Math. Proc. Camb. Phil. Soc. (1987), **102**, 25
Printed in Great Britain

Generalized triangle groups

BY GILBERT BAUMSLAG
Department of Mathematics, City College, New York, NY 10031, *U.S.A.*

JOHN W. MORGAN
Department of Mathematics, Columbia University, New York, NY 10027, *U.S.A.*

AND PETER B. SHALEN
Department of Mathematics, University of Illinois at Chicago, Chicago, IL 60680, *U.S.A.*

(*Received* 6 *August* 1986; *revised* 4 *November* 1986)

1. *Introduction*

A group G is called a *triangle group* if it can be presented in the form

$$G = \langle a, b; a^l = b^m = (ab)^n = 1 \rangle \quad (l, m, n > 1).$$

It is well-known that G is isomorphic to a subgroup of $PSL_2(\mathbb{C})$, that a is of order l, b is of order m and ab is of order n. If

$$s(G) = 1/l + 1/m + 1/n \tag{1}$$

then G contains the fundamental group of a positive genus orientable surface as a subgroup of finite index whenever $s(G) \leqslant 1$; in particular G is infinite. Furthermore, if $s(G) < 1$, the genus of the surface is greater than 1 and consequently G contains a free group of rank 2.

We are interested in the family of groups that can be presented in the form

$$G = \langle a, b; a^l = b^m = w^n = 1 \rangle \quad (l, m, n > 1),$$

where w is any element in the free group freely generated by a and b. Since w can be replaced by any one of its conjugates in such a presentation of G, we can always assume that w is a cyclically reduced word in a and b. If w is either a power of a or a power of b, G is a free product of two cyclic groups and so its structure is completely understood. Thus we concentrate on the cases where w is a more interesting word.

Here we term a group G a *generalized triangle group* if it can be presented in the form

$$G = \langle a, b; a^l = b^m = w^n = 1 \rangle \quad (l, m, n > 1) \tag{2}$$

where

$$w = a^{r_1} b^{s_1} \ldots a^{r_k} b^{s_k} \quad (k \geqslant 1, 0 < r_i < l, 0 < s_i < m). \tag{3}$$

Furthermore, we call a representation σ of G in a group X *special* if $\sigma(a)$, $\sigma(b)$, $\sigma(w)$ are of orders l, m and n respectively; if, in addition, $\sigma(G)$ is cyclic, σ will be called a *special cyclic representation* and if $\sigma(G)$ is a dihedral group, then σ will be called a *special dihedral representation*.

The following two theorems are analogues for generalized triangle groups of the results stated above for triangle groups.

https://doi.org/10.1017/S0305004100067013 Published online by Cambridge University Press

THEOREM A. *Let G be a generalized triangle group. Then there exists a special representation of G in $PSL_2(\mathbb{C})$. In particular, a, b and w have orders l, m and n respectively in G.*

It follows immediately from Theorem A that if G is a generalized triangle group then $G \neq 1$. This answers a question put to us by C. McA. Gordon (see Gordon and Luecke[5]). (For $n > 5$ this was established by González-Acuña and Short in [4], theorem 4·3.)

Let G be a generalized triangle group given by the presentation (2), (3), and let $s(G)$ be given by (1). The following theorem then holds.

THEOREM B. *If $s(G) \leqslant 1$, then G has a subgroup of finite index which maps onto Z; in particular, G is infinite. If $s(G) < 1$, then G contains a subgroup of finite index mapping onto a free group of rank 2; in particular, G contains a free subgroup of rank 2.*

The idea used in the proof of Theorem A is similar to one used by Ree and Mendelsohn in [10]. It makes crucial use of the fact that in $SL_2(\mathbb{C})$ a matrix A is of order $q > 2$ if and only if its trace is of the form

$$\operatorname{tr} A = \alpha + \alpha^{-1},$$

where α is a primitive qth root of unity.

The first part of Theorem B is proved using an estimate on the deficiency of a presentation (i.e. the difference between the number of generators and the number of relations) of a subgroup of finite index in a generalized triangle group. This estimate is derived by using Theorem A. The second part of Theorem B follows immediately from this estimate and a theorem of B. Baumslag and Pride[1].

There are two alternative criteria which can be used to show that a generalized triangle group is infinite. These are contained in

THEOREM C. *If the generalized triangle group G has a special cyclic representation, or if G has a special dihedral representation with at most one of l, m, n equal to 2, then G is infinite. In fact the kernel of the given special representation (is of finite index and) maps onto $Z \times Z$.*

Notice that if $l = m = 2$, then G is a finite dihedral group.

It is worth pointing out here that there are examples of generalized triangle groups which turn out to be infinite because of Theorem C but which do not satisfy the basic inequality of Theorem B. For example,

$$G = \langle a, b; a^2 = b^3 = ((ab)^4(ab^2))^2 = 1 \rangle$$

has a special cyclic representation and so is infinite. Similarly,

$$G = \langle a, b; a^2 = b^3 = (abab^2)^3 = 1 \rangle$$

has a special representation in the dihedral group of order 6 and so is also infinite, although again Theorem B does not apply.

After this work was completed Steven Boyer pointed out to us that he had independently, and at about the same time, proved that a generalized triangle group G is non-trivial by constructing a non-trivial representation of G in $SO(3)$ (Boyer[2]). Our methods do not yield such a representation of G. However we obtain a representation of G in $PSL_2(\mathbb{C})$ such that the elements a, b and w have orders l, m

https://doi.org/10.1017/S0305004100067013 Published online by Cambridge University Press

and n as specified by the defining relations of G. Boyer's work does not provide us with information about the orders of a, b and w.

2. *The proof of Theorem A*

Let F be the free group on a and b. We associate to each representation ρ of F in $SL_2(\mathbb{C})$ the point

$$(\operatorname{tr}\rho(a), \operatorname{tr}\rho(b), \operatorname{tr}\rho(ab)) \in \mathbb{C}^3.$$

The following two facts can be found, for example, in Horowitz [6]. Firstly, given any $(x, y, z) \in \mathbb{C}^3$ there exists a representation $\rho\colon F \to \mathbb{C}^3$ with $\operatorname{tr}\rho(a) = x$, $\operatorname{tr}\rho(b) = y$ and $\operatorname{tr}\rho(ab) = z$. Secondly, for each element $v \in F$ there is a polynomial $p_v(x, y, z)$ such that for any representation $\rho\colon F \to SL_2(\mathbb{C})$ we have

$$\operatorname{tr}\rho(v) = p_v(\operatorname{tr}\rho(a), \operatorname{tr}\rho(b), \operatorname{tr}\rho(ab)).$$

(From the point of view of geometric invariant theory $\mathbb{C}^3$ is the variety of characters of representations of F into $SL_2(\mathbb{C})$; cf. Culler and Shalen [3], Lubotzky and Magid [8].)

The following formula was pointed out to us by T. Jorgenson.

LEMMA 1. *Let w be the element of F given by* (3) *and let*

$$\lambda = \alpha + \alpha^{-1}, \quad \mu = \beta + \beta^{-1}$$

be complex numbers with

$$\alpha^2 \neq 1 \quad \text{and} \quad \beta^2 \neq 1.$$

Then

$$q_w(z) = p_w(\lambda, \mu, z)$$

is a polynomial of degree at most k. The coefficient of z^k is given by the formula

$$\prod_{i=1}^{k} \left(\frac{\alpha^{r_i} - \alpha^{-r_i}}{\alpha - \alpha^{-1}}\right)\left(\frac{\beta^{s_i} - \beta^{-s_i}}{\beta - \beta^{-1}}\right). \tag{4}$$

The proof of Lemma 1 is based on the following identities:

$$\operatorname{tr}(xy) + \operatorname{tr}(xy^{-1}) = \operatorname{tr}(x)\operatorname{tr}(y) = \operatorname{tr}(xy) + \operatorname{tr}(x^{-1}y) \quad (x, y \in SL_2(\mathbb{C})). \tag{5}$$

When $k = 1$, $w = a^{r_1}b^{s_1}$ and the result follows from (5) by induction on $r_1 + s_1$. (If $r_1 > 1$, we write $w = a(a^{r_1-1}b^{s_1})$; if $s_1 > 1$, we write $w = (a^{r_1}b^{s_1-1})b$.) For $k > 1$, the result again follows from (5) by induction on k. (We write $w = (a^{r_1}\ldots b^{s_{k-1}})(a^{r_k}b^{s_k})$.)

Theorem A can be readily deduced from the above remarks and Lemma 1. To this end, let G be a generalized triangle group defined by (2) and (3). Choose α and β to be primitive $2l$th and $2m$th roots of unity respectively and let

$$\lambda = \alpha + \alpha^{-1}, \quad \mu = \beta + \beta^{-1}.$$

Then, by Lemma 1, q_w is a polynomial in z of degree at most k with the coefficient of z^k given by (4). Since l and m are both greater than 1 we have $\alpha^2 \neq 1$ and $\beta^2 \neq 1$. So it follows from the conditions on the r_i and s_i and (4) that the coefficient of z^k is non-zero. Since $k \geqslant 1$, q_w is a polynomial of positive degree over $\mathbb{C}$, and so it takes on all values. In particular, if ζ is a primitive $2n$th root of unity there is a complex number ν with

$$q_w(\nu) = p_w(\lambda, \mu, \nu) = \zeta + \zeta^{-1}.$$

https://doi.org/10.1017/S0305004100067013 Published online by Cambridge University Press

Translating this information back into the language of representations, it follows that there is a representation ρ of F into $SL_2(\mathbb{C})$ such that

$$\operatorname{tr}\rho(a) = \lambda, \operatorname{tr}\rho(b) = \mu, \operatorname{tr}\rho(ab) = \nu \quad \text{and} \quad \operatorname{tr}\rho(w) = \zeta + \zeta^{-1}.$$

Since l, m and n are greater than 1, $\rho(a)$, $\rho(b)$, $\rho(w)$ are of orders $2l$, $2m$, $2n$ respectively. Now consider the sequence

$$F \xrightarrow{\rho} SL_2(\mathbb{C}) \xrightarrow{\pi} PSL_2(\mathbb{C}),$$

where π is the canonical homomorphism from $SL_2(\mathbb{C})$ onto $PSL_2(\mathbb{C})$. Since there is unique element of order 2 in $SL_2(\mathbb{C})$ and it is in the kernel of π, the orders of $\pi(\rho(a))$, $\pi(\rho(b))$ and $\pi(\rho(ab))$ are l, m and n respectively. Thus $\pi\rho$ factors yielding the required special representation of G into $PSL_2(\mathbb{C})$, which completes the proof of Theorem A.

3. *The proof of Theorem B*

If X is a group and S is a subset of X, then we denote the subgroup generated by S by $\operatorname{gp}(S)$ and the normal closure of S in X by $\operatorname{gp}_X(S)$. The proof of Theorem B depends on the corollary of the lemma that follows.

LEMMA 2. *Let X be a group, Y a normal subgroup of X of finite index j in X. Furthermore, let w be an element of X which is of order n modulo Y. Then*

$$\operatorname{gp}_X(w^n) = \operatorname{gp}_Y(s_i w^n s_1^{-1} \quad (i = 1, \ldots, j/n)),$$

where $s_1, \ldots, s_{j/n}$ is a set of representatives of the left cosets of $\operatorname{gp}(Y, w)$ in X.

Proof. Since $\{s_i w^k \mid i = 1, \ldots, j/n, k = 0, \ldots, n-1\}$ is a complete set of representatives of the right cosets of Y in X we have

$$\begin{aligned}\operatorname{gp}_X(w^n) &= \operatorname{gp}_Y((s_i w^k) w^n (s_i w^k)^{-1} \quad (i = 1, \ldots, j/n, k = 0, \ldots, n-1)) \\ &= \operatorname{gp}_Y(s_i w^n s_i^{-1} \quad (i = 1, \ldots, j/n)).\end{aligned}$$

Retaining the hypotheses and notation of Lemma 2, it is easy to deduce the following:

COROLLARY 3. *If Y can be presented in the form*

$$Y = \langle u_1, \ldots, u_p; r_1, \ldots, r_q \rangle,$$

then $$Y/\operatorname{gp}_X(w^n) = \langle u_1, \ldots, u_p; r_1, \ldots, r_q, \quad \{s_i w^n s_i^{-1} \mid i = 1, \ldots, j/n\} \rangle,$$

where each of $s_i w^n s_i^{-1}$ is here assumed to be written as a word in the given generators $u_1, \ldots, u_p$ of Y.

We are now in a position to begin the proof of Theorem B. We first concoct a special representation of the generalized triangle group G onto a finite group. To do this, note that by Theorem A, there is a special representation σ of G in $PSL_2(\mathbb{C})$. Let H be the inverse image of $\sigma(G)$ in $SL_2(\mathbb{C})$ and let c be the generator of the centre of $SL_2(\mathbb{C})$. It follows from a theorem of Mal'cev [9] that H is residually finite. Hence there is a normal subgroup K of finite index in H such that c does not belong to K. It follows that there are elements a', b' and w' in H covering $\sigma(a)$, $\sigma(b)$ and $\sigma(w)$, whose orders

https://doi.org/10.1017/S0305004100067013 Published online by Cambridge University Press

modulo K are $2l$, $2m$ and $2n$ respectively. Clearly the composition of the maps in the sequence

$$G \to \sigma(G) \cong H/\mathrm{gp}\,(c) \to H/\mathrm{gp}\,(K, c)$$

is a special representation ϕ, say, of G onto a finite group.

Now consider the group

$$X = \langle a, b;\ a^l = b^m = 1 \rangle.$$

Thus X is a free product of two finite cyclic groups. There is a canonical epimorphism $\epsilon\colon X \to G$. Consider the sequence

$$X \overset{\epsilon}{\to} G \overset{\phi}{\to} H/\mathrm{gp}\,(K, c).$$

Let $Y = \mathrm{Ker}\,(\phi\epsilon)$. Then Y is a normal subgroup of finite index, say j, in X and

$$Y \cap \mathrm{gp}\,(a) = 1 = Y \cap \mathrm{gp}\,(b).$$

Hence, by the Kurosh subgroup theorem [7], Y is a free group of finite rank r. Now Wall[11] has defined for each group W in a certain class of groups a rational number $\chi(W)$ called its Euler characteristic. This class includes, in particular, the infinite cyclic group and all finite groups and it is closed under free products with finitely many factors and subgroups of finite index. Moreover if W is a finite group, then $\chi(W) = 1/|W|$, if $W = A * B$ is a free product, then $\chi(W) = \chi(A) + \chi(B) - 1$, if W is a free group of finite rank r, then $\chi(W) = 1 - r$ and if V is a subgroup of finite index j in W then $\chi(V) = j\chi(W)$ (these facts are due to Wall[11]). Thus it follows that

$$\chi(X) = 1/l + 1/m - 1.$$

On the other hand we can compute the Euler characteristic of X in terms of that of Y:

$$\chi(X) = \chi(Y)/j = (1 - r)/j.$$

Hence
$$1/l + 1/m - 1 = (1 - r)/j.$$

Consequently
$$r = 1 - j(1/l + 1/m - 1). \tag{6}$$

Now our generalized triangle group G is obtained by adding the relation $w^n = 1$ to X:

$$G \cong X/\mathrm{gp}_X(w^n).$$

Since $\mathrm{gp}_X(w^n) \subset Y$, the factor group $Y/\mathrm{gp}_X(w^n)$ can be viewed as a subgroup of G of finite index. As already noted, Y is free of rank r given by (5). Thus, by Corollary 3, $Y/\mathrm{gp}_X(w^n)$ can be defined on r generators subject to j/n relations. So the deficiency d of this presentation is given by

$$d = r - j/n = 1 - j(1/l + 1/m + 1/n - 1) = 1 - j(s(G) - 1).$$

If $s(G) \leqslant 1$ then $d \geqslant 1$ which means that $Y/\mathrm{gp}_X(w^n)$ maps onto the infinite cyclic group. Moreover $d \geqslant 2$ if $s(G) < 1$, which means that H contains a subgroup of finite index mapping onto a free group of rank two by a theorem of B. Baumslag and Pride[1]. This then completes the proof of Theorem B.

https://doi.org/10.1017/S0305004100067013 Published online by Cambridge University Press

4. *The proof of Theorem C*

Let H be a group given by the presentation

$$H = \langle a, b; a^s = w_1^t = w_2^u = \ldots = 1 \rangle,$$

where s, t, u and so on are all at least 2. A *special representation* of H is a representation in which a has order s, w_1 has order t, w_2 has order u and so on.

Let $\tilde{L} \subset SL_2(\mathbb{C})$ be the group of lower triangular matrices and let $U \subset \tilde{L}$ be the unipotent subgroup. Let V and L be the projections of U and $\tilde{L}$ respectively in $PSL_2(\mathbb{C})$. Then $V \cong \mathbb{C}$ and $L/V \cong \mathbb{C}^*/\{\pm\}$.

The proof of Theorem C hinges on the following:

LEMMA 4. *Suppose that H has a special representation σ into a finite cyclic group such that $\sigma(b)$ has order greater than 2. Then H has a special representation ρ into L with $\rho(H) \cap V$ non-cyclic. Moreover $\rho^{-1}(V)$ is the kernel of σ.*

Proof. Let

$$\tilde{H} = \langle \tilde{a}, \tilde{b}; \tilde{a}^{2s} = \tilde{w}_1^{2t} = \tilde{w}_2^{2u} = \ldots = 1 \rangle$$

where $\tilde{w}_1$ is obtained from w_i by replacing a by $\tilde{a}$ and b by $\tilde{b}$. We view the given special representation σ as a map from H into $\mathbb{C}^*$. Choose $\tilde{\sigma}(\tilde{a})$ and $\tilde{\sigma}(\tilde{b})$ to be square roots of $\sigma(a)$ and $\sigma(b)$ respectively and to have twice the respective orders of $\sigma(a)$ and $\sigma(b)$. These choices determine a representation $\tilde{\sigma}\colon \tilde{H} \to \mathbb{C}^*$ since $\tilde{\sigma}(\tilde{w}_i)^2 = \sigma(w_i)$. By replacing some of the w_i by $a^s w_i$ in the original presentation of H, if necessary, we can arrange $\tilde{\sigma}$ to be a special representation of $\tilde{H}$ in a finite cyclic group. If we now view $\mathbb{C}^*$ as a subgroup of the diagonal matrices in $SL_2(\mathbb{C})$, this provides us with a homomorphism (again denoted $\tilde{\sigma}$) from $\tilde{H}$ into the subgroup of diagonal matrices of $SL_2(\mathbb{C})$ such that

$$\operatorname{tr}(\tilde{\sigma}(\tilde{w}_1)) = \alpha_1 + \alpha_1^{-1}, \operatorname{tr}(\tilde{\sigma}(\tilde{w}_2)) = \alpha_2 + \alpha_2^{-1}, \ldots,$$

where α_1 is a primitive $2t$th root of unity, α_2 is a primitive $2u$th root of unity and so on. Now $\tilde{a}$ maps under $\tilde{\sigma}$ to a matrix A of the form

$$\begin{bmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{bmatrix}.$$

Set

$$A_1 = \begin{bmatrix} \alpha & 0 \\ 1 & \alpha^{-1} \end{bmatrix}.$$

The mapping

$$\tilde{a} \to A_1, \quad \tilde{b} \to \tilde{\sigma}(\tilde{b}) = B, \ldots$$

extends to a homomorphism $\tilde{\rho}$ from $\tilde{H}$ into $SL_2(\mathbb{C})$ since $\tilde{\rho}(\tilde{w}_i)$ has the same trace as $\tilde{\sigma}(w_i)$ and so the defining relations of $\tilde{H}$ are carried into relations in $SL_2(\mathbb{C})$. Now by hypothesis, B has order greater than 4. Suppose that

$$B = \begin{bmatrix} \beta & 0 \\ 0 & \beta^{-1} \end{bmatrix}.$$

Observe then that

$$A_1 B A_1^{-1} B^{-1} = \begin{bmatrix} 1 & 0 \\ \alpha^{-1}(1-\beta^{-2}) & 1 \end{bmatrix} \quad \text{and} \quad A_1 B^2 A_1^{-1} B^{-2} = \begin{bmatrix} 1 & 0 \\ \alpha^{-1}(1-\beta^{-4}) & 1 \end{bmatrix}.$$

Since $\beta^4 \neq 1$, it follows that $\alpha^{-1}(1-\beta^{-2})$ and $\alpha^{-1}(1-\beta^{-4})$ generate a non-cyclic

https://doi.org/10.1017/S0305004100067013 Published online by Cambridge University Press

subgroup of $\mathbb{C}$. The representation $\tilde{\rho}$ factors to give the required special representation ρ.

We are now in a position to prove Theorem C. Suppose then that

$$G = \langle a, b; a^l = b^m = w^n = 1 \rangle$$

is a generalized triangle group with a special cyclic representation σ. Notice that it follows from the hypothesis that σ is special, that at least one of l and m is greater than 2. Thus, by Lemma 4, G has a special representation $\rho: G \to L$ with $\rho(G) \cap V$ non-cyclic and $\rho^{-1}(V) = \operatorname{Ker} \sigma$. This proves the first part of Theorem C.

Now suppose that G has a special dihedral representation σ and that at most one of l, m, n is 2. Since any generating set of a dihedral group must include an element of order 2, it follows that either l or m is 2. Suppose that $l = 2$ and that $m > 2$, $n > 2$. It follows that $\sigma(b)$ and $\sigma(w)$ lie in the same cyclic subgroup of index 2. The preimage H in G of this cyclic subgroup is then of index 2 and it follows from the Reidemeister–Schreier method that H can be presented in the form

$$H = \langle b_0, b_1; b_0^m = b_1^m = w_0^n = w_1^n = 1 \rangle$$

where $b_0 = b$, $b_1 = aba^{-1}$, w_0 is w re-expressed as a word in b_0, b_1 and $w_1 = awa^{-1}$ is similarly re-expressed. Now σ, restricted to H, is a special cyclic representation. Thus, by Lemma 4, there is a special representation $\rho: H \to L$ with $\operatorname{Ker} \sigma = \rho^{-1}(V)$ and with $\rho(H) \cap V$ non-cyclic. This proves the last part of Theorem C.

This work was carried out, in part, at MSRI in June, 1985. We thank MSRI and the NSF for their support.

REFERENCES

[1] B. Baumslag and S. J. Pride. Groups with two more generators than relators. *J. London Math. Soc.* (2) **17** (1978), 425–426.

[2] S. Boyer. On proper powers in free products and Dehn surgery. Preprint (1986).

[3] M. Culler and P. B. Shalen. Varieties of group representations and splittings of 3-manifolds. *Annals of Math.* **117** (1983), 109–146.

[4] F. González-Acuña and H. Short. Knot surgery and primeness. *Math. Proc. Cambridge Philos. Soc.* **99** (1986), 89–102.

[5] C. McA. Gordon and J. Luecke. Only integral Dehn surgeries can yield reducible manifolds. *Math. Proc. Cambridge Philos. Soc.* **102** (1987), 97–101.

[6] R. Horowitz. Characters of free groups represented in the two-dimensional linear group. *Comm. Pure & Appl. Math.* **25** (1972), 635–649.

[7] A. G. Kurosh. Die Untergruppen der freien Produkte von beliebigen Gruppen. *Math. Ann.* **109** (1934), 647–660.

[8] A. Lubotzky and A. Magid. Varieties of representations of finitely generated groups. To appear.

[9] A. I. Mal'cev. On faithful representations of infinite groups of matrices. *American Math. Soc. Translations* (2) **45** (1965), 1–18.

[10] R. Ree and N. S. Mendelsohn. Free subgroups of groups with a single defining relation. *Arch. Math.* **19** (1968), 577–580.

[11] C. T. C. Wall. Rational Euler characteristics. *Proc. Cambridge Philos. Soc.* **57** (1961), 182–184.

https://doi.org/10.1017/S0305004100067013 Published online by Cambridge University Press