ILLINOIS JOURNAL OF MATHEMATICS
Volume 25, Number 2, Summer 1981

DISCRETENESS CRITERIA AND HIGH ORDER GENERATORS FOR SUBGROUPS OF $SL(2, R)$

BY
COURTNEY DOYLE AND DAVID JAMES

1. Introduction

In 1939 and 1940 Lauritzen [3] and Nielson [11] used noneuclidean geometry to prove:

(*) A nonabelian group in $SL(2, R)$ which contains at most hyperbolics and $\pm I$ must be discrete.

Later in 1940 Siegel found a striking proof [12] of the more general theorem:

(**) If the matrices in a group Γ in $SL(2, R)$ do not all have a real (or infinite) fixed point in common, and if they do not all map a pair of real points onto that very pair, then Γ is discrete provided Γ contains no sequence of elliptics converging to I.

Authors presenting these or similar ideas now all use Siegel's proof, but some [10] present (**) while others [6], [7], [9] prefer the more easily grasped but weaker (*).

A formulation permitting a proof somewhat shorter than those for (*) and (**) is found by casting Siegel's proof, with minor modifications, upon class (i) of Theorem 1 of this paper. The result is Theorem 5 below which applies to a slightly wider collection of subgroups, and surprisingly also has more explicit conclusions; (*) and (**) are direct consequences of this theorem and its corollary respectively. The few subgroups for which this theorem does not apply are of a particularly simple nature, and are easily analyzed (classes (ii), (iii), and (iv) of Theorem 1).

The classification of subgroups into the four classes mentioned above also allows us to prove the existence of high order generators for subgroups of $SL(2, R)$ (Theorems 2 and 4).

In this and the next two sections all matrices are assumed to be in the Special Linear group $SL(2, R)$ of two by two matrices with real entries and determinants one.

If

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

Received January 19, 1978.

then MZ denotes $(az + b)/(cz + d)$. The *trace*, tr M, is $a + d$, and the solutions to $Mz = z$, the *fixed points*, are found to be

$$z = (a - d \pm \sqrt{(\operatorname{tr} M)^2 - 4})/2c.$$

F. Klein therefore categorized matrices into three types:

hyperbolic if $|\operatorname{tr} M| > 2$ (fixed points distinct—real or infinite);
parabolic if $|\operatorname{tr} M| = 2$ (fixed points coincide—real or infinite);
elliptic if $|\operatorname{tr} M| < 2$ (fixed points complex conjugates—nonreal).

Each hyperbolic is conjugate to

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix}$$

for some real $\lambda \neq -1, 0, 1$. Each parabolic is conjugate to

$$\begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -1 & \beta \\ 0 & -1 \end{pmatrix}$$

for some real $\beta \neq 0$. Each elliptic is conjugate to

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

for some real θ, not a multiple of π. To verify these three statements, one need only select the conjugator matrix to carry the fixed points to 0, ∞ for a hyperbolic, to ∞ for a parabolic, and to i, $-i$ for an elliptic.

2. Classification

LEMMA 1. *If a group* Γ *contains*

$$M = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix}$$

with real $\lambda \neq -1, 0, 1$, *then* Γ *is contained in one of two classes:*

H1. *Groups generated by hyperbolics, nor all of which have the same pair of fixed points.*

H2. *Groups with at most two types, namely entries zero on the main diagonal, and entries zero off the main diagonal.*

Proof. If

$$N = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then tr $M^nN = \lambda^n a + \lambda^{-n} d$. If not both a and d are zero then for some n, $|\operatorname{tr} M^nN| > 2$, that is, M^nN is hyperbolic. Hence Γ can be generated by hyperbolics together with those of the form

$$K = \begin{pmatrix} 0 & -k \\ 1/k & 0 \end{pmatrix}.$$

Suppose Γ is not in H2, that is, in addition to diagonal matrices, and matrices like K, Γ contains

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

where a or d is nonzero and b or c is nonzero. If K is in a list of generators, it can be replaced in the list with B and BK, and each of these can be replaced with M and a hyperbolic by the reasoning above. So Γ can be generated by hyperbolics. Not all these generators can have the same fixed points as M since then all would be diagonal and Γ would be in H2. So Γ is in H1.

LEMMA 2. *If a group Γ contains*

$$\begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -1 & \beta \\ 0 & -1 \end{pmatrix}$$

with real $\beta \neq 0$, then either Γ is conjugate to H1, *or Γ is contained in the following class:*

P1. *Groups with at most two types,*

$$\begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 & \delta \\ 0 & -1 \end{pmatrix}.$$

Proof. Let

$$M = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

If $c \neq 0$, then since tr $M^nN = a + n\beta c + d$, n can be selected so that M^nN is hyperbolic. There is a conjugate of M^nN which is diagonal, and Lemma 1 can be applied to the corresponding conjugate group. Class H2 cannot contain parabolics, so the conjugate group is in H1.

If $c = 0$, then the determinant condition forces $d = 1/a$, so again Γ contains a hyperbolic unless $a = d = \pm 1$.

The analysis is identical for the case

$$M = \begin{pmatrix} -1 & \beta \\ 0 & -1 \end{pmatrix}.$$

LEMMA 3. *If a group Γ contains*

$$M = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

for some real θ, not a multiple of π, then either Γ is conjugate to H1 *or* H2, *or Γ is contained in the following class.*

E1. *groups with at most the type*

$$\begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}, \quad \phi \text{ real.}$$

Proof. If Γ contains a hyperbolic or parabolic ($\neq \pm I$), then Lemma 1 or Lemma 2 is applicable to some conjugate group, and Γ is conjugate to H1, H2, or P1. The last cannot occur since by assumption, Γ contains more than parabolics.

Suppose Γ contains at most elliptics and $\pm I$, and let

$$N = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

By utilizing $ad - bc = 1$, one calculates the trace of the commutator $MNM^{-1}N^{-1}$ to be

$$2\cos^2\theta + (a^2 + b^2 + c^2 + d^2)\sin^2\theta.$$

But $a^2 + b^2 + c^2 + d^2 = 2 + (a-d)^2 + (b+c)^2$, so the absolute value of the trace is at least two. If the trace exceeds two, Γ contains a hyperbolic, contrary to assumption. So either $\sin\theta = 0$, or both $a = d$ and $b = -c$. The assumption that θ is not a multiple of π eliminates the first case. In the second,

$$N = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

and $a^2 + b^2 = 1$, so there is a ϕ such that $a = \cos\phi$ and $b = \sin\phi$.

We now prove a classification theorem.

If Γ consists only of I or $\pm I$, it lies in P1. For every other Γ, a conjugate group can be found which contains a matrix which is a type in the hypothesis of one of the lemmas. So Γ is conjugate to a group in H1, H2, P1, or E1. Since determinant, trace, order and matrix type are invariant under conjugation, the lemmas give the following:

THEOREM 1. *Every group is in one of the following classes:*

(i) *Groups generated by hyperbolic matrices, not all of which have the same pair of fixed points.*

(ii) *Groups which consist of at most two types, hyperbolics all of which have identical fixed points x and y, and elliptics with trace zero which interchange x and y and which when squared equal $-I$,*

(iii) *Groups which consists of parabolics, all of which have the same fixed point,*

(iv) *Groups which consists of elliptics, all of which have the same two fixed points.*

Also I is required in each group, and $-I$ is permitted.

Groups in (iii) or (iv) are commutative. In (ii), the commutative groups are those which contain only the first type, or those which consist of four matrices, I, $-I$, M, $-M$, where M is of the second type. No group in (i) is commutative. These facts follow from Theorem 1E in [7] that two real bilinear transformations, neither the identity, commute if and only if they have the same set of fixed points. It is not difficult to verify further that for any two matrices M and N in (ii), $(MN)^2 = (NM)^2$, but that any group in (i) contains a pair of matrices for which this fails.

3. High order generators

The minimal positive integer n such that $M^n = I$ is the *order* of M. If no such n exists, the order is infinite. Every elliptic is conjugate to a matrix of the form

$$M = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

and since

$$M^n = \begin{pmatrix} \cos n\theta & \sin n\theta \\ -\sin n\theta & \cos n\theta \end{pmatrix},$$

M has order n exactly when $\theta = 2\pi k/n$ for some integer k relatively prime to n. Hyperbolics and parabolics cannot have finite order.

THEOREM 2. *If Γ contains a matrix L of finite order n, then Γ can be generated by matrices of order n or greater (possibly infinite).*

Proof. Note that $-I$ is not needed as a generator since $-I = (-L)L^{-1}$ for any L. Without loss of generality, Γ is in H1, H2, P1, or E1 since order is invariant under conjugation. The theorem is trivial except for E1.

Suppose Γ in E1 contains

$$L = \begin{pmatrix} \cos 2\pi k/n & \sin 2\pi k/n \\ -\sin 2\pi k/n & \cos 2\pi k/n \end{pmatrix}$$

with order n. Since k is relatively prime to n, there exist integers x and y such that $xk + yn = 1$. Then $N = L^x$ has the form of L but with $k = 1$. Suppose a generator M has order $m < n$, that is, M has the form of L but with $2\pi h/m$. Then

$$MN^{-j} = \begin{pmatrix} \cos 2\pi(h/m - j/n) & \sin 2\pi(h/m - j/n) \\ -\sin 2\pi(h/m - j/n) & \cos 2\pi(h/m - j/n) \end{pmatrix}.$$

Select j so $0 \le h/m - j/n < 1/n$. If the equality occurs, $M = N^j$, and in the list of generators, M can be replaced by N. If the inequality is strict, the order of MN^{-j} exceeds n, and M can be replaced by N and MN^{-j}. The theorem is proved.

We remark that infinite groups of finite order matrices do exist, for instance

$$\Gamma = \left\{ \begin{pmatrix} \cos r\pi & \sin r\pi \\ -\sin r\pi & \cos r\pi \end{pmatrix} : r \text{ rational} \right\}.$$

However if there is a uniform bound n on the orders of the matrices in Γ, then Γ is a period group with period $n!$ so a theorem of Burnside [1, p. 25] implies Γ is finite.

THEOREM 3. *Any finite group Γ contains at most elliptics and $\pm I$, and is cyclic.*

Proof. Only elliptics and $\pm I$ have finite order, so we assume Γ is in E1 or H2. Finite Γ in H2 must be $\{I, M\}$ or $\{I, -I, M, -M\}$, both cyclic, where M has zeros on the main diagonal, because if Γ contains another elliptic N then MN is hyperbolic, forcing Γ to be infinite. In E1 let L be a matrix with the highest occurring order, n, and $N = L^x$ as in the proof of Theorem 2. Every matrix of order n in E1 is a power of N. In particular this includes the generators of Γ guaranteed by Theorem 2. See [2, p. 133] for a different proof of this theorem.

THEOREM 4. *If a group Γ contains a matrix L of infinite order, then Γ can be generated by matrices of infinite order, together with at most one M such that $M^2 = -I$. (This M is necessary only if Γ is in class* (ii) *of Theorem* 1 *and must further satisfy* trace $M = 0$.)

Proof. By conjugation we may assume Γ is in H1, H2, P1, or E1, and the theorem is trivial for H1 and P1. As previously noted, $-I$ is not needed as a generator.

A matrix in H2 with infinite order has the form

$$L = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix} \quad \text{with } \lambda \neq -1, 0, 1.$$

We need to show that if

$$M = \begin{pmatrix} 0 & -b \\ 1/b & 0 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 0 & -\beta \\ 1/\beta & 0 \end{pmatrix},$$

then N can be generated by matrices of infinite order together with M. If $b = \beta$ then $N = M$. If $b = -\beta$ then $N = -M = (-L)L^{-1}M$. If $b \neq \pm\beta$ then

$$MN = \begin{pmatrix} -b/\beta & 0 \\ 0 & -\beta/b \end{pmatrix}$$

has infinite order, and $N = M^{-1}(MN)$.

A matrix in E1 with infinite order has the form

$$L = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

with θ not a rational multiple of π. We need show if

$$N = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$

has finite order, that is, if ϕ is a rational multiple of π, then N is the product of matrices of infinite order. But $N = L^{-1}(LN)$ and

$$LN = \begin{pmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}$$

has infinite order since $\theta + \phi$ is not a rational multiple of π.

4. Principal-circle groups

The results of Section 3 are summarized as follows: Let Γ be a subgroup of $SL(2, R)$.

(i) If there is a uniform bound on the orders of the elements in Γ, then Γ is finite.

(ii) If Γ is finite, it is cyclic.

(iii) If Γ contains an element of finite order n, then Γ can be generated by elements of order n or greater.

(iv) If any element of Γ has infinite order, then Γ can be generated by elements of infinite order together with at most one element which when squared equals $-I$.

These properties are invariant under conjugation and so apply to principal-circle subgroups of $SL(2, C)$, and may also be translated into information about birational transformation groups. They apply as well as to groups faithfully represented by subgroups of $SL(2, R)$. See a paper of Lehner [4] for two such examples.

Another result of Lehner [5] relates to Theorem 1. He generalizes a result of Knopp [8] to prove any horocyclic group of genus g possesses a system of generators consisting entirely of parabolic and elliptic elements, if and only if $g = 0$.

5. Discreteness

A group Γ of matrices is *discrete* if Γ does not contain any sequence of matrices $M_n(\neq I)$ which converges entrywise to I.

We ask which of the matrix principal-circle groups Γ are discrete. Theorem 2H of [7] states that if Γ is discrete, and all elements of Γ have a common fixed point, then Γ (when considered as a bilinear transformation group) must be cyclic. Thus for Γ in class (iv) of Theorem 1, Γ is discrete if and only if it is finite cyclic.

Γ in class (iii) is discrete if and only if Γ is cyclic or is a cyclic group with $-I$ adjoined.

For Γ in (ii), a similar argument yields that the subgroup of principal diagonal matrices with positive entries must be cyclic. So Γ is discrete if and only if it is generated by at most two matrices, one hyperbolic and the other either $-I$ or an elliptic which interchanges the fixed points of the hyperbolic, has trace zero, and which when squared equals $-I$.

We now consider discreteness for Γ in class (i).

THEOREM 5. *Suppose a principal-circle group Γ in class* (i) *of Theorem* 1 *fails to be discrete. Then either there is a sequence of parabolics converging to I or there are two sequences, one of hyperbolics and the other of elliptics, each converging to I.*

Proof. Since discreteness is invariant under conjugation, we may assume Γ is in H1. Therefore Γ contains

$$L = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix} \quad \text{with } \lambda \neq -1, 0, 1,$$

$$K = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{neither diagonal nor antidiagonal,}$$

and Γ contains a sequence of matrices

$$M_n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix} \to I.$$

We may assume the sequence M_n has no diagonal matrices, by replacing any diagonal M_n by the non-diagonal KM_nK^{-1}. The new sequence still converges to I. The commutator

$$N_n = LM_nL^{-1}M_n^{-1} = \begin{pmatrix} 1 + (1 - \lambda^2)b_n c_n & (\lambda^2 - 1)a_n b_n \\ (\lambda^{-2} - 1)c_n d_n & 1 + (1 - \lambda^{-2})b_n c_n \end{pmatrix}$$

has trace $2 - b_n c_n(\lambda - 1/\lambda)^2$, and $N_n \to I$.

If $b_n c_n > 0$ for infinitely many n, then the subsequence of N_n for which this holds consists of elliptics, and the corresponding M_n are hyperbolics, since $a_n d_n = 1 + b_n c_n > 1$ which forces $|a_n + d_n| > 2$.

If $b_n c_n = 0$ for infinitely many n, then the corresponding N_n are parabolics. No N_n is I, because neither a_n nor d_n is zero (since $a_n d_n - b_n c_n = 1$), and only one of b_n or c_n is zero (since M_n is not diagonal).

If $b_n c_n < 0$ for infinitely many n, the corresponding N_n are hyperbolics. Let

$$P_n = LN_nL^{-1}N_n^{-1},$$

so that tr $P_n = 2 + (\lambda - 1/\lambda)^4 a_n b_n c_n d_n$ and $P_n \to I$. Since $a_n d_n \to 1$ and $b_n c_n \to 0$ and are negative, the P_n are elliptic from some point on.

COROLLARY. *Suppose Γ satisfies the hypotheses of Theorem 5.*

(a) *If there is a point which is left fixed by every matrix, then there is a sequence of parabolics converging to I.*

(b) *If there is no such point, there must be a sequence of hyperbolics and one of elliptics, each converging to I.*

Proof. By conjugation, we assume Γ is in $SL(2, R)$. Hyperbolics have real (or infinite) fixed points, and elliptics do not, so in (a), Γ contains no elliptics, and the corollary is immediate.

Part (*b*) is immediate from the proof of Theorem 5, unless every

$$\begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix} \to I$$

has $b_n c_n = 0$ from some point on. If this last condition is fulfilled either infinitely many $b_n = 0$ or infinitely many $c_n = 0$. Without loss of generality, assume the latter. From the sequence N_n of Theorem 5 we extract $L_n \to I$ which have the form

$$L_n = \begin{pmatrix} 1 & \beta_n \\ 0 & 1 \end{pmatrix}$$

where $\beta_n = (\lambda^2 - 1)a_n b_n \neq 0$, so $\beta_n \to 0$. By assumption there is a

$$K = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

which does not fix infinity, that is, $c \neq 0$. Now

$$L'_n = KL_nK^{-1} = \begin{pmatrix} 1 - ac\beta_n & a^2\beta_n \\ -c^2\beta_n & 1 + ac\beta_n \end{pmatrix} \to I,$$

so $a^2\beta_n = 0$ from some point on since $-c^2\beta_n \neq 0$. Thus $a = 0$, so

$$L_nL'_n = \begin{pmatrix} 1 - c^2\beta_n & \beta_n \\ -c^2\beta_n & 1 \end{pmatrix} \to I.$$

Part (b) follows from our first remark.

We remark that for Γ in class (i) satisfying part (a) of the corollary, Γ cannot be discrete since it is not cyclic even when considered as a bilinear transformation group. So the assumption "Γ fails to be discrete" is redundant for this half of the corollary.

In summary, for Γ in classes (ii), (iii), or (iv) of Theorem 1, discreteness criteria were easily found. For Γ in class (i), Theorem 5 and its corollary apply.

BIBLIOGRAPHY

1. C. W. CURTIS and I. REINER, *Representation theory of finite groups and associative algebras*, vol. II, Interscience, New York, 1962.
2. L. FORD, *Automorphic functions*, Chelsea, New York, 1951.
3. S. LAURITZEN, *En Saetning om Grupper af lineaere Substitutioner*, Mat. Tidsskr. B, 1939, pp. 69–76.
4. J. LEHNER, *Representations of a class of finite groups*, Mich. Math. J., vol. 7 (1960), 233–236.
5. ———, *On the generation of discontinuous groups*, Pacific J. Math., vol. 13 (1963), pp. 169–170.
6. ———, *Discontinuous groups and automorphic functions*, Math. Surveys, vol. 8, Amer. Math. Soc., Providence, R. I., 1964.
7. ———, *A short course in automorphic functions*, Holt, Rinehart, and Winston, New York, 1966.
8. M. KNOPP, *Construction of a class of modular functions and forms*, I and II, Pacific J. Math., vol. 11 (1961), pp. 275–293, 661–678.
9. H. MAASS, *Lectures on modular functions of one complex variable*, Tata Institute of Fundamental Research, Bombay, 1964.
10. W. MAGNUS, *Discrete groups*, New York University Notes, New York, 1952.
11. J. NIELSEN, *Uber Gruppen linearer Transformationen*, Mitt. Math. Ges. Hamburg, vol. 8 (1940), pp. 82–104.
12. C. L. SIEGEL, *Bemerkung zu einem Satze von Jakob Nielsen*, Mat. Tidsskr. B, 1940, pp. 66–70.

3555 HURON AVENUE
LAKEPORT, MICHIGAN
THE UNIVERSITY OF MICHIGAN–DEARBORN
DEARBORN, MICHIGAN