Elementary Groups and Jörgensen's Inequality

Juan Miguel Ogarrio

November 13, 2009

1 Elementary Groups

Definition 1 *A subgroup* Γ *of PSL(2,**R**) is called* elementary *if there exists some* x *such that* Γx*, the* Γ *orbit of* x*, is finite.*

Clearly, any subgroup in which all elements share some fixed point is elementary.

We've noted how an element of PSL(2,**R**) takes elements from **H** to **H**. However, it follows directly from the closure of the real numbers that PSL(2,**R**) takes elements from **R** to **R**. If we look at the limiting behavior when dealing with ∞, we can say that PSL(2,**R**) takes elements from $\mathbf{R}\bigcup\{\infty\}$ to $\mathbf{R}\bigcup\{\infty\}$.

In other words, orbits stay in $\mathbf{R}\bigcup\{\infty\}$ or in **H**, though in either case the group acts on the closure of **H**.

Definition 2 *If* g, h *are elements of a group, then the* commutator *is the element* $ghg^{-1}h^{-1}$*, denoted by* $[g, h]$.

Clearly, if g and h commute, the commutator is just the identity. We will use the commutator as a tool to prove various theorems, and it is a critical concept in Jörgensen's inequality.

1.1 Characterization of Elementary Groups

Lemma 1 *Let* G *be a subgroup of PSL(2,**R**) containing besides the identity only elliptic elements. Then all elements of* G *have the same fixed point, and the subgroup is abelian, cyclic, and elementary.*

We recall that we can always conjugate all elements of G so that some element in the subgroup "behaves nicely". So if we look at the unit disc model, where elements of PSL(2,**R**) take the form $\begin{bmatrix} a & b \\ \bar{b} & \bar{a} \end{bmatrix}$, we can conjugate G so that $g \epsilon G$ is expressed by the matrix $\begin{bmatrix} u & 0 \\ 0 & \bar{u} \end{bmatrix}$ (i.e. it fixes the origin). Then take any other $h \epsilon G$ so that $h = \begin{bmatrix} a & c \\ \bar{c} & \bar{a} \end{bmatrix}$.

We want to show that g and h have the same fixed points, so two important assumptions are that $g \neq Id$ and $g \neq h$.

We first find the commutator matrix to be $[g,h] = \begin{bmatrix} |a|^2 - u^2|c|^2 & -ac + acu^2 \\ -\bar{a}\bar{c} + \bar{a}\bar{c}\bar{u}^2 & |a|^2 - \bar{u}^2|c|^2 \end{bmatrix}$, so that $tr[g,h] = 2|a|^2 - |c|^2(u^2 + \bar{u}2) = 2|a|^2 - |c|^2(-4Im(u)^2 + 2) = 2(|a|^2 - |c|^2) + |c|^2 4Im(u)^2 = 2 + |c|^2 4Im(u)^2$.

Since g contains no hyperbolic elements, this value can be no greater than 2, which means that either $|c| = 0$ or $Im(u) = 0$. In the latter case, g is just the identity, which contradicts our assumptions, so $|c| = 0$, which entails that h also fixes the origin. Since this works for arbitrary h, then all elements in G share g's fixed point, and so all elements have the same fixed points, which we know entails G is abelian and cyclic, and since one point has an orbit of order 1, then G is elementary.

We can now state a theorem which describes all elementary Fuchsian groups.

Theorem 1 *Any elementary Fuchsian group is either cyclic or conjugate in PSL(2,**R**) to a group generated by* $g(z) = kz (k > 1), h(z) = -1/z$.

Consider an elementary Fuchsian group Γ. We prove this theorem considering three possible cases: Γ has an orbit of order 1, Γ has an orbit of order 2 in $\mathbf{R} \bigcup \{\infty\}$, or any other situation.

Case 1 This is equivalent to saying that all elements of Γ share 1 fixed point in the closure of **H**. If this point is in **H**, then all elements in Γ are elliptic (as parabolic and hyperbolic elements have no fixed points in **H**), so our lemma shows us Γ is cyclic.

If the fixed point is in $\mathbf{R} \bigcup \{\infty\}$, then all elements are either parabolic or hyperbolic. We wish to show they must all be of the same type.

Assume the contrary, and conjugate Γ so that the fixed point is ∞. Then if we pick a hyperbolic element g and a parabolic element h, $g(z) = \lambda z$, $h(z) = z + k$ for $\lambda > 1, k \neq 0$. If we couldn't choose such a k, there would be no non-identity parabolic elements, and for the same reason we can choose a non-zero λ, and take the inverse if necessary to satisfy the inequality condition. Then the element $g^{-n}hg^{n}(z) = z + \lambda^{-n}k$.

Then $\|g^{-n}hg^n\| = (2^2+(\lambda^{-n}k)2)^{1/2}$, which is bounded, since $\lambda > 1$. Then we can extract a converging subsequence of distinct terms which contradicts discreteness. So there can only be parabolic or hyperbolic elements.

If there are only parabolic elements, then all elements of Γ have the same fixed points and so Γ is cyclic. If all elements are hyperbolic, we wish to show that the second fixed point is also shared by all of them, which once again would imply Γ is cyclic.

Conjugate so that some non-identity element f is represented by the matrix $\begin{bmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{bmatrix}$, which fixes 0 and ∞, and some other parabolic element g is represented by $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If we want g to fix only 0, then $b = 0, c \neq 0, a \neq 0$, and $d = 1/a$, so that $[f, g] = \begin{bmatrix} 1 & 0 \\ c/a(1/\lambda^2 - 1) & 1 \end{bmatrix}$ We know $c \neq 0$, otherwise g fixes ∞, and $\lambda \neq 1$, otherwise f is the identity, both of which contradict our assumptions. But then if we let $t = c/a(1/\lambda^2-1)$, g has the form $g(z) = \frac{z}{tz+1}$, which has a fixed point only when $tz = 0$. But since $t \neq 0$, this has only one fixed point, and thus is parabollic, a contradiction. So all hyperbolic elements in Γ have the same fixed points, and so the group is cyclic.

Case 2 Suppose Γ has an orbit of order 2 in $\mathbf{R}\bigcup\{\infty\}$. Then an element in the group either fixes each of these points or interchanges them. Then there can be no parabolic elements, for if we conjugate in order to express such an element as $f(z) = z + k$ for $k \neq 0$, it's clear that any z is either a fixed point or has an infinite orbit, as $f^n(z) = z + nk$ is an element of the group for any n and only ∞ is fixed by such a group of transforms.

If there are just hyperbolic elements, then the points in the orbit correspond to the fixed points of these elements, as hyperbolic elements can't interchange two points. To see why, express such an element as $f(z) = \lambda z$. Then f^2 fixes the two points, which means λ is 1, and f is just the identity. It follows that the group is cyclic. If all elements are elliptic, the lemma we proved earlier also yields this result.

Now consider the case in which there are both elliptic and hyperbolic elements. The hyperbolic elements must fix the two points, and the elliptic elements must alternate them. We can see that the latter claim is true by looking at the Poincar model and thinking of elliptic transformations as rotations, in which case fixing two points on the boundary are equivalent to the identity.

Let us conjugate Γ so that the two points of the orbit are 0 and ∞. In order to fix the two points, a hyperbolic element must be of the form $\begin{bmatrix} a & 0 \\ 0 & 1/a \end{bmatrix}$, and in order to alternate them, an elliptic element must be of the form $\begin{bmatrix} 0 & b \\ -1/b & 0 \end{bmatrix}$. We can certainly conjugate so that at least one elliptic element is of the form $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. We can use this transform so that if $\begin{bmatrix} 0 & \beta \\ -1/\beta & 0 \end{bmatrix}$ is any elliptic element in the group, there is also some hyperbolic element $\begin{bmatrix} \beta & 0 \\ 0 & -1/\beta \end{bmatrix}$ obtained by composition which in a sense corresponds to the elliptic element.

By discreteness, all the hyperbolic elements are generated by some element $\begin{bmatrix} k^{1/2} & 0 \\ 0 & 1/k^{1/2} \end{bmatrix}$, and since we can associate a hyperbolic element to each elliptic element by means of composition with $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the whole group is generated by these two transforms, which yields the desired result.

Case 3 Γ has an orbit of order 2 in **H** or an orbit of order $k \geq 3$ in the closure of **H**. Since parabolic and hyperbolic elements can have at most 2 fixed points on $\mathbf{R} \bigcup \{\infty\}$, and all points in **H** of these transforms have infinite orbits, then this means Γ contains only elliptic elements, and so is cyclic.

1.2 Non-Elementary Groups

The following results facilitate the discussion of Jörgensen's inequality.

Theorem 2 *A non-elementary subgroup of PSL(2,**R**) must contain a hyperbolic element.*

Call the subgroup Γ, and suppose it contains no hyperbolic elements. Then by the first lemma we showed, it must contain a parabolic element. Let us conjugate Γ so that this element fixes ∞, and is of the form $f = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Pick any other $g \epsilon \Gamma$ so that $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $f^n g = \begin{bmatrix} a+nc & b+nd \\ c & d \end{bmatrix}$, and $tr^2 = (a+d+nc)^2$. Since there are no hyperbolic elements, this value is less than or equal to 4 for any n. But since we can take n to be as large as we want, this implies $c = 0$. Then g also fixes ∞, and since g is arbitrary, every element of Γ fixes infinity, which means that Γ is elementary, which contradicts our assumptions.

Corollary 1 *Every non-elementary group contains infinitely many hyperbolic elements with distinct fixed points.*

It is rather easy to show that there are infinitely many hyperbolic elements in which one fixed point is distinct. We take our existent hyperbolic element, call it h, with fixed points a and b. Then all we need is a transform g in PSL(2,$\mathbf{R}$) that takes a third point c to either a or b, so that $g^{-1}hg$ fixes c.

However, we wish to show the stronger result: that we can find infinitely many hyperbolic elements with pairwise disjoint fixed points. So pick a $T \epsilon \Gamma$ hyperbolic, with fixed points a and b, and conjugate with some g as above so that $S = g^{-1}Tg$ has at least one fixed point which is different from those of T. Then we have two alternatives: $\{a, b\}$ and $\{S(a), S(b)\}$ are disjoint, or one element intersects both sets (both can't intersect as this would mean S exchanges the two points, which we've shown isn't possible for a hyperbolic transform).

Suppose they are disjoint. Then $H = STS^{-1}$ is hyperbolic and shares no fixed points with T. Suppose we were to take $STS^{-1}(a)$. Then S^{-1} sends a to some c different from a and b, T and T sends c to some d. If S sends d to a, it means $c = d$, which is a contradiction. The same holds for b. So the fixed points of H are not a or b.

We then construct a sequence T^nHT^{-n}, and claim that each term has fixed points which are pair-wise distinct. The reasoning is pretty much the same as above. If T^nHT^{-n} has a pair of fixed points, then they can't be the fixed points of T, so T^{-1} sends them to something different, and then T^nHT^{-n} maps them to a point which will not get inverted by T, so $T^{n+1}HT^{-(n+1)}$ can't have the same fixed points.

Now suppose $\{a, b\} \bigcap \{S(a), S(b)\}$ is non-empty. We can assume without loss of generality that a lies in the intersection. Then $[T, H] = THT^{-1}H^{-1} = TSTS^{-1}T^{-1}ST^{-1}S^{-1}$. If $S(a) = a$, this element clearly fixes a. If $S(b) = a$, then T and its inverse both fix a, and S and its inverse just send b to a and a to b, respectively. So H fixes a.

We wish to show that $[T, H]$ is parabolic, i.e. it fixes solely a. So we conjugate Γ so that the fixed points of T are 0 and ∞, and the fixed point it shares with H is ∞. Then $T = \begin{bmatrix} u & 0 \\ 0 & 1/u \end{bmatrix}$ for $u > 1$, and $H = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ for some a, b, c, d. If H is to fix ∞, then $c = 0$, and since 0 is not a fixed point, $b \neq 0$. This means H has the form $H = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$ for $b \neq 0$. We can then compute $[T, H]$ to be $\begin{bmatrix} 1 & b(u^2 - 1) \\ 0 & 1 \end{bmatrix}$, so that its transform is of the form $H(z) = z + t$ for non-zero t, and other than ∞, every point is fixed.

Call $P = [T, H]$. Since Γ is non-elementary, a has an infinite orbit, and we can find an element U such that U doesn't map a to a or b, so that U^{-1} doesn't map either a or b to a, and so $Q = UPU^{-1}$ fixes neither a nor b, but is still parabolic. Then Q^nTQ^{-n} is hyperbolic for all n, and for sufficiently high values of n, both a and b will be mapped to points far away from themselves, so that T will not fix them, and the set $\{a,b\}\bigcap\{Q^nTQ^{-n}(a), Q^nTQ^{-n}(b)\}$ will be empty, which reduces to the earlier case.

Theorem 3 *If a subgroup of PSL(2,**R**) contains no elliptic elements, it is either elementary or discrete.*

Call the subgroup Γ, and assume it's non-elementary. Then it contains a hyperbolic element h, which we can conjugate so it's given by the matrix $h = \begin{bmatrix} u & 0 \\ 0 & 1/u \end{bmatrix}$. We wish to show that for any sequence $g_n \to Id$, $g_n = Id$ for sufficiently large n.

Let $g_n = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}$ be such a sequence. Then $[h, g_n] = \begin{bmatrix} a_nd_n - b_nc_nu^2 & a_nb_n(u^2-1) \\ c_nd_n(1/u^2-1) & a_nd_n - b_nc_n/u^2 \end{bmatrix}$, so that $tr[h, g_n] = 2a_nd_n - b_nc_n(u^2 - 1/u^2) = 2a_nd_n - b_nc_n((u-1/u)^2+2) = 2(a_nd_n - b_nc_n) - b_nc_n(u-1/u)^2 = 2 - b_nc_n(u-1/u)^2$. Since $g_n \to Id$, $[h, g_n] \to Id$ too, which means this value tends to 2. Since there are no elliptic elements, $|tr[h, g_n]| \geq 2$, this means that for large n, $b_nc_n \leq 0$.

Now let us define $f_n = [h, g_n] = \begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix}$. As the calculations are the same as above, we know $tr[h, f_n] = 2 - B_nC_n(u-1/u)^2 = 2-(a_nb_nc_nd_n(u^2 - 1)(1/u^2-1)(u-1/u)^2 = 2 - (a_nb_nc_nd_n(2-(u^2+1/u^2)(u-1/u)^2) = 2 + a_nb_nc_nd_n(u-1/u)^4 = 2 + (1+b_nc_n)b_nc_n(u-1/u)^4$. So for large n, where b_n and c_n are close to 0, $b_nc_n \geq 0$ so that the above expression remains greater than or equal to 2. If we combine the results, this means that for sufficiently large n, $b_nc_n = 0$, so either $c_n = 0$, in which case g_n fixes ∞, or $b_n = 0$, in which case it fixes 0.

Either way, g_n will share a fixed point with h, as those are the fixed points of h in this conjugation. If we pick three hyperbolic elements with distinct fixed points, h_1, h_2 and h_3, then we can obtain the same results without changing our sequence g_n, so that g_n, for sufficiently large n, shares a fixed point with three elements such that these fixed points are all distinct. Then g_n must be the identity, as otherwise it can have at most 2 fixed points. So in the case that Γ is non-elementary, it is discrete.

2 Jörgensen's Inequality

Jörgensen's inequality is a theorem that places a large constraint on the way non-elementary, discrete groups behave. First, we define $< T, S >$ to be the group generated by elements T, S. The theorem states that if a discrete group generated by two such elements is non-elementary, then at least one of this elements must differ considerably from the identity.

Theorem 4 *If $T, S \epsilon PSL(2, \mathbf{R})$, and $< T, S >$ describes a discrete, non-elementary group, then $|Tr(T)^2 - 4| + |Tr(TST^{-1}S^{-1}) - 2| \geq 1$. The lower bound is best possible.*

The proof will proceed by contradiction. But first, we show two lemmas.

Lemma 2 *Suppose $T, S \epsilon PSL(2, \mathbf{R})$ and $T \neq Id$. Define $S_0 = S, S_1 = S_0 T S_0^{-1}, \ldots, S_{r+1} = S_r T S_r^{-1}, \ldots$. If, for some n, $S_n = T$, then $< T, S >$ is elementary.*

Suppose T has a fixed point a. Then since S_r for $r > 0$ are conjugate to T, they also have one fixed point. We have $S_{r+1} S_r(a) = S_r T S_r^{-1} S_r(a) = S_r(a)$, so S_{r+1} fixes $S_r(a)$. Since $T = S_n$, it follows that $S_{n-1}(a) = a$, and by induction, S_r fixes a for all r. Since a is fixed by the generators of $< T, S >$, it is fixed by all elements in the group, and so the group is elementary.

Now suppose T has exactly two fixed points. Then it is hyperbolic, and we can conjugate everything so that $T(z) = kz$ for some $k \neq 1$. Then S_r all have two fixed points. Recalling that $T(0) = S_{n-1}(0)$ and $T(\infty) = S_{n-1}(\infty)$, the same reasoning as above lets us conclude that S_r all map $\{0, \infty\}$ onto the same set (in fact, these are the fixed points for all but S_0). Since these two points are fixed by the generators of $< T, S >$, the same holds for all elements in the group, and so it's elementary.

Lemma 3 *Let $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be two matrices in PSL(2,**R**). Then the inequality in Jörgensen's inequality holds if and only if $|c| \geq 1$.*

Clearly, this reduces to showing that $|Tr(TST^{-1}S^{-1}) - 2| \geq 1$ The proof is purely computational.

$[T, S] = \begin{bmatrix} d(a+c) - c(-a-c+b+d) & a(-a-c+b+d) - b(a+c) \\ cd - c(-c+d) & -bc + a(-c+d) \end{bmatrix}$

so that

$Tr[T, S] = ad + cd + ac + c^2 - bc - cd - bc - ac + ad = 2(ad - bc) + c^2 = 2 + c^2$.

It follows that the equality holds when $|c|^2 \geq 1$, which of course is true if and only if $|c| \geq 1$.

2.1 Proof of the theorem

We first get rid of a trivial case: if T is of order 2, then $Tr(T) = 0$ so that the equality holds.

For the other cases, we define S_r as in the first lemma of the previous section. We wish to show that if the inequality doesn't hold, then $T = S_n$ for some n, which implies that $< T, S >$ is elementary.

Case 1: T is parabolic Since the trace is invariant under conjugation, we may assume $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, where $c \neq 0$ (otherwise ∞ is fixed by both S,T and the group is elementary). Suppose that the inequality doesn't hold. Then our earlier lemma shows that $|c| < 1$. If we write $S_n = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}$, then we obtain from $S_{n+1} = S_n T S_n^{-1}$:

$$\begin{bmatrix} a_{n+1} & b_{n+1} \\ c_{n+1} & d_{n+1} \end{bmatrix} = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_n & -b_n \\ -c_n & a_n \end{bmatrix} = \begin{bmatrix} 1 - a_n c_n & a_n^2 \\ -c_n^2 & 1 + a_n c_n \end{bmatrix} \quad (1)$$

By induction, we get that for $n > 0, c_n = -c^{2n}$, which goes to 0 as $n \to \infty$, since $|c| < 1$. We also have by induction that $|a_n| \leq n + |a|$. Since a_n is linearly bounded, even if $a_n \to \infty$, $a_n c_n \to 0$. Since this means $a_n \to 1$, $\begin{bmatrix} 1 - a_n c_n & a_n^2 \\ -c_n^2 & 1 + a_n c_n \end{bmatrix} \to \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, and by discreteness, $T = S_n$ for some n.

Case 2: T is hyperbolic We may assume that $T = \begin{bmatrix} u & 0 \\ 0 & 1/u \end{bmatrix} S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ for $u > 1$. Neither b nor c are 0, otherwise T and S share a fixed point, and the group is elementary. So $bc \neq 0$.

Note that $[T, S] = \begin{bmatrix} a_n d_n - b_n c_n u^2 & a_n b_n (u^2 - 1) \\ c_n d_n (1/u^2 - 1) & a_n d_n - b_n c_n / u^2 \end{bmatrix}$

If the inequality fails, then

$\mu = |Tr(T)^2 - 4| + |Tr[T, S] - 2| = |(u + 1/u)^2 - 4| + |2 - bc(u - 1/u)^2 - 2| = |u - 1/u|^2 + |bc||u - 1/u|^2 = (1 + |bc|)|u - 1/u)^2 < 1$.

We also can use $S_{n+1} = S_n T S_n^{-1}$ to write:

$\begin{bmatrix} a_{n+1} & b_{n+1} \\ c_{n+1} & d_{n+1} \end{bmatrix} = \begin{bmatrix} a_n d_n u - b_n c_n / u & a_n b_n (u - 1/u) \\ c_n d_n (u - 1/u) & a_n d_n / u - b_n c_n u \end{bmatrix}$. Then $b_{n+1} c_{n+1} = -b_n c_n (1 + b_n c_n)(u - 1/u)^2$

By induction, $|b_n c_n| \leq \mu^n |bc|$, so $b_n c_n \to 0$. Since $a_n d_n = 1 + b_n c_n$, then it goes to 1, and so $a_n \to u, d_n \to 1/u$

$|\frac{b_{n+1}}{b_n}| = |a_n(1/u - u)| \to |u(1/u - u)| \leq \mu^{1/2}|u|$. So $|\frac{b_{n+1}}{u^{n+1}}| < \mu^{1/2} |\frac{b_n}{u^n}$, and so $\frac{b_n}{u^n} \to 0$. Likewise, $c_n u^n \to 0$.

Therefore, if we define a series $T^{-n} S_{2n} T^n = \begin{bmatrix} a_{2n} & b_{2n}/u^{2n} \\ c_{2n} u^{2n} & d_{2n} \end{bmatrix} \to T$.

Since $< T, S >$ is discrete, for large $n, T^{-n} S_{2n} T^{n} = T$, which means $S_{2n} = T$, and so the conditions for our lemma are satisfied.

Case 3: T is elliptic Reduces to Case 2 in the unit disc model.

To show that the lower bound is best possible, we consider the group generated by $T(z) = z + 1$ and $S(z) = -1/z$. $< T, S >$ is clearly non-elementary, and since these elements actually belong to PSL(2,**Z**), the group is discrete (should this statement prove too much of a leap of faith, a geometric intuition of the transforms should suffice to see that they are discrete, as T is a translation with discrete steps, and S is a reflection across i). Since $[T, S] = \frac{2z+1}{z+1}$ has trace 3, equality is satisfied.