Vol. 69, No. 3 DUKE MATHEMATICAL JOURNAL © March 1993

JORGENSEN'S INEQUALITY FOR DISCRETE GROUPS IN NORMED ALGEBRAS

SHMUEL FRIEDLAND AND SA'AR HERSONSKY

0. Introduction. Let $\mathscr{A}$ be a normed algebra with identity over $\mathbf{C}$ with the norm $|\cdot|$. Let $\Gamma \subset \mathscr{A}$ be a group of invertible elements in $\mathscr{A}$. We then regard Γ as a topological group with the topology induced by the norm $|\cdot|$. Assume that $a, b \in \mathscr{A}$ are invertible. Denote by $\langle a, b\rangle$ the group generated by a and b. Suppose that $\langle a, b\rangle$ is a discrete group. What can one say about a, b? The best known result in this area is Jorgensen's inequality [Jor]. Let $\mathscr{A} = M_2(\mathbf{C})$ be the algebra on 2×2 complex valued matrices and assume $\langle a, b\rangle$ is a subgroup of the special linear group of $SL_2(\mathbf{C}) \subset M_2(\mathbf{C})$. Then the sharp inequality of Jorgensen claims that if $\langle a, b\rangle$ do not generate an elementary group then

$$|(trace(a))^2 - 4| + |trace([a, b]) - 2| \geqslant 1. \tag{0.1}$$

Here, $[a, b] = aba^{-1}b^{-1}$. We call $\langle a, b\rangle$ an elementary group if and only if $\langle a, b\rangle$ has a nilpotent subgroup of a finite index. Note that (0.1) is invariant with respect to conjugacy in $GL_2(\mathbf{C})$. Jorgensen's inequality translates immediately to Kleinian groups—discrete groups of Möbius transformations of the Riemann sphere. Jorgensen's inequality was recently generalized by Martin [Mar1] to nonelementary discrete groups of Möbius transformations of any dimension $n > 2$. A simple version of Martin's inequality can be stated in our terms as follows. If $\langle a, b\rangle \subset \mathscr{A}$ generate a discrete nonelementary group then

$$\max(|a - 1|, |b - 1|) \geqslant 2 - \sqrt{3}. \tag{0.2}$$

To obtain the corresponding results for Möbius transformations one recalls that orientation preserving Möbius transformations are isomorphic to the group $SO^+(1, n)(\mathbf{R}) \subset M_{n+1}(\mathbf{C})$. In that case, the norm $|\cdot|$ is assumed to be the spectral norm. In [Mar2] Martin's inequalities are used to obtain new lower bounds for the volume of all hyperbolic n-manifolds.

The object of this paper is twofold. We first study necessary conditions for discreteness of the group $\langle a, b\rangle \subset \mathscr{A}$. By doing that we believe that one can get similar results for discrete groups acting on other homogenous spaces, for example the Siegel upper half plane. We also show that there are other variants of Martin's inequality (0.2). Second, we claim that if a, b belong to some classical groups, for example $SO^+(1, n)(\mathbf{R})$ or $Sp(n, \mathbf{R})$ then the inequality (0.2) can be improved by

Received 17 July 1992. Revision received 10 September 1992.

replacing the constant $2-\sqrt{3} \sim 0.2679$ by $\tau \sim 0.2971$. This is done in the general setting by considering normed algebras with an involution.

We now survey briefly the contents of the paper. In §1 we consider the basic iterations:

$$x_{m+1} = [a, x_m], \qquad m = 0, \ldots, x_0 = b. \tag{0.3}$$

We give conditions on a and b which imply $\lim_{m\to\infty} x_m = 1$. We call the group $\langle a, b\rangle$ an a-nilpotent group if the iterations (0.3) stop at 1 after a finite number of steps, i.e., $x_m = 1$ for some m. We then state a generalization of Jorgensen's inequality for non-a-nilpotent groups. We show that our condition implies (0.2). In §2 we improve our bounds to certain groups in normed algebra with an involution. In particular, we replace the constant $2-\sqrt{3}$ in (0.2) by the constant τ. In §3 we discuss our results for the algebras of $n \times n$ complex valued matrices. In §4 we estimate from below the radius of the largest ball in a hyperbolic n-manifold following Martin [Mar2] and using our bounds. In the last section we give generalizations of Shimuzu-Leutbecher and Jorgensen inequalities to certain discrete nonelementary subgroups of the symplectic group as in [Her1].

1. Iterations in normed algebras. Let $\mathscr{A}$ be a normed algebra with identity over **C** with the norm $|\cdot|$. That is, $|\cdot|$ is a submultiplicative norm on $\mathscr{A}$ and $|1| = 1$. For any linear bounded operator $T\colon \mathscr{A} \to \mathscr{A}$ we let $\|T\| = sup_{|x|\leqslant 1} |T(x)|$. The following lemma is a basic tool in our arguments.

LEMMA 1.1. *Let $a \in \mathscr{A}$ be an invertible element. Associate with a the bounded linear operator*

$$\hat{a}\colon \mathscr{A} \to \mathscr{A}, \qquad \hat{a}(x) = axa^{-1} - x. \tag{1.2}$$

Then $\hat{a}(1) = 0$ and $|\hat{a}(x)| \leqslant \|\hat{a}\|\,|x-1|$. Furthermore

$$\|\hat{a}\| \leqslant 2\min(|a-1|\,|a^{-1}|, |a^{-1}-1|\,|a|). \tag{1.3}$$

Proof. Clearly, $\hat{a}(1) = 0$. Then,

$$|\hat{a}(x)| = |\hat{a}(x-1)| \leqslant \|\hat{a}\|\,|x-1|.$$

The following inequalities yield (1.3).

$$|\hat{a}(x)| = |(ax - xa)a^{-1}| = |((a-1)x - x(a-1))a^{-1}| \leqslant 2|x|\,|a-1|\,|a^{-1}|,$$

$$|\hat{a}(x)| = |a(xa^{-1} - a^{-1}x)| = |a(x(a^{-1}-1) - (a^{-1}-1)x)| \leqslant 2|x|\,|a^{-1}-1|\,|a|.$$

Consider the following iterations:

$$x_{n+1} = [a, x_n], \qquad n = 0, 1, \ldots, x_0 = b. \tag{1.4}$$

We now study the case when $\lim_{n\to\infty} x_n = 1$.

THEOREM 1.5. *Let $\mathscr{A}$ be a normed algebra and assume that $a, b \in \mathscr{A}$ are invertible elements. Consider the iterations* (1.4). *Assume that $\|\hat{a}\| < 1$. If one of the following conditions holds*

$$|b - 1| < 1 - \|\hat{a}\|, \qquad |[a, b] - 1| < 1 - \|\hat{a}\| \tag{1.6}$$

then $\lim_{n\to\infty} x_n = 1$

Proof. Let $z \in \mathscr{A}$. Assume that $|z| < 1$. Recall that $1 + z$ is invertible and the following conditions hold

$$(1 + z)^{-1} = \sum_0^\infty (-z)^n, \qquad |(1 + z)^{-1}| \leqslant \frac{1}{1 - |z|}. \tag{1.7}$$

Set $z_n = x_n - 1$. We now claim that

$$|z_{n+1}| \leqslant \frac{\|\hat{a}\| \, |z_n|}{1 - |z_n|}. \tag{1.8}$$

Note that (1.4) is equivalent to

$$x_{n+1} - 1 = a x_n a^{-1} x_n^{-1} - 1 = \hat{a}(x_n) x_n^{-1}. \tag{1.4$'$}$$

Lemma 1.1 yields:

$$|x_{n+1} - 1| \leqslant \|\hat{a}\| \, |x_n - 1| \, |x_n^{-1}|. \tag{1.9}$$

Assume first that the first condition of (1.6) holds. That is $|z_0| < 1 - \|\hat{a}\|$. Assume by induction that $|z_n| < 1 - \|\hat{a}\|$. Use (1.9) and (1.7) to obtain (1.8). To this end consider the map

$$f\colon [0, r] \to [0, r], \qquad f(x) = \frac{\|\hat{a}\| x}{1 - x}, \qquad r = 1 - \|\hat{a}\|.$$

Note that $0, r$ are two fixed points of f. Furthermore, $f(x) < x$ on $(0, r)$. It then follows that any sequence

$$\{r_n\}_0^\infty, \qquad r_{n+1} = f(r_n), \qquad n = 1, \ldots, 0 < r_1 < r \tag{1.10}$$

decreases to the attracting point 0. The induction hypothesis $|z_n| < r$ implies that either $z_{n+1} = 0$ or $0 < |z_{n+1}| \leqslant f(|z_n|) < |z_n| < r$. As the sequence $\{r_n\}_1^\infty$ given in (1.10) converges to 0 we deduce that $\lim_{n\to\infty} z_n = 0$. This proves the theorem in the case that the first condition of (1.6) holds. Assume that the second condition of (1.6) holds. We consider the iterations (1.4) starting from $x_1 = [a, b]$. As the second condition of (1.6) yields that $|z_1| < r$ we deduce the theorem in a similar way. □

The following theorem can be considered as a generalization of Jorgensen's inequality.

THEOREM 1.11. *Let $\mathscr{A}$ be a normed algebra. Assume that $\langle a, b\rangle$ is a discrete non-a-nilpotent group. Then the following conditions hold*

$$\begin{gathered}\|\hat{a}\| + |[a, b] - 1| \geqslant 1,\\ \|\hat{a}\| + |b - 1| \geqslant 1.\end{gathered} \tag{1.12}$$

In particular,

$$\max(|[a, b] - 1|, |a - 1|) \geqslant 2 - \sqrt{3}, \quad \max(|b - 1|, |a - 1|) \geqslant 2 - \sqrt{3}. \tag{1.13}$$

Proof. Assume to the contrary that at least one of the conditions of (1.12) does not hold. According to Theorem 1.5 the iterations (1.4) converge to 1. As $\langle a, b\rangle$ is discrete it follows that $x_n = 1$ for some n. That is, $\langle a, b\rangle$ is a-nilpotent, contrary to our assumptions.

Let $\chi = \max(|[a, b] - 1|, |a - 1|)$. Assume to the contrary that $\chi < \kappa = 2 - \sqrt{3}$. As $|a - 1| < \kappa < 1$ from (1.3) and (1.7) we obtain:

$$\|\hat{a}\| < \frac{2\kappa}{1 - \kappa} = 1 - \kappa < 1 - |[a, b] - 1|.$$

This contradicts the first inequality of (1.12). The above contradiction proves the first inequality of (1.12). The second inequality is established in a similar manner. □

Let

$$B(a, r) = \{x: x \in \mathscr{A}, |x - a| \leqslant r\}, \qquad B_0(a, r) = \{x: x \in \mathscr{A}, |x - a| < r\}.$$

THEOREM 1.14. *Let A be a normed algebra. Assume that Γ is a discrete group. Then for any $0 < r < 2 - \sqrt{3}$ the group generated by the elements $\Gamma \cap B_0(1, r)$ is nilpotent. That is, there exists a k-so that for any sequence*:

$$a_i \in B_0(1, r), \qquad i = 0, \ldots, y_i = [a_i, y_{i-1}], \qquad i = 1, \ldots, y_0 = a_0, \tag{1.15}$$

$y_k = 1$. In particular, if $\langle a, b\rangle$ is a discrete group and

$$|a - 1| < 2 - \sqrt{3}, \qquad |b - 1| < 2 - \sqrt{3} \tag{1.16}$$

then $\langle a, b\rangle$ is a nilpotent group.

Proof. As Γ is a discrete group, there exists a positive ε so that $\Gamma \cap B_0(1, \varepsilon) = \{1\}$. Let $\kappa = 2 - \sqrt{3}$, $c = \frac{r}{\kappa} < 1$. We claim that

$$|y_i - 1| < c^{i+1}\kappa, \qquad i = 0, 1, \ldots. \tag{1.17}$$

We prove the above inequality by induction. As $y_0 = a_0 \in B_0(1, r)$, (1.17) holds trivially for $i = 0$. Assume that (1.17) holds for $i = n$. Then use (1.7) and the induction hypothesis to deduce

$$\begin{aligned}
|y_{n+1} - 1| &= |a_n y_n a_n^{-1} y_n^{-1} - 1| = |(a_n y_n - y_n a_n) a_n^{-1} y_n^{-1}| \\
&= |((a_n - 1)(y_n - 1) - (y_n - 1)(a_n - 1)) a_n^{-1} y_n^{-1}| \\
&\leqslant 2|a_n - 1||y_n - 1||a_n^{-1}||y_n^{-1}| \leqslant \frac{2|a_n - 1||y_n - 1|}{(1 - |a_n - 1|)(1 - |y_n - 1|)} \\
&< \frac{2c\kappa c^{n+1}\kappa}{(1 - c\kappa)(1 - c^{n+1}\kappa)} < \frac{c^{n+2}2\kappa^2}{(1 - \kappa)^2} = c^{n+2}\kappa.
\end{aligned}$$

The above inequality proves (1.17). Let k be the smallest positive integer so that $c^{k+1}\kappa < \varepsilon$. Then (1.17) yields that $y_k \in B_0(1, \varepsilon)$. Hence, $y_k = 1$ and the group generated by the elements $\Gamma \cap B_0(1, r)$ is nilpotent, e.g. [Hup, III.1.11].

Suppose that (1.16) holds. Then there exists $r < \kappa$ so that $a, b \in B_0(1, r)$. Consider the sequence (1.15) with each a_i equal to either a or b for $i = 0, 1, \ldots,$. As $y_k = 1$ we deduce that $\langle a, b \rangle$ is nilpotent. □

Theorem 1.14 is essentially due to Martin [Mar1]. That is, we showed that $B_0(x, r)$ is a Zassenhaus neighborhood for any $r < \kappa$. Note that in this generality, the closed ball $B(1, r)$ is a compact set if and only if $\mathscr{A}$ is a finite dimensional normed algebra.

2. Normed algebras with an involution. Let $\mathscr{A}$ be a normed algebra with identity over $\mathbf{C}$ with the norm $|\cdot|$. A real involution $^\dagger : \mathscr{A} \to \mathscr{A}$ satisfies the conditions:

$$\begin{gathered}
(\alpha a + \beta b)^\dagger = \alpha a^\dagger + \beta b^\dagger, \qquad (ab)^\dagger = b^\dagger a^\dagger, \qquad 1^\dagger = 1, \\
|a^\dagger| = |a|, \qquad a, b \in \mathscr{A}, \qquad \alpha, \beta \in \mathbf{C}.
\end{gathered} \tag{2.1}$$

A complex involution $^\dagger : \mathscr{A} \to \mathscr{A}$ satisfies the conditions:

$$\begin{gathered}
(\alpha a + \beta b)^\dagger = \bar{\alpha} a^\dagger + \bar{\beta} b^\dagger, \qquad (ab)^\dagger = b^\dagger a^\dagger, \qquad 1^\dagger = 1, \\
|a^\dagger| = |a|, \qquad a, b \in \mathscr{A}, \qquad \alpha, \beta \in \mathbf{C}.
\end{gathered} \tag{2.2}$$

We say that $\mathscr{A}$ is a normed algebra with an involution if either (2.1) or (2.2) holds. Let $j \in \mathscr{A}$ be an invertible element. We then denote by $O(j) \subset \mathscr{A}$ the following group

$$O(j) = \{a\colon a^{\dagger} j a = j,\ a \in \mathscr{A}\}. \tag{2.3}$$

We call $O(j)$ the j orthogonal group. Recall that j is called an isometry if

$$|ja| = |aj| = |a|, \qquad \forall a \in \mathscr{A}.$$

Clearly, j is an isometry if and only if j^{-1} is an isometry. We shall show later on that the classical spinor group and the (p, q) orthogonal groups are j orthogonal and j is an isometry, for corresponding matrices j. The aim of this section is to improve the results of §1 in the case $a, b \in O(j)$ and j is an isometry. Our main tool is the following obvious lemma:

LEMMA 2.4. *Let $\mathscr{A}$ be a normed algebra. Assume that a is an invertible element. Suppose furthermore that $a^{-1} = j^{-1} a j$ and j is an isometry. Then*

$$|a^{-1} - 1| = |a - 1|, \qquad |a^{-1}| = |a| \leqslant 1 + |a - 1|. \tag{2.5}$$

Let $\mathscr{A}$ be a normed algebra with an involution. Assume that $a \in O(j)$ and j is an isometry. Then the conditions (2.5) *hold.*

THEOREM 2.6. *Let $\mathscr{A}$ be a normed algebra with an involution. Assume that j is an invertible element and an isometry. Suppose that $a, b \in O(j)$. Consider the iterations* (1.4). *Assume that $\|\hat{a}\| < 1$. If one of the following conditions holds*

$$\begin{gathered} |b - 1| < \frac{1}{\|\hat{a}\|} - 1, \\ |[a, b] - 1| < \frac{1}{\|\hat{a}\|} - 1 \end{gathered} \tag{2.7}$$

then $\lim_{n\to\infty} x_n = 1$. Thus, if $\langle a, b\rangle \subset O(j)$ is a discrete non–a-nilpotent group, then

$$|b - 1| \geqslant \frac{1}{\|\hat{a}\|} - 1, \qquad |[a, b] - 1| \geqslant \frac{1}{\|\hat{a}\|} - 1. \tag{2.8}$$

In particular,

$$\max(|[a, b] - 1|, |a - 1|) \geqslant \tau, \quad \max(|b - 1|, |a - 1|) \geqslant \tau \tag{2.9}$$

Here τ is the unique positive solution of the cubic equation

$$2\tau(1 + \tau)^2 = 1, \quad \tau > 0.2971. \tag{2.10}$$

Proof. Our proof is a corresponding modification of the proofs of Theorem 1.5. We just point out the modifications one should make in these proofs. Use inequalities (1.9) and (2.5) to deduce

$$|z_{n+1}| \leqslant \|\hat{a}\| \, |z_n|(1 + |z_n|). \tag{1.8$'$}$$

Let f be the function

$$f\colon [0, r] \to [0, r], \qquad f(x) = \|\hat{a}\| x(1 + x), \qquad r = \frac{1}{\|\hat{a}\|} - 1.$$

As in the proof of Theorem 1.5 we obtain that either of the conditions of (2.7) implies that $\lim_{n\to\infty} x_n = 1$. Assume that $\langle a, b\rangle$ is a discrete non-a-nilpotent group. Then the arguments of the proof of Theorem 1.5 yield (2.8). Let $\chi = \max(|b - 1|, |a - 1|)$. Assume to the contrary that $\chi < \tau$. As $|a - 1| < \tau$ from (1.3) and (2.5) we obtain:

$$\|\hat{a}\| < 2\tau(1 + \tau).$$

Hence

$$\frac{1}{\|\hat{a}\|} - 1 > \frac{1}{2\tau(1 + \tau)} - 1 = \tau > |b - 1|.$$

This contradicts the first inequality of (2.8). The above contradiction proves the first inequality of (2.9). The second inequality is established in a similar manner. □

THEOREM 2.11. *Let $\mathscr{A}$ be a normed algebra with an involution. Assume that j is an invertible element and an isometry. Let $\Gamma \subset O(j)$ be a discrete group. Then for any $0 < r < \tau$, where τ is given by* (2.10), *the group generated by the elements $\Gamma \cap B_0(1, r)$ is nilpotent. That is, there exists a k so that for any sequence* (1.15) $y_k = 1$. *In particular, if $\langle a, b\rangle$ is a discrete group and*

$$|a - 1| < \tau, \qquad |b - 1| < \tau \tag{2.12}$$

then $\langle a, b\rangle$ is a nilpotent group.

Proof. As Γ is a discrete group, there exists a positive ε so that $\Gamma \cap B_0(1, \varepsilon) = \{1\}$. Let $c = r/\tau$. We claim that

$$|y_i - 1| < c^{i+1}\tau, \qquad i = 0, 1, \ldots. \tag{2.13}$$

We prove the above inequality by induction. As $y_0 = a_0 \in B_0(1, r)$, (2.13) holds trivially for $i = 0$. Assume that (2.13) holds for $i = n$. Then use Lemma 2.4 and the induction hypothesis to deduce

$$\begin{aligned} |y_{n+1} - 1| &= |a_n y_n a_n^{-1} y_n^{-1} - 1| = |(a_n y_n - y_n a_n) a_n^{-1} y_n^{-1}| \\ &= |((a_n - 1)(y_n - 1) - (y_n - 1)(a_n - 1)) a_n^{-1} y_n^{-1}| \\ &\leqslant 2|a_n - 1||y_n - 1||a_n^{-1}||y_n^{-1}| \\ &\leqslant 2|a_n - 1||y_n - 1|(1 + |a_n - 1|)(1 + |y_n - 1|) \\ &< 2c\tau c^{n+1}\tau(1 + c\tau)(1 + c^{n+1}\tau) < c^{n+2} 2\tau^2 (1 + \tau)^2 = c^{n+2}\tau. \end{aligned}$$

The above inequality proves (2.13). The rest of the proof is concluded as the proof of Theorem 1.14. □

We conclude this section with a variation of Theorem 2.11 which can be used to improve Martin's lower bound for the radius r_n of the largest ball contained in any n-hyperbolic manifold [Mar2].

Set

$$\tilde{B}(1, r) = \{x \colon x \in \mathrm{A}, |x||x - 1| \leqslant r\}, \qquad \tilde{B}_0(1, r) = \{x \colon x \in \mathscr{A}, |x||x - 1| < r\}.$$

Assume that $x \in O(j)$. From (2.5) we deduce

$$|x - 1| = |x^{-1} - 1| = |x^{-1}(1 - x)| \leqslant |x^{-1}||1 - x| = |x||x - 1|.$$

Hence

$$\tilde{B}(1, r) \subset B(1, r), \qquad \tilde{B}_0(1, r) \subset B_0(1, r). \tag{2.14}$$

Denote by ω the positive solution of

$$2\omega(2\omega^2 + 1) = 1, \qquad \omega \sim 0.3855. \tag{2.15}$$

THEOREM 2.16. *Let $\mathscr{A}$ be a normed algebra with an involution. Assume that j is an invertible element and an isometry. Let $\Gamma \subset O(j)$ be a discrete group. Then for any $0 < r < \omega$, the group generated by the elements $\Gamma \cap \tilde{B}_0(1, r)$ is nilpotent.*

Proof. Our proof is a slight modification of the proof of Theorem 2.11 and we point out the specific changes that should be made in the above proof. Let $c = r/\omega$. We claim that

$$|y_i||y_i - 1| < c^{i+1}\omega, \qquad i = 0, 1, \ldots$$

From the proof of Theorem 2.11 it follows that

$$|y_{n+1} - 1| \leqslant 2|a_n - 1||y_n - 1||a_n^{-1}||y_n^{-1}| = 2|a_n - 1||a_n||y_n - 1||y_n| < 2c^{n+2}\omega^2.$$

Use the last inequality of (2.5) and the above inequality to get

$$|y_{n+1}||y_{n+1}-1| \leqslant (|y_{n+1}-1|+1)|y_{n+1}-1| < (2\omega^2+1)2c^{n+2}\omega^2 = c^{n+2}\omega.$$

3. Matrix algebras. In this section we assume that $\mathscr{A} = M_n(\mathbf{C})$-algebra of $n \times n$ complex valued matrices. Denote by $\Lambda(A) = \{\lambda_1, \ldots, \lambda_n\}$ the n eigenvalues of A counted with their multiplicity and let $\rho(A) = \max_{1 \leqslant i \leqslant n} |\lambda_i|$ be the spectral radius of A. A is called diagonable if there exists $X \in GL_n(\mathbf{C})$ so that $A = XDX^{-1}$ where $D = diag(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix. In that case the columns of X form a set of linearly independent eigenvectors of A. One views $M_n(C) = Hom(\mathbf{C}^n, \mathbf{C}^n)$ where $A(x) = Ax$, $x \in \mathbf{C}^n$, $A \in M_n(\mathbf{C})$. Let $|\cdot|_p$ be the l_p norm on $\mathbf{C}^n$

$$|(x_1, \ldots, x_n)^T|_p = \left(\sum_1^n |x_i|^p\right)^{1/p}, \qquad 1 \leqslant p \leqslant \infty.$$

It is well known that the conjugate norm of $|\cdot|_p$ is $|\cdot|_q$, $p^{-1}+q^{-1}=1$. Let

$$|X|_p = \max_{|x|_p \leqslant 1} |Xx|_p, \qquad X \in M_n(\mathbf{C})$$

be the corresponding operator norm on $M_n(\mathbf{C})$. Thus, $M_n(\mathbf{C})$ is a normed algebra with respect to any $|\cdot|_p$, $1 \leqslant p \leqslant \infty$. Clearly, the topology induced by any of this norm is the standard topology on $M_n(\mathbf{C}) \equiv \mathbf{C}^{n^2}$. It is easy to see that $|\overline{X}|_p = |X|_p$. It is also known that

$$|X^*|_p = |X^T|_p = |X|_q, \qquad p^{-1}+q^{-1}=1. \tag{3.1}$$

See, for example, [Fri] for this and other basic results needed here. The spectral norm $|X|_2$ is equal to $\sqrt{\rho(AA^*)} = \sqrt{\rho(A^*A)}$. The group of isometries of the l_2 norm is the group of the unitary matrices $U_n \subset GL_n(\mathbf{C})$. Hence U_n is also the group of isometries of the spectral norm. The natural real involution on $M_n(\mathbf{C})$ is $A \mapsto A^T$ where A^T is the transposed matrix of A. The natural complex involution is $A \mapsto A^*$ where A^* is the conjugate transpose of A. The corresponding norm of $M_n(\mathbf{C})$ which is invariant under the above involutions is the spectral norm. Moreover, $M_n(\mathbf{C})$ is a C^* algebra under the spectral norm and the involution $A \mapsto A^*$. (Note that according to (3.1) the norms $|\cdot|_p$, $p \neq 2$ are not invariant under any of the above involutions.)

Let $J \in GL_n(\mathbf{C})$. We then let

$$\begin{aligned} O(J) &= \{A,\ A^TJA = J,\ A \in GL_n(\mathbf{C})\}, \\ U(J) &= \{A,\ A^*JA = J,\ A \in GL_n(\mathbf{C})\}. \end{aligned} \tag{3.2}$$

We call $O(J)$ and $U(J)$ the J orthogonal and J unitary groups respectively. The classical $(m, n-m)$ orthogonal group—$SO(m, n-m)$ group corresponds to a diagonal J with m "1"'s and $n-m$ "-1"'s on the diagonal. The classical spinor group $Sp(n)$ corresponds to the matrix 2×2 block matrix J:

$$J = (J_{ij})_1^2, \qquad J_{11} = J_{22} = 0, \qquad J_{12} = -J_{21} = I \in GL_n(\mathbf{C}).$$

In the both cases J is a unitary matrix. Thus, we can apply Theorem 2.11 to the discrete nonelementary subgroups $\langle a, b\rangle$ contained either in $SO(1, n)$ or $Sp(2n)$. Assume first that $a, b \in SO(1, n)$. As the constant τ appearing in (2.12) is definitely bigger than $2 - \sqrt{3}$ one can improve the recent lower bounds of Martin [Mar2] for the volume of all hyperbolic n-manifolds. We hope that in the case $a, b \in Sp(2n)$ Theorem 2.11 will give lower bound for the volume of manifolds whose universal covering spaces are Siegel upper half planes.

Consider again the iterations (0.3) on $GL_n(\mathbf{C})$. For a fixed $A \in GL_n(\mathbf{C})$ let

$$\Phi\colon GL_n(\mathbf{C}) \to GL_n(\mathbf{C}), \qquad \Phi(X) = AXA^{-1}X^{-1}. \tag{3.3}$$

According to the standard notation of dynamical systems let

$$\Phi^{\circ 0} = Id, \qquad \Phi^{\circ 1} = \Phi, \qquad \Phi^{\circ(m+1)} = \Phi^{\circ m}(\Phi), \qquad m = 1, \ldots, .$$

Thus, the m-th element X_m of the sequence (0.3) is given by $\Phi^{\circ m}(X_0)$. Note that $GL_n(\mathbf{C})$ is an algebraic group and map (3.3) is a holomorphic rational map. Furthermore, the identity matrix I is a fixed point of Φ. It is now clear that Theorem 1.5 gives sufficient conditions for I to be an isolated attracting point. Furthermore, the inequalities (1.6) estimate the basin of the attraction of I. The following theorem gives the exact condition when I is an attractor:

THEOREM 3.4. *Let $A \in GL_n(\mathbf{C})$ and consider the map (3.3). Then the Jacobian map $J(\Phi)(I)\colon M_n(\mathbf{C}) \to M_n(\mathbf{C})$ at the fixed point I is equal to $\hat{A}$. In particular, I is an isolated hyperbolic attracting point if and only if*

$$\max_{\lambda_i, \lambda_j \in \Lambda(A)} \left|\frac{\lambda_i}{\lambda_j} - 1\right| < 1. \tag{3.5}$$

Proof. As the Lie algebra of $GL_n(\mathbf{C})$ is $M_n(\mathbf{C})$, it follows that around I we have the equalities:

$$X = I + Z, X^{-1} = I - Z + \cdots, \Phi(X) = A(I+Z)A^{-1}(I - Z + \cdots)$$
$$= I + (AZA^{-1} - Z) + \cdots \Rightarrow J(\Phi)(I) = \hat{A}.$$

We claim that the n^2 eigenvalues of $\hat{A}$ are $(\lambda_i/\lambda_j) - 1$, $i, j = 1, \ldots, n$. Assume first that A is diagonable, i.e., $A = TDT^{-1}$. Conjugating by T we may assume that

$A = D = diag(\lambda_1, \ldots, \lambda_n)$. Let $Z = (z_{ij})_1^n \in M_n(\mathbf{C})$. Then

$$\hat{D}(Z) = (\lambda_i z_{ij} \lambda_j^{-1} - z_{ij})_1^n = ((\lambda_i \lambda_j^{-1} - 1) z_{ij})_1^n .$$

That is, $\hat{D}$ is a diagonal matrix acting on the n^2 dimensional space vector space $M_n(\mathbf{C})$. Thus, the n^2 eigenvalues of $\hat{D}$ are $(\lambda_i/\lambda_j) - 1$, $i, j = 1, \ldots, n$. This proves our claim in the case A is a diagonable matrix. The continuity argument implies the above result for any A. Hence, if I is an isolated attracting point we must have that all the eigenvalues of $\hat{A}$ do not exceed 1 in absolute value. By the definition of a hyperbolic attractive point we must have the strict inequalities (3.5). □

It is of interest to study the dynamics of the map Φ in general and in particular, the domain of the attraction of I under the conditions (3.5). Perhaps it could lead to a generalization of Jorgensen's inequality involving the spectrum of the matrices in question rather than their norms. Consider next the subgroup $O(J) \in GL_n(\mathbf{C})$. Obviously, $O(J)$ is an algebraic group. Furthermore, if $A \in O(J)$ then Φ: $O(J) \to O(J)$. In this case the $\Phi(X) = AXJ^{-1}A^T X^T J$ is a "quadratic" map.

THEOREM 3.6. *Let $A \in GL_n(\mathbf{C})$. Assume that $|\cdot|$ is a vector norm on $M_n(\mathbf{C})$ (not necessarily submultiplicative). Then*

$$\max_{1 \leqslant i, j \leqslant n} \left| \frac{\lambda_i}{\lambda_j} - 1 \right| \leqslant \|\hat{A}\| . \tag{3.7}$$

If A is diagonable, then there exists an operator norm on $M_n(\mathbf{C})$ such that the equality holds in the inequality (3.7). *If A is not diagonable then for any $\varepsilon > 0$ there exists an operator norm on $M_n(\mathbf{C})$ such that the right-hand side of* (3.7) *minus its left-hand side is less than ε.*

Proof. From the proof of Theorem 3.4 it follows that the eigenvalues of the operator $\hat{A}$ are $(\lambda_i/\lambda_j) - 1$, $i, j = 1, \ldots, n$. Recall that any operator norm of T: $V \to V$ bounds the spectral radius of T. This observation yields the inequality (3.7).

Assume that A is diagonable. W.l.o.g. (Without loss of generality) we may assume that $A = diag(\lambda_1, \ldots, \lambda_n)$. Thus, $\hat{A}$ is a diagonal matrix acting on the matrix $Z = (z_{ij})_1^n$. Recall that the l_1 and l_∞ norms of Z are given by the formula:

$$|(z_{ij})_1^n|_1 = \max_{1 \leqslant j \leqslant n} \sum_{i=1}^n |z_{ij}|, \qquad |(z_{ij})_1^n|_\infty = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^n |z_{ij}| . \tag{3.8}$$

A straightforward computation shows that for the above norms we have the equality in the equality (3.7). In the case A is not diagonable one proves the ε results using the continuity argument. □

Comparing Theorems 3.4 and 3.6 we see that for some choices of operator norm on $M_n(\mathbf{C})$ Theorem 1.5 is sharp.

Problem 3.9. Do we have equality in (3.7) for a diagonal matrix $A = diag(\lambda_1, \ldots, \lambda_n)$ if we choose the matrix operator norm to be the $|\cdot|_2$?

We now consider the problem when the iterations (0.3) stop after one or two iterations. We will assume first that $\mathscr{A}$ is an algebra over $\mathbf{C}$. Let $x \sim y \Leftrightarrow y = zxz^{-1}$, i.e., x and y are conjugate.

Definition 3.10. An invertible element $a \in \mathscr{A}$ is called regular if the following condition holds

$$x \sim a, \qquad [a, [a, x]] = 1 \Rightarrow [a, x] = 1.$$

Equivalently, for any regular a if $x \sim a$ and xax^{-1} commutes with a then x commutes with a. Note that

$$axax^{-1} = xax^{-1}a \Leftrightarrow x^{-1}axa = ax^{-1}ax \Leftrightarrow a^{-1}x^{-1}a^{-1}x = x^{-1}a^{-1}xa^{-1}.$$

Thus, a is regular if and only if a^{-1} is regular.

LEMMA 3.11. *Let $\mathscr{A}$ be an algebra and assume that $a, b \in \mathscr{A}$ are two invertible elements. Consider the iterations* (1.4). *Assume that $x_n = 1$ for some n. If a is regular then either $[a, b] = 1$, i.e., the iterations stop after one iteration at* 1, *or $[a, b] \neq 1$, $[a, [a, b]] = 1$, i.e., the iterations stop at* 1 *after two iterations.*

Proof. Clearly, if $[a, b] = 1$ then the iterations stop at 1 after one iteration. Assume that $[a, b] \neq 1$. Suppose that $x_{n-1} \neq 1, x_n = 1$. We claim that $n = 2$. Assume to the contrary that $n > 2$. As $1 = [a, x_{n-1}]$ we deduce that x_{n-1} commutes with a. Hence, $a^{-1}x_{n-1} = (x_{n-2}a^{-1})a^{-1}(x_{n-1}a^{-1})^{-1}$ commutes with a^{-1}. Furthermore, $x_{n-2}a^{-1} = (ax_{n-3})a^{-1}(ax_{n-3})^{-1}$. That is, $y = x_{n-2}a^{-1}$ is similar to a^{-1} and $ya^{-1}y^{-1}$ commutes with a^{-1}. Since a is regular we deduce that a^{-1} is regular. Hence, y commutes with a^{-1}. Therefore, x_{n-2} commutes with a. That is, $x_{n-1} = [a, x_{n-2}] = 1$ which contradicts our assumptions. □

THEOREM 3.12. *Let $A \in GL_n(\mathbf{C})$. Assume that A has pairwise distinct eigenvalues. Suppose furthermore that the eigenvalues of A do not have a set of k eigenvalues which are all k-th roots of unity times some complex number $c \in \mathbf{C}^*$, for some $k > 1$. Then A is regular.*

Proof. As A has pairwise distinct eigenvalues, A is similar to a diagonal matrix D with pairwise distinct diagonal entries. W.l.o.g. we may assume that $A = D$. Let $\mathscr{C}(D) \subset M_n(\mathbf{C})$ be the center of D. Next observe if $X \in \mathscr{C}(D)$ then X is a diagonal matrix. Suppose that $XDX^{-1} = D_1$ where D_1 is a diagonal matrix. As D and D_1 are similar it follows that the set of diagonal entries of D and D_1 are the same. Since D_1 is diagonal it follows that every column of X^{-1} is an eigenvector of D. As D is diagonal with pairwise distinct entries, every eigenvector of D is a multiple of a basis vector $(\delta_{1i}, \ldots, \delta_{ni})^T$. It then follows that X^{-1} is a monomial matrix. That is, $X = D_2P$ where D_2 is a diagonal matrix and P is a permutation matrix. We claim that $P = I$, i.e., X is diagonal. Indeed, for each cycle of P of length k we have that X has k eigenvalues which are all k-*th* roots of unity times c. As X is similar to A our assumptions on the spectrum of A yield that $k = 1$. Thus, X is diagonal, i.e., $X \in \mathscr{C}(D)$. □

LEMMA 3.13. *Let $A \in GL_2(\mathbf{C})$, $trace(A) \neq 0$. Then A is regular.*

Proof. If $A = \lambda I$ then $\mathscr{C}(A) = M_2(\mathbf{C})$, hence A is regular. If A has two pairwise distinct eigenvalues then A is regular by Theorem 3.12 unless the sum of the two eigenvalues is zero, i.e., $trace(A) = 0$. Assume that the minimal polynomial of A is $(x - \lambda)^2$. That is, A is similar to an upper triangular matrix U with λ on the diagonal such that $U - \lambda I \neq 0$. Then any matrix in $V \in \mathscr{C}(U)$ which is similar to U is of the form $V = \lambda I + a(U - \lambda I)$, $a \neq 0$. The assumption that $XUX^{-1} = V$ implies that V is upper triangular. If X is similar to U it follows that the two diagonal entries of X are equal to λ. Hence, $X \in \mathscr{C}(U)$. That is, U is regular. □

Note that the matrices satisfying the conditions of Lemma 3.13 correspond to Möbius transformations which are not rotations. We remark that if the minimal polynomial of $A \in GL_3(\mathbf{C})$ is $(x - 1)^3$ then according to our computation A is not regular. It is of interest to characterize regular matrices in $GL_n(\mathbf{C})$ for $n > 2$. We conclude this section with the following conjecture.

Conjecture 3.14. *Let n be a positive integer. Then there exists an integer $\alpha(n)$ so that the following condition holds. Assume that $A, B \in GL_n(\mathbf{C})$ and consider the iterations* (0.3). *Then either $X_{\alpha(n)} = 1$ or $X_m \neq I$, $m = 0, 1, \ldots$.*

4. Balls in hyperbolic manifolds. In this section we estimate from below r_n—the largest round ball in a hyperbolic n-manifold following [Mar2]. Unfortunately, there is an error in Corollary 3.3 of [Mar2] which invalidates Martin's inequality for r_n. In what follows we modify some of the lemmas of §3 in [Mar2] incorporating our results in the previous sections to obtain a lower bound on r_n. We adopt the notation of [Mar2] unless stated otherwise.

Let $\mathbf{H}^n$ be the hyperbolic n-space. We are going to use here two standard models of $\mathbf{H}^n$. First, identify $\mathbf{H}^n$ with the open Euclidean unit ball $B^n \subset \mathbf{R}^n$ with the Riemannian metric

$$ds^2 = \frac{4}{(1 - |x|^2)^2}|dx|^2.$$

For $x, y \in B^n$ let $dist(x, y)$ be the hyperbolic distance between x and y. Recall that

$$dist(0, x) = ln\frac{1 + |x|}{1 - |x|}, \qquad x \in B^n.$$

Let $\mathscr{M}_n$ be the group of hyperbolic isometries of B^n. The second model of $\mathbf{H}^n$ is the hyperboloid model $K \subset \mathbf{R}^{n+1}$. See for example [Bea]. Then $f \in \mathscr{M}_n$ is represented by $A \in O(1, n)$. Recall that any n-hyperbolic manifold is B^n/Γ where $\Gamma \subset \mathcal{M}_n$ is a discrete torsion free subgroup.

LEMMA 4.1. *Let $A \in O(n)$. Then for each $Q \geqslant 4$ there is $B \in O(n)$ such that for some $1 \leqslant q \leqslant Q^{\lfloor n/2 \rfloor}$, $B^{2q} = I$ and*

$$\|A - B\| \leqslant \frac{\pi}{qQ}.$$

Proof. Recall that $O(n)$ is $SO(1) \oplus SO(n)$ where $SO(m)$ is the group of $m \times m$ real orthogonal matrices with the determinant equal to one. Thus, w.l.o.g. we may assume that $A, B \in SO(n)$. Recall that the eigenvalues of A are either equal to ± 1 or come in pairs $e^{\sqrt{-1}\pi\theta}, e^{-\sqrt{-1}\pi\theta}$, $0 < \theta < 1$. Let $\theta_1, \ldots, \theta_m \in (0, 1)$, $m \leqslant \lfloor n/2 \rfloor$ be the θ's corresponding to the nonreal eigenvalues of A. As $A \in SO(n)$ it follows that A must have $2k$ eigenvalues equal to -1. That is, there exists an orthogonal decomposition

$$\mathbf{R}^n = \left(\bigoplus_1^{m+k} U_i\right) \oplus \left(\bigoplus_1^{m+k} U_i\right)^{\perp}$$

where each U_i is an invariant subspace of A of dimension 2 such that:

(i) the restriction of A to U_i is a rotation by $\varepsilon_i \pi \theta_i$, $\varepsilon_i = \pm 1$, $i = 1, \ldots, m$;

(ii) the restriction of A to U_i is a rotation by π for $i = m+1, \ldots, m+k$;

(iii) the restriction of A to $(\bigoplus_1^{m+k} U_i)^{\perp}$ is the identity.

Use the Dirichlet theorem to deduce the existence of $1 \leqslant q \leqslant Q^m$ and $p_1, \ldots, p_m \in \mathbf{Z}$ so that

$$\left|\theta_i - \frac{p_i}{q}\right| \leqslant \frac{1}{qQ}, \qquad i = 1, \ldots, m.$$

Let $B \in SO(n)$ be obtained from A as follows. Assume that $U_1, \ldots, U_{m+k}$ and $(\bigoplus_1^{m+k} U_i)^{\perp}$ are invariant subspaces of B. Furthermore, the restriction of B to U_i is a rotation by $\varepsilon_i \pi p_i/q$, $i = 1, \ldots, m$. On the invariant subspaces of $U_{m+1}, \ldots, U_{m+k}$, $(\bigoplus_1^{m+k} U_i)^{\perp}$ the action of B coincides with the action of A. Hence, $B^{2q} = I$ and $\|A - B\| \leqslant \pi/qQ$. □

Remark. The above lemma is the correct version of Corollary 3.3 of [Mar2]. Indeed, Corollary 3.3 claims that $B^q = I$. This is obviously false if $A \in O(n)$ has $2k \geqslant 2$ eigenvalues equal to -1 and $n - 2k \geqslant 1$ eigenvalues equal to 1. (In that case $B = A$.)

LEMMA 4.2. *Suppose that f is a hyperbolic isometry and that*

$$\frac{1 + |f(0)|}{1 - |f(0)|} = r.$$

Let $A \in O(1, n)$ be the matrix corresponding to f and $B \in O(n)$. Then for each $q \geqslant 1$

$$\|A^q - B^q\| \leqslant \frac{r^q - 1}{r - 1}\|A - B\|.$$

Proof. It is known, e.g., Proof of Lemma 3.1 in [Mar2], that $\|A\| = r$. As $B \in O(n)$ we get $\|B^i\| = 1$. Use the identity

$$A^q - B^q = (A - B)A^{q-1} + B(A - B)A^{q-2} + \cdots + B^{q-2}(A - B)A + B^{q-1}(A - B),$$

the triangle inequality, the inequality $\|A^i\| \leqslant \|A\|^i$ and the above inequalities to deduce the lemma. □

THEOREM 4.3. *Let $\Gamma \subset \mathcal{M}_n$ be a discrete torsion free nonelementary group. Then there exists $o \in B^n$ so that for any $f \in \Gamma$*

$$dist(o, f(o)) \geqslant \delta, \qquad \delta = \frac{0.005}{17^{\lfloor n/2 \rfloor}}. \tag{4.4}$$

Proof. By abuse of notation we view $\Gamma \subset O(1, n)$. Use the arguments of Theorem 2.4 in [Mar2] and Theorem 2.16 to deduce the existence of $\alpha \in O(1, n)$ so that for any $I \neq C \in \alpha\Gamma\alpha^{-1}$ we have the inequality

$$\|C\|\,\|C - I\| \geqslant \omega > 0.3854. \tag{4.5}$$

We claim that (4.4) holds with $o = \alpha^{-1}(0)$. Assume to the contrary that (4.4) does not hold for some $f \in \Gamma$. Then f is represented by $\tilde{A} \in O(1, n)$. Set $A = \alpha\tilde{A}\alpha^{-1}$, $r = \|A\|$. As in the proof of Lemma 3.1 in [Mar2] we deduce that $r < e^{\delta}$. Furthermore, there exists $O \in O(1, n)$ so that $\|A - O\| \leqslant r(r - 1)$. Apply Lemma 4.1 to O with $Q = 17$. We then deduce the existence of an elliptic $B \in O(1, n)$ of order $2q$ so that

$$\|O - B\| \leqslant \frac{\pi}{qQ}, \qquad \|O^{2q} - B^{2q}\| = \|O^{2q} - I\| \leqslant \frac{2\pi}{Q}, \qquad 1 \leqslant q \leqslant Q^{\lfloor n/2 \rfloor}.$$

Use Lemma 4.2 to deduce

$$\|A^{2q} - O^{2q}\| \leqslant (r^{2q} - 1)r.$$

Use the triangle inequality and the fact that $O \in O(n)$ and the existence of B to deduce

$$\|A^{2q}\| \leqslant 1 + (r^{2q} - 1)r, \qquad \|A^{2q} - I\| \leqslant \frac{2\pi}{Q} + (r^{2q} - 1)r.$$

Thus

$$\|A^{2q}\|\,\|A^{2q}-I\| \leqslant (1+(r^{2q}-1)r)\left(\frac{2\pi}{Q}+(r^{2q}-1)r\right)$$

$$< (1+(e^{2Q\lfloor n/2\rfloor\delta}-1)e^{\delta})\left(\frac{2\pi}{Q}+(e^{2Q\lfloor n/2\rfloor\delta}-1)e^{\delta}\right).$$

Use the assumption that $Q = 17$, the value of δ and the inequality $n \geqslant 2$ to obtain

$$\|A^{2q}\|\,\|A^{2q}-I\| < \left(\frac{2\pi}{17}+(e^{0.01}-1)e^{0.005/17}\right)(1+(e^{0.01}-1)e^{0.005/17}) < 0.384.$$

The above inequality contradicts (4.5). The proof of the theorem is completed. □

THEOREM 4.6. *Let $M = B^n/\Gamma$ be a hyperbolic n-manifold. Then any fundamental domain corresponding to M contains a hyperbolic ball of radius*

$$r = \frac{0.0025}{17^{\lfloor n/2\rfloor}}.$$

5. The symplectic group. In this section we extend Shimuzu-Leutbecher and Jorgensen inequalities to certain discrete nonelementary subgroups $\langle A, B\rangle$ of the symplectic group

$$Sp(n,\mathbf{R}) = O(J)\cap GL_{2n}(\mathbf{R}) = U(J)\cap GL_{2n}(\mathbf{R}),$$
$$J = (J_{ij})_1^2,\quad J_{11}=J_{22}=0,\quad J_{12}=-J_{21}=I. \tag{5.1}$$

We follow closely some ideas and arguments of the second named author in [Her1]. In [Her1] one expresses Möbius transformations in $\hat{\mathbb{R}}^n$ as 2×2 matrices whose entries are Clifford numbers, i.e., Vahlen matrices. Here we use the natural partition of symplectic matrices as 2×2 block matrices which satisfy special identities stated below.

The following lemma follows straightforward from the definition of $U(J)$.

LEMMA 5.2. *Let $A = (A_{ij})_1^2 \in U(J)$, $A_{ij}\in M_n(\mathbf{C})$, $i,j = 1,2$. Then*

$$A^{-1} = ((A^{-1})_{ij})_1^2,\qquad (A^{-1})_{11} = A_{22}^*,\qquad (A^{-1})_{22} = A_{11}^*,$$
$$(A^{-1})_{12} = -A_{12}^*,\qquad (A^{-1})_{21} = -A_{21}^*,$$
$$A_{11}A_{22}^* - A_{12}A_{21}^* = A_{22}A_{11}^* - A_{21}A_{12}^* = A_{11}^*A_{22} - A_{21}^*A_{12}$$
$$= A_{22}^*A_{11} - A_{12}^*A_{21} = I, \tag{5.3}$$

$$(A_{11}A_{12}^*)^* = A_{11}A_{12}^*, \qquad (A_{21}A_{22}^*)^* = A_{21}A_{22}^*,$$
$$(A_{11}^*A_{21})^* = A_{11}^*A_{21}, \qquad (A_{22}^*A_{12})^* = A_{22}^*A_{12}. \tag{5.3}$$

$A \in U(J)$ is called a translation

$$A = (A_{ij})_1^2, \qquad A_{11} = A_{12} = A_{22} = I, \qquad A_{21} = 0.$$

For $C \in M_n(\mathbf{C})$ let $|C| = \rho(C^*C)^{1/2}$ be the spectral norm of C. The following lemma can be considered as a generalization of Shimizu-Leutbecher inequality. See [Her1, Theorem A].

THEOREM 5.4. *Let $A = (A_{ij})_1^2$, $B = (B_{ij})_1^2 \in U(J)$. Assume that A is a translation and $B_{21} \neq 0$. If $\langle A, B\rangle$ is a discrete group then $|B_{21}| \geqslant 1$.*

Proof. Let

$$\Psi(X) = XAX^{-1}, \quad X \in GL_{2n}(\mathbf{C}), \quad X_m = (X_{m,ij})_1^2 = \Psi^{\circ m}(B), \quad m = 0, 1, \ldots. \tag{5.5}$$

Set:

$$X_{m,11} = a_m, X_{m,12} = b_m, X_{m,21} = c_m, X_{m,22} = d_m;\ a_m, b_m, c_m, d_m \in M_n(\mathbf{C}). \tag{5.6}$$

Use (5.3) to deduce

$$\begin{aligned} a_{m+1} &= a_m d_m^* - (a_m + b_m)c_m^* = 1 - a_m c_m^*, \\ b_{m+1} &= -a_m b_m^* + (a_m + b_m)a_m^* = a_m a_m^*, \\ c_{m+1} &= c_m d_m^* - (c_m + d_m)c_m^* = -c_m c_m^*, \\ d_{m+1} &= -c_m b_m^* + (c_m + d_m)a_m^* = 1 + c_m a_m^*. \end{aligned} \tag{5.7}$$

In particular,

$$c_{m+1} = -(c_0 c_0^*)^{2^m}, \qquad |c_{m+1}| = |(c_0 c_0^*)|^{2^m} = |c_0|^{2^{m+1}}, \qquad m = 0, \ldots. \tag{5.8}$$

Assume to the contrary that $|c_0| < 1$. Then $\lim_{m\to\infty} c_m = 0$. We claim that $\lim_{m\to\infty} X_m = A$. First note that from the first identity of (5.7) we get the inequality

$$|a_{m+1}| \leqslant 1 + |a_m||c_m^*| = 1 + |a_m||c_0|^{2^m}.$$

Use the induction and the assumption $|c_0| < 1$ to deduce that

$$|a_{m+1}| \leqslant |a_0||c_0|^{m+1} + \sum_{i=0}^{m} |c_0|^i.$$

In particular the sequence $\{|a_m|\}_0^\infty$ is bounded by some positive constant K. Hence

$$|a_{m+1} - 1| \leqslant |a_m|\,|c_m^*| \leqslant K|c_0|^{2^m} \Rightarrow \lim_{m\to\infty} a_m = 1.$$

The other equalities of (5.7) yield that $\lim_{m\to\infty} X_m = A$. Finally, the assumption that $c_0 = B_{21} \neq 0$ implies that $c_m \neq 0$, $m = 1, \ldots,$. That is $X_m \neq A$, $m = 0, 1, \ldots$ which contradicts the assumption that $\langle A, B\rangle$ generates a discrete group. □

$A \in Sp(n, \mathbf{R})$ is called superhyperbolic if

$$A = Q(\tau I_n \oplus \tau^{-1} I_n)Q^{-1}, \qquad Q \in Sp(n, \mathbf{R}), \qquad 0 < \tau < 1. \tag{5.9}$$

In order to give a generalization of Jorgensen's inequality we need to view $GL_{2n}(\mathbf{C})$ (more precisely $PGL_{2n}(\mathbf{C})$) as a group of Möbius transformations acting on the Grassmanian manifold $\mathscr{G}_{2n,n}$. Recall that $\mathscr{G}_{2n,n}$ is the projective variety of all n dimensional subspaces of $\mathbf{C}^{2n}$. As usual let $M_{p,q}(\mathbf{C}) \approx Hom(\mathbf{C}^q, \mathbf{C}^p)$ be the vector space of all $p \times q$ complex valued matrices. Then a point $P \in \mathscr{G}_{2n,n}$ is represented by a matrix $P \in M_{2n,n}$ of rank n, i.e. $rank(P) = n$. That is, the corresponding n dimensional subspace of $\mathbf{C}^{2n}$ is the subspace spanned by the n columns of P. Thus, $P' \in M_{2n,n}$ represent the same point in $\mathscr{G}_{2n,n}$ if and only if $P' = PT$, $T \in GL_n(\mathbf{C})$. We adopt the following block notation for P

$$P = [P_1, P_2] \Leftrightarrow P = (P_{ij})_{i,j=1}^{i=2,j=1}, \qquad P_{11} = P_1, \qquad P_{21} = P_2 \in M_n(\mathbf{C}).$$

Then $GL_{2n}(\mathbf{C})$ acts naturally on $\mathscr{G}_{2n,n}$ by multiplication.

$$A\colon \mathscr{G}_{2n,n} \to \mathscr{G}_{2n,n}, \qquad P \mapsto AP, \qquad P \in \mathscr{G}_{2n,n}, \qquad A \in GL_{2n}(\mathbf{C}).$$

It now follows that A and αA, $\alpha \in \mathbf{C}^*$ represent the same transformation on $\mathscr{G}_{2n,n}$. We now show that the above action corresponds to the standard representation of Möbius maps. Let

$$\mathscr{G}'_{2n,n} = \{P\colon P = [X, I],\ X \in M_n(\mathbf{C})\}.$$

Here I is the $n \times n$ identity matrix. Thus, $\mathscr{G}_{2n,n} \supset \mathscr{G}'_{2n,n} \sim M_n(\mathbf{C})$ is one of the affine charts of $\mathscr{G}_{2n,n}$. Assume that $A = (A_{ij})_1^2 \in GL_{2n}(\mathbf{C})$. Then $A[X, I] = [A_{11}X + A_{12}, A_{21}X + A_{22}]$. Thus, if $A_{21}X + A_{21}$ is nonsingular we get $[X, I] \mapsto [(A_{11}X + A_{12})(A_{21}X + A_{22})^{-1}, I]$. This is exactly the Möbius transformation on the first coordinate of $P = [X, I]$.

We now discuss briefly the fixed points of $A \in GL_{2n}(\mathbf{C})$ using the standard tools of multilinear algebra. Consult for example with [Mar]. First note that $\mathscr{G}_{2n,n}$ is the *n-th* wedge product $\bigwedge_1^n \mathbf{C}^{2n}$. Then the action of A on $\bigwedge_1^n \mathbf{C}^{2n}$ is given by $C_n(A) = A \wedge A \wedge \cdots \wedge A$ the *n-th* compound of A. Recall that $C_n(A)$ is $N \times N$ matrix with

$N = \binom{2n}{n}$, where each entry of $C_n(A)$ is a corresponding $n \times n$ minor of A. Thus, any fixed point $P \in \mathscr{G}_{2n,n}$ of A corresponds to an eigenvector of $C_n(A)$ which represents an n-dimensional invariant subspace of A. Recall

$$\Lambda(C_n(A)) = \{\lambda, \lambda = \lambda_{i_1} \dots \lambda_{i_n}, 1 \leqslant i_1 < \cdots < i_n \leqslant 2n\}, \qquad \Lambda(A) = \{\lambda_1, \dots, \lambda_{2n}\}.$$

Thus, P is an isolated fixed point of A if and only if P is an eigenvector of $C_n(A)$ corresponding to a simple eigenvalue. Let A be of the form (5.9). It now follows that $C_n(A)$ has exactly two simple eigenvalues τ^n, τ^{-n}. That is, the corresponding Möbius transformation $A: \mathscr{G}_{2n,n} \to \mathscr{G}_{2n,n}$ has exactly two fixed isolated points. For $n > 1$, A has other fixed points which form irreducible varieties of positive complex dimension.

Let $A, B \in GL_{2n}(\mathbf{C})$ and denote by $\langle A, B\rangle$ the group of Möbius transformations generated by A and B. We call $\langle A, B\rangle$ an elementary group if $\langle A, B\rangle$ has an invariant orbit consisting of a finite number of points. See for example [Bea]. The following theorem is a generalization of Jorgensen's inequality. See [Her1, Theorem B].

THEOREM 5.10. *Let $A, B \in Sp(n, \mathbf{R})$. Assume that A is superhyperbolic with an eigenvalue $0 < \tau < 1$. If $\langle A, B\rangle$ is a discrete nonelementary group then*

$$\left(\tau - \frac{1}{\tau}\right)^2 + \max_{\nu \in \Lambda(BAB^{-1}A^{-1})} \left|\nu + \frac{1}{\nu} - 2\right| \geqslant 1. \tag{5.11}$$

Proof. Suppose to the contrary (5.11) does not hold

$$\left(\tau - \frac{1}{\tau}\right)^2 + \max_{\nu \in \Lambda(BAB^{-1}A^{-1})} \left|\nu + \frac{1}{\nu} - 2\right| < 1. \tag{5.11$'$}$$

W.l.o.g. we assume that

$$\begin{gathered} A = (A_{ij})_1^2, \quad A_{11} = \tau I, \quad A_{22} = \tau^{-1} I, \quad A_{12} = A_{21} = 0, \\ B_{m+1} = B_m A B_m^{-1} = (B_{m+1,ij})_1^2, \quad B_0 = B, \quad m = 1, \dots, \\ B_{m,11} = a_m, \quad B_{m,12} = b_m, \quad B_{m,21} = c_m, \quad B_{m,22} = d_m, \quad m = 0, 1, \dots \end{gathered} \tag{5.12}$$

Recall that the matrices B_m, $m = 0, 1, \dots$, are real. Thus, $B_{m,ij}^* = B_{m,ij}^T$. From the identities (5.3) and we obtain the recursive relations:

$$\begin{aligned} a_{m+1} &= \tau a_m d_m^* - \tau^{-1} b_m c_m^*, \qquad & b_{m+1} &= b_m a_m^*(\tau^{-1} - \tau), \\ c_{m+1} &= c_m d_m^*(\tau - \tau^{-1}), \qquad & d_{m+1} &= \tau^{-1} d_m a_m^* - \tau c_m b_m^*. \end{aligned} \tag{5.13}$$

From the above identities one deduces that the matrices b_m, c_m, $m = 1, \dots,$ are real symmetric matrices. Combine the above equalities with (5.3) to obtain

$$b_{m+1}c_{m+1}^* = -(1 + b_m c_m^*)b_m c_m^*(\tau - \tau^{-1})^2. \tag{5.14}$$

Let

$$f: [0, r] \to [0, r], \quad f(x) = x(1 + x)(\tau - \tau^{-1})^2, \quad r = (\tau - \tau^{-1})^{-2} - 1 > 0. \tag{5.15}$$

(Note that (5.11)′ implies the inequality $(\tau - \tau^{-1})^2 < 1$, i.e. $r > 0$.) The two fixed points of f are 0, r. Furthermore, $f(x)$ is an increasing function and $f(x) < x$ on $(0, r)$. We claim (5.11)′ yields that $\rho(b_0 c_0^*) < r$. Assume for a moment that this is correct. It is well known that one can introduce an operator norm on $M_n(\mathbf{C})$ so that $|b_0 c_0^*| < \rho(b_0 c_0^*) + \varepsilon$ for any positive ε. (See, for example, [H-J, 5.6.10].) Choose a positive ε so that $|b_0 c_0^*| < r$. The submultiplicativity of an operator norm and (5.14) yield

$$|b_{m+1}c_{m+1}^*| \leqslant f(|b_m c_m^*|) \leqslant \cdots \leqslant f^{\circ(m+1)}(|b_0 c_0^*|).$$

Hence, $\lim_{m\to\infty} b_m c_m^* = 0$. Use (5.3) to obtain $a_m d_m^* = I + b_m c_m^*$. Thus, $\lim_{m\to\infty} a_m d_m^* = I$. As $(a_m d_m^*)^* = d_m a_m^*$ we also deduce that $\lim_{m\to\infty} d_m a_m^* = I$. The equalities (5.13) and the limits that we showed to exist yield that

$$\lim_{m\to\infty} a_m = \tau I, \qquad \lim_{m\to\infty} d_m = \tau^{-1} I. \tag{5.16}$$

The inequality (5.11)′ yields $|\tau - \tau^{-1}| < 1$. Use (5.16) and (5.13) to obtain

$$|b_{m+1}\tau^{-(m+1)}|_2 \leqslant c|b_m \tau^{-m}|_2, \qquad |c_{m+1}\tau^{m+1}|_2 \leqslant c|c_m \tau^m|_2, \qquad m > M,$$

for some $0 < c < 1$. Hence

$$\lim_{m\to\infty} b_m \tau^{-m} = \lim_{m\to\infty} c_m \tau^m = 0.$$

It then follows that $\lim_{m\to\infty} A^{-m}B_{2m}A^m = A$. As $\langle A, B\rangle$ is discrete we deduce that $B_{2m} = A$ for some m. We claim that $\langle A, B\rangle$ is elementary. Indeed, as $B_{2m-1}A = AB_{2m-1}$ it then follows that the set $\{P, Q\}$ consisting of the two isolated fixed points of A remains invariant under the action of B_{2m-1}. Continue this argument inductively to deduce that $\{P, Q\}$ is invariant under the action of B_i, $i = 2m - 1, \dots, 0$. Hence, A and B have an invariant finite orbit $\{P, Q\}$ which contradicts that $\langle A, B\rangle$ is nonelementary. To end the proof of the theorem we need to show that (5.11)′ yields the inequality $\rho(b_0 c_0^*) < r$. More precisely, assume that

$$\left(\tau - \frac{1}{\tau}\right)^2 + \max_{\nu \in \Lambda(BAB^{-1}A^{-1})} \left|\nu + \frac{1}{\nu} - 2\right| < 1 - \varepsilon, \qquad 0 < \varepsilon. \tag{5.11″}$$

We will then show that

$$(1 + \rho(b_0 c_0^*))(\tau - \tau^{-1})^2 \leqslant 1 - \varepsilon. \tag{5.17}$$

Consider the matrix B_1. As B_1 is similar to A it follows that

$$(B_1 - \tau I)(B_1 - \tau^{-1} I) = 0. \tag{5.18}$$

In particular:

$$\begin{aligned} a_1^2 - (\tau + \tau^{-1})a_1 + I &= -b_1 c_1 = -b_1 c_1^*, \\ d_1^2 - (\tau + \tau^{-1})d_1 + I &= -c_1 b_1 = -(b_1 c_1^*)^*. \end{aligned} \tag{5.19}$$

Assume first all the eigenvalues of $b_1 c_1$ which are equal to the eigenvalues of $c_1 b_1$ are pairwise distinct and are different from zero. It then follows that if x is a nontrivial eigenvalue of $b_1 c_1$, that is $b_1 c_1 x = \gamma x$, $x \neq 0$ then $a_1 x = \alpha x$. Note that $c_1 b_1 (c_1 x) = \gamma (c_1 x)$. As $b_1 c_1$ is a nonsingular matrix we obtain that $c_1 x \neq 0$. Deduce from (5.19) that $d_1(c_1 x) = \delta(c_1 x)$. Let

$$y_1 = (x^T, 0)^T, \qquad y_2 = (0, (c_1 x)^T)^T.$$

It now follows that $span\{y_1, y_2\}$ is an invariant subspace of B_1 and B_1 is represented by 2×2 matrix with the entries α, 1; γ, δ. Note that

$$\alpha + \delta = \tau + \tau^{-1}, \qquad \alpha\delta - \gamma = 1. \tag{5.20}$$

Observe next that $span\{y_1, y_2\}$ is also an invariant space of $B_1 A^{-1}$ with representation matrix $\alpha\tau^{-1}$, τ^{-1}; $\gamma\tau$, $\delta\tau$. Thus, if $\nu \in Spec(BAB^{-1}A^{-1})$ it follows that

$$\nu + \nu^{-1} = \alpha\tau^{-1} + \delta\tau. \tag{5.21}$$

Use the equality (5.14) for $m = 0$, the fact that $c_1^* = c_1$ and the assumption that the eigenvalues of $b_1 c_1 = b_1 c_1^*$ are pairwise distinct to deduce that x is also an eigenvector of $b_0 c_0^*$: $b_0 c_0^* x = \gamma_0 x$. We claim that $\alpha = \tau + (\tau - \tau^{-1})\gamma_0$. Indeed, the first equality of (5.13) for $m = 0$ reads:

$$a_1 = \tau a_0 d_0^* - \tau^{-1} b_0 c_0^* = \tau b_0 c_0^* + \tau I - \tau^{-1} b_0 c_0^* \Rightarrow a_1 x = (\tau + (\tau - \tau^{-1})\gamma_0)x.$$

The equality $\alpha + \delta = \tau + \tau^{-1}$ yields that $\delta = \tau^{-1} - (\tau - \tau^{-1})\gamma_0$. Hence

$$\nu + \nu^{-1} - 2 = \alpha\tau^{-1} + \delta\tau = -\gamma_0(\tau - \tau^{-1})^2.$$

We thus showed

$$\max_{\nu \in \Lambda(BAB^{-1}A^{-1})} |\nu + \nu^{-1} - 2| = \rho(b_0 c_0^*)(\tau - \tau^{-1})^2.$$

Hence, (5.11)″ implies (5.17). In particular, $\rho(b_0c_0^*) < r$ where r is defined in (5.15). It is left to show that (5.11)″ yields (5.17) for any $b_0c_0^*$. Clearly, we can approximate the matrix B by $\tilde{B}$ so that (5.11)″ holds for A and $\tilde{B}$. Let the corresponding blocks of $\tilde{B}$ be $\tilde{a}_0$, $\tilde{b}_0$, $\tilde{c}_0$, $\tilde{d}_0$. We can assume that $\tilde{b}_0\tilde{c}_0^*$ have pairwise distinct eigenvalues such that $(1 + \tilde{b}_0\tilde{c}_0^*)\tilde{b}_0\tilde{c}_0^*$ have also pairwise distinct eigenvalues different from zero. As A, $\tilde{B}$ satisfy the inequality (5.11)″ our arguments show that (5.17) holds for $\tilde{b}_0\tilde{c}_0^*$. Now the continuity argument shows the inequality (5.11)″ implies (5.17) for any b_0, c_0. The proof of the theorem is completed. □

Acknowledgement. The second named author is a graduate student at the University of Illinois at Chicago. He thanks Professor Peter Shalen, his supervisor, for his advice and support.

References

[Bea] A. F. Beardon, *The Geometry of Discrete Groups*, Graduate Texts in Math **91**, Springer-Verlag, New York, 1983.

[Fri] S. Friedland, *Nonoscillation, disconjugacy and integral inequalities*, Mem. Amer. Math. Soc. 7 (1976), #176.

[Her1] S. Hersonksy, *A generalization of Shimizu-Leutbecher lemma and the Jorgenson inequality to Möbius transformations in* $\hat{R}^n$, to appear in Proc. Amer. Math. Soc.

[Her2] ———, *Covolume estimates for discrete hyperbolic isometries having parabolic elements*, to appear in Michigan Math J.

[HJ] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, Massachusetts, 1985.

[Hup] B. Huppert, Endliche Gruppen I, Springer-Verlag, New York, 1967.

[Jor] T. Jorgensen, *On discrete groups of Möbius transformations*, Amer. J. Math. **98** (1976), 739–749.

[Mar] M. Marcus, *Finite Dimensional Multilinear Algebra I, II*, Pure Appl. Math. **23**, Marcel Decker, New York, 1973, 1975.

[Mar1] G J. Martin, *On discrete Möbius groups in all dimensions: A generalization of Jorgensen's inequality*, Acta Math. **163** (1989), 253–289.

[Mar2] ———, *Balls in hyperbolic manifolds*, J. London Math. Soc. **40** (1989), 257–265.

Department of Mathematics, University of Illinois at Chicago, Box 4348, Chicago, Illinois 60680-4348, USA