic to $\sigma(D)$. (See [1, p. 356] for the case $g > 1$, for example.) This conclusion contradicts Proposition 2.2. The proof is complete.

In the remainder of this section, the integers p, q, and r are such that $\sigma(D)$ is a (p, q, r)-group with $p \leq q \leq r$.

LEMMA 3.2. *The integers* ℓ, m, *and* n *each divide one of* p, q, *and* r.

Proof. Let the standard generators of the (p, q, r)-group $\sigma(D)$ be α and β. It is enough to prove that each element of $\sigma(D)$ of finite order is conjugate to an integral power of α, β, or $\alpha\beta$. Take an element of finite order greater than 1, and conjugate it by a member of $\sigma(D)$ so that its fixed point x is in the standard fundamental polygon for $\sigma(D)$. Then x is not an interior point, because interior points (especially those near x) are inequivalent. It is not an edge point other than a vertex for the same reason (or, if the element has order 2, because points near x on the edge are inequivalent). Therefore x is a vertex, and the conjugated element is a power of α, β, $\alpha\beta$, or $\beta\alpha$. That power must be a rational power, and the denominator of the exponent must be 1. In fact, otherwise $\sigma(D)$ would possess an element, a nontrivial integral power of which is α, β, $\alpha\beta$, or $\beta\alpha$, contrary to the fact that points in a fundamental polygon near a vertex are all inequivalent. Since $\alpha\beta$ and $\beta\alpha$ are conjugate, the lemma is proved.

Now apply Theorem 1.1 to the polygon in Figure 1. By part (i), the translates by $\sigma(D)$ of this polygon cover most points of the upper half plane t times, where t is independent of the point. We shall call t the *covering number* of the polygon. Most points in a neighborhood of $FP(\sigma AB)$ are covered s times just by the powers of $\sigma(AB)$. Hence $t \geq s$. The area of the polygon in Figure 1 and the area of a standard fundamental region for $\sigma(D)$ are both computable, and part (ii) of Theorem 1.1 thus gives us the important relation

$$1 - \left(\frac{1}{\ell} + \frac{1}{m} + \frac{s}{n}\right) = t\left[1 - \left(\frac{1}{p} + \frac{1}{q} + \frac{1}{r}\right)\right]. \tag{8}$$

We shall describe a partitioning of the polygon in Figure 1 that will give one further condition on D. By the remarks preceding Proposition 2.2, $\sigma(AB)$ rotates directions at $FP(\sigma AB)$ counterclockwise through $2s\pi/n$. Since s and n are relatively prime, we can choose k so that $sk \equiv 1 \pmod{n}$. Then σ of $C = (AB)^k$ rotates directions at $FP(\sigma AB)$ counterclockwise through $2\pi/n$. Form the powers C^j with $0 \leq j \leq s$, and then fix j. In Figure 1, draw the geodesic from $FP(\sigma AB)$ at an angle of $j\pi/n$ measured clockwise from side 2. This geodesic meets the imaginary axis at a point x between $FP(\sigma A)$ and $FP(\sigma B)$ (strictly between them, if $0 < j < s$). Let the arc from $FP(\sigma AB)$ to x be called side 5, and let the reflection of side 5 in the imaginary axis be called side 6. The quadrilateral 1234 has thereby been partitioned into quadrilaterals 1256 and 6534 (one of them is degenerate if $j = 0$ or $j = s$). We claim that $\sigma A^{-1}C^j(\text{side } 5) = \text{side } 6$. In fact, sides 5 and 6 have the same length, and hence so do $\sigma C^j(\text{side } 5)$ and $\sigma A(\text{side } 6)$. If we show that $\sigma C^j(\text{side } 5)$ and $\sigma A(\text{side } 6)$ make the same oriented angle with $\sigma A(\text{side } 1) = \text{side } 2$, then we shall have the result $\sigma C^j(\text{side } 5) = \sigma A(\text{side } 6)$, that is,

$$\sigma A^{-1} C^j(\text{side } 5) = \text{side } 6\,.$$

Thus we need only observe that

$$\text{angle}(\sigma A(\text{side } 1), \sigma A(\text{side } 6)) = \text{angle}(\text{side } 1, \text{side } 6) = \text{angle}(\text{side } 5, \text{side } 2)$$
$$= \text{angle}(\text{side } 2, \sigma C^j(\text{side } 5)).$$

Let $0 \leq j \leq s - 1$, and consider the above construction for j and $j + 1$. In an obvious notation, quadrilateral $6_j\, 5_j\, 5_{j+1}\, 6_{j+1}$ has its sides matched in pairs by members of $\sigma(D)$. Specifically,

$$\sigma A^{-1} C^j (\text{side } 5_j) = \text{side } 6_j \quad \text{and} \quad \sigma A^{-1} C^{j+1} (\text{side } 5_{j+1}) = \text{side } 6_{j+1}.$$

Also, $\sigma A^{-1} C^j$ and $\sigma A^{-1} C^{j+1}$ together generate $\sigma(D)$, and so by Theorem 1.1 quadrilateral $6_j\, 5_j\, 5_{j+1}\, 6_{j+1}$ receives a covering number relative to $\sigma(D)$. As j varies, we get a partition of quadrilateral 1234 into s pieces. Each of the pieces has a covering number, and the sum of these covering numbers is t, almost by definition.

Each of the s pieces is composed of a triangle and its reflection. If one of these triangles has angle ϕ at a vertex y, then an elliptic transformation that fixes y and rotates directions at y through 2ϕ is a member of $\sigma(D)$. In particular, one of the s quadrilaterals has covering number 1 if and only if the angles of its component triangles (the two triangles are congruent) are all submultiples of π. (The "if"-part of the statement follows from Poincaré's Theorem, since $\sigma A^{-1} C^j$ and $\sigma A^{-1} C^{j+1}$ generate all of $\sigma(D)$.)

At this point, we have developed all the tools we need, and the proof will be completed in four steps.

Step 1. If $2 \leq t \leq 6$, the only possible values for t and s are (2, 2), (3, 2), (4, 3), and (6, 4), and these possibilities lead only to Cases III, IV, V, and VI, respectively. For the proof we observe that $2 \leq s \leq t \leq 6$. The individual cases are somewhat alike, and as samples we shall treat only the cases $(t, s) = (2, 2), (3, 2), (6, 4), (6, 5)$, and (6, 6).

If $(t, s) = (2, 2)$, the quadrilateral 1234 divides into two pieces, and each is forced to have covering number 1. The angles of the component triangles must be submultiples of π, and in particular the four angles at x on the imaginary axis must each be $\pi/2$. Since the areas of the component quadrilaterals must be equal when the covering numbers are equal, we must have $\pi/\ell = \pi/m$. That is, ℓ must equal m, and n (being prime to $s = 2$) must be odd. This is Case III.

If $(t, s) = (3, 2)$, the quadrilateral 1234 divides into two pieces, one with covering number 1 and the other with covering number 2. Consider the former quadrilateral. The angle at x of a component triangle is a submultiple of π. Thus the angle at x of a component triangle of the other quadrilateral is at least $\pi/2$. But that angle must also be a submultiple of 2π, since the second quadrilateral has covering number at most 2, and that angle cannot be $\pi/2$ since otherwise the second quadrilateral would have covering number 1. Hence the angle is $2\pi/3$. The other angles of a component triangle of the second quadrilateral are submultiples of π, and the case $(t, u) = (2, 2)$ implies that these angles are equal. Bisecting the second quadrilateral as in the (2, 2)-case, we find that a component triangle of one of the pieces has angles $\pi/2$, $\pi/3$, and π/n. A component triangle of the first quadrilateral has angles $\pi/3$, π/n, and π/ℓ or π/m. Hence either ℓ or m equals 2, and the other is n. The integer n is odd, since it is prime to $s = 2$. This is Case IV.

If $t = 6$ and $s = 4$, 5, or 6, we appeal to equation (8). Lemma 3.2 shows that ℓ, m, and n are not greater than r. Since $s \geq 4$, the left side is increased if we replace it with $1 - 6/r$. We see that

$$1 - \frac{6}{r} \geq 6 - \frac{6}{r} - 6\left(\frac{1}{p} + \frac{1}{q}\right) \tag{9}$$

with equality only if $\ell = m = n = r$ and $s = 4$. Inequality (9) says that $p^{-1} + q^{-1} \geq 5/6$, and this is possible only if $p = 2$ and $q = 3$. We are forced into Case VI.

The other possibilities for (t, s) with $t \leq 6$ are handled by arguments similar to those for (2, 2) and (3, 2), but some of the other cases are a little more involved. We omit the proofs.

Step 2. If $t > 6$, then $p = 2$ and $q = 3$. In fact, direct calculation shows that only if $p = 2$ and $q = 3$ is it possible to satisfy the condition imposed by Lemma 3.2, the requirement $s > 1$, and the inequalities

$$1 - \left(\frac{1}{\ell} + \frac{1}{m} + \frac{s}{n}\right) \geq 7\left[1 - \left(\frac{1}{p} + \frac{1}{q} + \frac{1}{r}\right)\right] > 0 \tag{10}$$

implied by equation (8).

Step 3. If $t > 6$, the only possibilities for $(t, s; \ell, m, n; p, q, r)$ are

(VIIa) (10, 2; 9, 9, 9; 2, 3, 9),

(VIIb) (18, 2; 7, 7, 7; 2, 3, 7),

(VIIc) (10, 2; 3, 7, 7; 2, 3, 7),

(VIId) (9, 3; 8, 8, 8; 2, 3, 8),

(VIIe) (12, 3; 7, 7, 7; 2, 3, 7).

For the proof, we recall that $p = 2$ and $q = 3$. Inequality (10) and the inequalities $\ell, m, n \leq r$ and $r \geq 7$ lead to the inequality $6(5 - s) \geq 7$. Hence $s = 2$ or $s = 3$.

Suppose $s = 2$. If $n \leq r/2$ or if neither ℓ nor m is greater than $r/2$, inequality (10) leads to the contradiction $r \leq 6$. Thus by Lemma 3.2 we may assume that $m = n = r$. Equation (8) becomes

$$\frac{t - 3}{n} = \frac{1}{6}(t - 6) + \frac{1}{\ell}. \tag{11}$$

If $\ell = n$, the only solutions (n, t) to (11) with $t > 6$ and n odd (recall that s is prime to n) are (7, 18) and (9, 10), and these give VIIb and VIIa. When $\ell \leq n/2$, equation (11) gives $n \leq 12$, and so $n = 7$, 9, or 11. Case-by-case inspection for these values of n gives VIIc as the only solution to (11) with $t > 6$ and $\ell \leq n/2$.

Suppose $s = 3$. Unless $\ell = m = n = r$, inequality (10) and Lemma 3.2 lead to the contradiction $r \leq 6$. With $\ell = m = n = r$, equation (8) becomes $n = 6(t - 5)/(t - 6)$, and the only solutions (n, t) with $n \geq 7$, $t > 6$, and n prime to $s = 3$ are those in VIId and VIIe.

Step 4. In Step 3, only VIIc can occur, and it is Case VII of Theorem 2.3. In fact, the values of s, ℓ, m, and n in VIIa and VIIb are covered by Case III; they lead to discrete groups, but t is 2 for them. Clearly VIIc is Case VII of the theorem.

Consider VIId. As in Step 1, the quadrilateral 1234 divides into three parts. The covering numbers of the edge quadrilaterals must be 1, 2, 3, 4, or 6, because in each case two of the angles of a component triangle are submultiples of π (namely, $\pi/8$). The covering number 1 is impossible, because two of p, q, and r would have to be 8; covering numbers 3 and 6 are impossible because 8 is even; and covering number 4 is impossible because the angles of the triangle would have to be $\pi/3$, $\pi/8$, and $3\pi/8$. Therefore the covering numbers for the three pieces are 2, 5, 2. The two edge triangles are then congruent, and since the sum of the angles in the middle triangle is less than π, the angles are $\pi/3$, $\pi/3$, and $\pi/8$. Since these are all submultiples of π, the middle quadrilateral has covering number 1, which is a contradiction.

Consider VIIe. Again quadrilateral 1234 divides into three parts. Covering numbers 1 and 4 are impossible for the edge quadrilaterals, just as in VIId, and covering number 3 is impossible because the angles of a component triangle would have to be $\pi/2$, $\pi/7$, and $2\pi/7$ (whereas two of them are $\pi/7$). Thus the only possibilities for the covering numbers are 2, 8, 2 and 2, 4, 6. Possibility 2, 8, 2 does not occur for the same reason as with 2, 5, 2 in VIId. Possibility 2, 4, 6 does not occur, because elementary geometry shows that the triangle associated with covering number 2 and the one with covering number 6 are congruent, in contradiction to part (ii) of Theorem 1.1.

4. GROUPS WITH A PARABOLIC GENERATOR

If one generator of a doubly generated subgroup of G is parabolic and the other is elliptic or parabolic, the question of discreteness is more easily settled than when both generators are elliptic. In this section, we state the results and mention what tools are used, but we only sketch the proofs.

We may assume that the parabolic matrix generators in question have trace -2, because the replacement of a parabolic generator by its negative leads to the same subgroup of G*, and hence discreteness is not affected. For a matrix in G of trace -2, the signs of the upper right and lower left entries do not change when the matrix is conjugated by a member of G. (It is enough to verify this assertion for the matrix $(-1, \pm 1 \mid 0, -1)$.) Also, the signs are opposite, except that one of the entries may be 0. We shall say a parabolic matrix is *normalized* if it has trace -2, its upper right entry is nonnegative, and its lower left entry is nonpositive.

In considering discreteness for the elliptic-parabolic case, we may certainly assume that the elliptic generator has finite order. We may then arrange, without loss of generality, that both generators are normalized.

PROPOSITION 4.1. *If* A *is a normalized elliptic matrix and* B *is a normalized parabolic matrix, then the subgroup of* G *that they generate is discrete if and only if* AB *is hyperbolic, or is parabolic, or is elliptic with extreme negative trace.*

For the proof, conjugate the group generated by A and B so that the fixed point of $\sigma(A)$ is at i and the fixed point of $\sigma(B)$ is at ∞. The picture is then the same as in Figure 1, 2, or 3, except that $FP(\sigma B)$ is ∞ and the arcs leading to it are vertical. Moreover, trace AB < 2, and if trace AB ≤ -2, then Figure 2 or Figure 3 applies, and the same argument as in the case where both generators are elliptic gives discreteness. If AB is elliptic and θ in the modified Figure 1 is a submultiple of π, the group is discrete. If θ is a rational multiple of π but not a submultiple, then we can generate the group by A and a power $(AB)^k$ such that $\sigma(AB)^k$ rotates directions at $FP(\sigma AB)$ through a smaller angle than $\sigma(AB)$. The result is that translates of a compact portion of the polygon in the figure cover all of U. Since U/D cannot be compact if D is a discrete subgroup of G* containing parabolic elements, the group in question is not discrete.

For the case where both generators are parabolic, we may assume that both generators are normalized. The case where their fixed points are the same is simple, and we assume that the fixed points are distinct. They we can conjugate the normalized generators so that they are of the form

$$A = \begin{pmatrix} -1 & 0 \\ -1 & -1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} \text{ with } b > 0. \tag{12}$$

The trace of AB is 2 - b.

PROPOSITION 4.2. *If* A *and* B *are normalized parabolic matrices of the form* (12), *then the subgroup of* G *that they generate is discrete if and only if* $b \geq 4$ *or* $2 - b = 2\cos(\pi + 2\pi/n)$ *for an integer* $n \geq 3$.

In other words, the exceptional case III of Theorem 2.3 occurs also in the case where both generators are parabolic, but there are no other exceptional cases. In the proof, trace $AB \geq 2$ cannot occur and trace $AB \leq -2$ is handled as in the preceding cases. If AB is elliptic and the group is discrete, the angle that corresponds to θ in Figure 1 must be $2\pi/n$, since otherwise we can give an argument similar to the one in the elliptic-parabolic case to show that U modulo σ of the group is compact. This means that $2 - b$ must be of the form $2\cos(\pi + 2\pi/n)$ if the group is discrete. On the other hand, arguments similar to those at the beginning of Section 3 show that the group is discrete if $2 - b$ is $2\cos(\pi + 2\pi/n)$.

The referee points out that the situation in Proposition 4.2 was considered by Brenner [2], who proved that the group in question (call it D) is free if $b \geq 4$. Now if this group D is free, then $\sigma(D)$ is free, and Maskit's Theorem 4 in [5] shows the group is discrete. Hence Brenner's result implies Proposition 4.2 above if $b \geq 4$. Maskit's theorem also answers a question raised by Brenner: Is there any value of b with $0 < b < 4$ for which D is free? The answer is negative, because $\sigma(D)$ would have to be discrete, and the elliptic matrix AB would therefore have to be of finite order.

REFERENCES

1. L. Bers, *Spaces of Riemann surfaces*. Proc. Internat. Congress Math. 1958, 349-361, Cambridge Univ. Press, New York, 1960.

2. J. L. Brenner, *Quelques groupes libres de matrices*. C. R. Acad. Sci. Paris 241 (1955), 1689-1691.

3. L. Greenberg, *Maximal Fuchsian groups*. Bull. Amer. Math. Soc. 69 (1963), 569-573.

4. J. Lehner, *Discontinuous groups and automorphic functions*. Mathematical Surveys, No. 8, American Mathematical Society, Providence, R.I., 1964.

5. B. Maskit, *Construction of Kleinian groups*. Proc. Conf. Complex Analysis (Minneapolis, 1964), 281-296. Springer, New York, 1965.

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
and
Cornell University, Ithaca, New York 14850