

REMARKS ON THE BOUNDARY MAPPING FOR A FUCHSIAN GROUP

STEPHEN AGARD

If Γ, Γ' are discrete Möbius groups acting discontinuously in the upper half space $U = U^{n+1}$, and if there is a homeomorphism $f_0: \mathcal{R} = \Gamma \backslash U \rightarrow \mathcal{R}' = \Gamma' \backslash U$, then the situation induces a mapping $\hat{f}: U \rightarrow U$ with the *compatibility property*

$$\hat{f} \circ A \circ \hat{f}^{-1} \in \Gamma' \quad (\text{all } A \in \Gamma),$$

and an isomorphism $\Theta: \Gamma \rightarrow \Gamma'$ in which $A \mapsto \hat{f} \circ A \circ \hat{f}^{-1}$. It may happen that \hat{f} extends to $\mathbf{R}^n = \partial U^{n+1}$, inducing there a mapping $f: \mathbf{R}^n \rightarrow \mathbf{R}^n$ (the *boundary mapping*) with the same compatibility property, at least for the restriction of A to \mathbf{R}^n . Since a Möbius transformation which fixes U^{n+1} is completely determined by its values on \mathbf{R}^n , the status of the isomorphism Θ is the same whether f or \hat{f} is used.

Theorems of *Mostow-rigidity type* draw conclusions about f or Θ from various assumptions about \mathcal{R}, Γ, f_0 . It is not my purpose here to review all the relations among f, Θ, f_0 as this has been done in many places in the literature. However, it is appropriate to point out that Θ determines f to the extent that f maps the attracting (respectively repelling) fixpoint of nonelliptic $A \in \Gamma$ on that of $\Theta(A)$, and so *at best*, Θ determines f only on the limit set, which is the closure of the set of fixed points. Therefore, while the most general type of conclusion would be that Θ is conjugation by a Möbius transformation, in the context that f is continuous and Γ is of the first kind, the appropriate corresponding conclusion is that f itself is Möbius.

While no doubt the primary motivation for Mostow-rigidity theorems was the conclusion about f_0 (that $f_0: \mathcal{R} \rightarrow \mathcal{R}'$ is homotopic to a conformal map), this inference as well as the labor involved in establishing the appropriate properties of f (existence, uniqueness, continuity, univalence, quasiconformality), are a distraction. I shall therefore refrain from further comments in this area, and simply assume for f the property whose consequences I would like to study.

In this note I wish to clarify a remark by G. D. Mostow which is to be found on page 178 of his book [5]: If $n=1$, if Γ has finite covolume, and if f has a derivative which is finite and nonzero at *even a single point*, then f is linear fractional.

The situation for $n=1$ is of special interest because of the striking contrast to the cases $n \geq 2$, where if Γ has finite covolume, f is *always* Möbius [4]. The effect is to render discrete the Teichmüller spaces for many manifolds of dimension ≥ 3 . The fertility of Teichmüller spaces of Riemann surfaces (dim 2) supports the contrast.

A rigidity theorem for the case $n=1$ was proved by Kuusalo [3], in what seems to me to have been the pioneering paper. He showed that if Γ acts ergodically on $\mathbf{R} \times \mathbf{R}^*$ there are only the two alternatives: either f is linear fractional, or f is singular ($f'=0$ a.e.).

Variations of this have continued to appear in the literature. For example, in a posthumous article in 1979, the late Rufus Bowen [2] gave a "new proof" that if \mathcal{R} is compact, Γ without torsion, and if f is absolutely continuous, then f is linear fractional.

At issue in the present note is the question of whether, as remarked by Mostow, the existence of a positive finite derivative at a single point, suffices to force f to be Möbius. And if not always, then to what points or groups can such a statement be applied?

We make two steps toward a resolution of this question by showing: f *always* has finite, positive derivative at every parabolic fixpoint; however, the statement *does* apply to hyperbolic fixpoints in general, and to all points for those groups for which \mathcal{R} is compact. The latter argument was largely provided to me by Dennis Sullivan during his 1982 visit to the IMA in Minneapolis.

To make this article completely self-contained, we begin with a review of procedures by which the Bowen/Kuusalo theorem can be proved, using an early (1931) density theorem of P. J. Myrberg [6].

1. *The Theorems of Myrberg, Bowen, and Kuusalo.* We shall work in $\bar{\mathbf{R}}$, the one point compactification of \mathbf{R} . The Möbius group $M(\mathbf{R})$ may be identified with $SL(2, \mathbf{R})/\{I, -I\}$ through the homomorphism $\Phi: SL(2, \mathbf{R}) \rightarrow M(\mathbf{R})$, in which

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto A, \quad \text{where } A(x) = \frac{ax+b}{cx+d} \quad \begin{cases} x \in \bar{\mathbf{R}} \\ ad-bc = 1. \end{cases}$$

The topology is induced by the entry-wise topology in $SL(2, \mathbf{R})$. Thus $M(\mathbf{R})$ is a 3-manifold.

These Möbius transformations have extensions in $M(\mathbf{U})$, the Möbius transformations of the upper half plane. It is only necessary to replace x by the complex variable z . As such, the fixpoint classification is familiar: A is *hyperbolic* if it has two fixpoints in $\bar{\mathbf{R}}$; *parabolic* if it has one fixpoint (necessarily in $\bar{\mathbf{R}}$); *elliptic* if it has two fixpoints outside \mathbf{R} (necessarily complex conjugates). In these cases A is respectively conjugate in $M(\mathbf{R})$ to a dilation $h_\lambda: x \mapsto \lambda x$, $\lambda > 1$; a translation $t_a: x \mapsto x+a$, $a \in \mathbf{R}$; or a mapping whose Φ -inverse contains a matrix in $SO(2)$. Denoting the classes respectively by H, P, E , we have the disjoint decomposition $M(\mathbf{R}) = H \cup P \cup E \cup \{\text{id}\}$, and $P \cup \{\text{id}\} = \partial E = \partial H$.

* This condition is equivalent to the property known as "divergence type", and is implied by finite covolume.

Regarding hyperbolics, I will denote by $A_{pq}(p, q \in \overline{\mathbf{R}})$ the set of mappings A with attracting fixpoint $p=P(A)$, and repelling fixpoint $q=N(A)$. The number $\lambda=\lambda(A)$ is known as the *multiplier*.

The non-Euclidean *lines* in U are the intersections with U of circles or lines perpendicular to \mathbf{R} . For $p, q \in \overline{\mathbf{R}}$, the one joining p and q will be denoted by l_{pq} . A ray $r=[a, p)$ is designated by that portion of l_{pq} between $a \in l_{pq} \cap U$ and $p \in \overline{\mathbf{R}}$. Of course the lines are permuted by $A \in M(U)$, as are the rays. The lines are globally parametrized by their endpoints as the subset $\{(p, q): p < q \cong \infty\} \subseteq \mathbf{R} \times \overline{\mathbf{R}}$, with a topology inherited from the latter. To say that a sequence of rays $[a_n, p_n)$ converges to a hyperbolic line l will mean that $p_n \rightarrow p$, $a_n \rightarrow q$, and $l=l_{pq}$.

It is clear that if a sequence $\{A_n\} \subseteq M(\mathbf{R})$ has $A_n \rightarrow A \in M(\mathbf{R})$, then $A_n(z) \rightarrow A(z)$ for every $z \in C$, and that $A_n(l) \rightarrow A(l)$ for any hyperbolic line l .

We usually assume that a subgroup $\Gamma \subseteq M(\mathbf{R})$ is *discrete* in the topology of $M(\mathbf{R})$, in which case Γ acts discontinuously in U ; the 2-manifold $\mathcal{R}=\Gamma \backslash U$ is a Riemann surface; the projection $\pi: U \rightarrow \mathcal{R}$ is a covering branched at the elliptic fixpoints of Γ ; and the mappings $A \in \Gamma$ are the cover transformations. We say that Γ is of the *first kind* if the hyperbolic fixpoints are dense in \mathbf{R} . This is a consequence of several possible assumptions, the most common being either that Γ has *finite covolume*, or that \mathcal{R} is *compact*. The first means that the coset space $\Gamma \backslash M(\mathbf{R})$ has finite invariant measure, and implies among other things that the non-Euclidean area for each fundamental region for Γ is finite. It allows a finite number of inequivalent parabolic fixpoints, and is implied by the compactness of \mathcal{R} , which however excludes all parabolics. Finite covolume implies as well the algebraic condition that Γ is *finitely generated*. First kind and finite generation are utterly unrelated, but together imply finite covolume.

We introduce three definitions:

(i) A point $p \in \mathbf{R}$ is a *Myrberg-density point* ($p \in M_y(\Gamma)$) if the Γ -images of some ray $r=[a, p)$ are dense in the hyperbolic lines of U .

(ii) A point $p \in \mathbf{R}$ is a *Mostow-density point* ($p \in M_o(\Gamma)$) if ΓA_{pq} is dense in $M(\mathbf{R})$ for at least one $q \neq p$.

(iii) A point $p \in \mathbf{R}$ is a *Mostow-rigidity point* ($p \in M_r(\Gamma)$) if the conditions: f compatible with Γ , $f'(p) \in (0, \infty)$; imply that f is linear fractional.

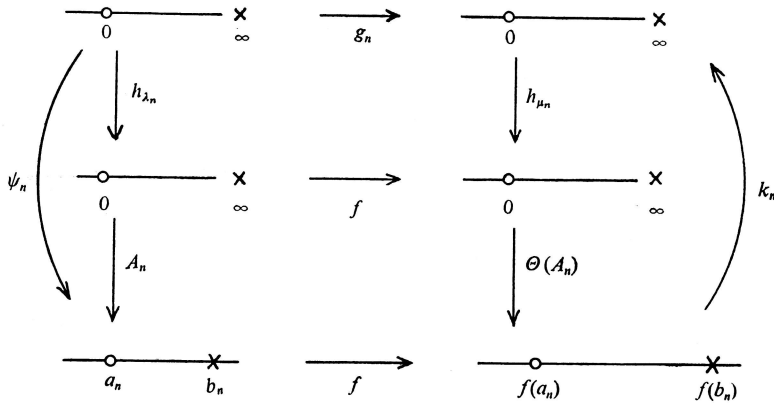
Remarks. First, it is clear that for any $A \in M(\mathbf{R})$, and any of the sets M_i ($i=y, o, r$), that $M_i(A\Gamma A^{-1})=A(M_i(\Gamma))$. This observation allows us to normalize many situations. Next, we observe in connection with M_y , that if $\Gamma(r)$ is dense in lines, then by projection on \mathcal{R} , $\pi(r)=\pi(\Gamma(r))$ is dense in \mathcal{R} . Because any two rays $[a, p)$ and $[b, p)$ (collinear or not) are asymptotic at p in the hyperbolic metric, the condition that $\pi(r)$ be dense on \mathcal{R} is independent of a . This is a crucial observation, because by contraposition, if $\pi(r)$ is *not* dense on \mathcal{R} for *some* ray $r=[a, p)$, then $p \notin M_y(\Gamma)$.

Theorem 1. $M_y = M_o \subseteq M_r$, for any discrete group $\Gamma \subseteq M(\mathbf{R})$.

Proof. To show $M_y \subseteq M_o$, suppose that $p \in M_y(\Gamma)$, that $q \neq p$, and that $r = [a, p] \subseteq l_{pq}$ is a ray whose images are dense in lines. Given $A \in M(\mathbf{R})$, label $s = A(p)$, $t = A(q)$, and select $A_n \in \Gamma$ with $A_n(p) \rightarrow s$, $A_n(a) \rightarrow t$. Thus, relative positions on r are stretched by A_n away from p and toward q . Using premultiplication by suitable elements $B_n \in A_{pq}$ with compensating multipliers tending to ∞ , the convergence $A_n B_n \rightarrow A$ is easily achieved.

Conversely, if ΓA_{pq} is dense in $M(\mathbf{R})$, if $[a, p] \subseteq l = l_{pq}$, and if a target line l_{st} is selected, fix some $A \in M(\mathbf{R})$ with $A(p) = s$, $A(q) = t$, and obtain sequences $A_n \in \Gamma$, $B_n \in \Gamma_{pq}$ with $A_n B_n \rightarrow A$. Under B_n the ray $[a, p]$ is merely contracted along l . But because Γ is discrete, the multipliers $\lambda(B_n)$ must tend to ∞ , so the contraction is right down to p itself. Thus B_n^{-1} stretches r out (in the limit) over all of l , and we see $A_n(r) = A_n B_n (B_n^{-1}(r)) \rightarrow A(l) = l_{st}$. Thus $p \in M_y(\Gamma)$, and we have shown $M_o \subseteq M_y$.

Finally, to understand the inclusion $M_o \subseteq M_r$, suppose $p \in M_o$. We may assume $p = 0$, that $\Gamma A_{0\infty}$ is dense in H , that $f'(0) = c \in (0, \infty)$, and as well that f fixes $0, 1, \infty$. Observe that $A_{0\infty}$ may be identified with the inverse dilations $\{h_\lambda : 0 < \lambda < 1\}$. Because $\Gamma A_{0\infty}$ is dense in $M(\mathbf{R})$, there exist numbers $\lambda_n \in (0, 1)$, and elements $A_n \in \Gamma$, with $A_n h_{\lambda_n} \rightarrow \text{id}$. We may assume by discreteness that $\lambda_n \rightarrow 0$. Consider this diagram, in which the progress of 0 (\circ) and ∞ (\times) are noted:



Here, $k_n \in M(\mathbf{R})$ is a normalized mapping carrying $f(a_n) \mapsto 0$, $f(b_n) \mapsto \infty$, specifically

$$k_n(x) = -f(b_n) \left(\frac{x - f(a_n)}{x - f(b_n)} \right),$$

with a simple modification if $f(b_n)$ is already ∞ : $k_n(x) = x - f(a_n)$.

We observe that $k_n(x) \rightarrow x$ (uniformly on compact sets) precisely because $f(a_n) \rightarrow 0$, $f(b_n) \rightarrow \infty$, which facts are assured by the continuity of f and the condition $\psi_n \rightarrow \text{id}$. The numbers μ_n are then completely determined by $h_{\mu_n}^{-1} = k_n \Theta(A_n)$. It follows from the two representations

$$h_{\mu_n}^{-1} \circ f \circ h_{\lambda_n} = g_n = k_n \circ f \circ \psi_n$$

and the established convergences, that

$$\lim_{n \rightarrow \infty} \frac{f(\lambda_n x)}{\mu_n} = f(x).$$

Next, it follows from the condition $c = f'(0)$, and by pruning the sequences so that λ_n/μ_n has a limit L (finite or infinite) that

$$Lcx = \lim \frac{\lambda_n x}{\mu_n} c = \lim \frac{\lambda_n x}{\mu_n} \frac{f(\lambda_n x)}{\lambda_n x} = f(x).$$

Precisely because c is positive, it now follows that L is finite, and $f(x) \equiv Lcx \equiv x$, the last because $f(1) = 1$. Finally, when the previous normalizations are taken into account, we have the relaxed conclusion for the original context: $f \in M(\mathbf{R})$. \square

Remarks. (i) As soon as Möbius groups $M(\mathbf{R}^n)$, $n \geq 2$, are considered, the classification $H \cup P \cup E$ needs to be revised. In addition to hyperbolics, one needs to consider *loxodromics* (L), which are conjugate in $M(\mathbf{R}^n)$ to mappings of the form uh_λ where u is orthogonal. The terminology *hyperbolic* for the case $u = \text{id}$ is used to some advantage still in $n = 2$, but rarely (exceptions in [11]) to advantage in $n \geq 3$. For our purpose, a decomposition

$$M(\mathbf{R}^n) = L \cup P \cup E \cup \{\text{id}\}$$

is preferred. (Now only L has codimension zero.)

As for the set M_r , it is appropriate to replace the hypothesis $f'(p) \in (0, \infty)$ by the hypothesis that f have nonsingular total differential, and the conclusion $f \in M(\mathbf{R})$ by the condition that f be conjugate in $M(\mathbf{R}^n)$ to a linear map. With these modifications, there is little problem to carry through the proof of Theorem 1 in any dimension [1].

(ii) P. J. Myrberg [6] proved that M_y has full measure in \mathbf{R} for certain special kinds of Fuchsian groups ($n = 1$). Mostow proved that M_o has full measure in \mathbf{R}^n assuming finite covolume. The first treatment of the case $n = 1$ occurs in his book [5], which includes other generalizations not directly relevant here. It is precisely the extensive scope of his book which has obscured the situation as regards the case $n = 1$. Lost in the shuffle, one might say.

(iii) Sullivan [9] introduced the term *Mostow-rigid* to describe groups for which, in the context of the introduction, Θ is *always* conjugation in $M(\mathbf{R}^n)$. It was the

main thrust of Mostow's earlier paper [4] that this is so for groups of finite covolume in dimensions $n \geq 2$. The proof required three steps: (1) the boundary mapping f is quasiconformal, therefore (if $n \geq 2$) differentiable with positive Jacobian a.e.; (2) M_o has full measure in \mathbf{R}^n , hence at least one of the points with nonsingular total differential belongs to $M_o \subseteq M_r$, and therefore f is conjugate in $M(\mathbf{R}^n)$ to a linear map; finally (3) a linear map which is compatible with a discrete group having at least three points in the limit set, is necessarily conformal. See [1]. Thus the *simplicity* of the case $n=1$ is that step (3) is not required. The *problem* is that step (1) has an inapplicable inference, but the saving feature for Teichmüller space theory, is that $f'(x)=0$ a.e. is a possibility.

(iv) To explain his stronger remark, Mostow in essence asserted ([5], page 178)* that under the circumstances, *any* point p was a density point (M_o), and thus inferred that even a single point with $f'(p) \in (0, \infty)$ was sufficient to conclude that f was linear fractional. This has to be an oversight, because it is clear that fixpoints cannot be Myrberg density points. Indeed, if r is a ray terminating at a parabolic fixpoint, then $\pi(r)$ tends to the corresponding "puncture" on \mathcal{R} . Likewise, if p, q are fixed for $A \in \Gamma \cap \mathbf{H}$, then $\pi(l_{pq})$ is compact on \mathcal{R} , as is $\pi(r)$ for any ray r lying in l_{pq} . In either case $\pi(r)$ is not dense on \mathcal{R} , and p is not in M_y .

(v) On the other hand, as soon as $f'(p) \in (0, \infty)$ on a set of positive measure, then one of these points belongs to $M_y \subseteq M_r$, and $f \in M(\mathbf{R})$. Thus Bowen's theorem follows from Myrberg's, since the groups Myrberg considered included compact \mathcal{R} , Γ without torsion. Kuusalo's theorem, with its sharp dichotomy, also follows for any groups for which M_y has full measure. This now includes groups of divergence type [1], and probably more [9].

(vi) It is of interest to note that rigidity of the group, while no longer prevalent in case $n=1$, still occurs. For example, the modular group Γ_0 and its congruence subgroup are both rigid. Sheingorn [7] is investigating ways to characterize the points of $M_y(\Gamma_0)$, and has in particular found ways of identifying some points (other than fixpoints) which are definitely not in $M_y(\Gamma_0)$. His method promises to be very difficult to apply to other groups. Myrberg himself reported that (at least for genus zero) the complement of M_y is always an uncountable set of Cantor type, from which it would follow that more than fixpoints are involved.

2. *Fixpoints.* We study the behavior of the boundary map near fixpoints for the group. Observe that no assumptions about the group are required in this section.

Theorem 2. *Let f be the boundary map for the isomorphism Θ and assume that f is not linear fractional.*

- (1) *If $p \in \mathbf{R}$ is any parabolic fixpoint for Γ , then $f'(p)$ exists, and lies in $(0, \infty)$.*
- (2) *If $p \in \mathbf{R}$ is fixed for hyperbolic $A \in \Gamma$, then*

* In Mostow's notation, $\triangleleft A(q, p)$ is what I have called A_{pq} .

- (a) $f'(p)=0$ if $\lambda(A)<\lambda(\Theta(A))$,
 (b) $f'(p)=\infty$ if $\lambda(A)>\lambda(\Theta(A))$,
 (c) $f'(p)$ does not exist if $\lambda(A)=\lambda(\Theta(A))$.

Proof. First suppose p is a parabolic fixpoint. By various conjugations in $M(\mathbf{R})$, we may assume $p=\infty=f(p)$, that Γ, Γ' both contain the translation $T: x \mapsto x+1$, and that $\Theta(T)=T$. We may further assume that 0 is not equivalent to ∞ under Γ or Γ' and that $f(0)=0$. The effect of these normalizations is to ensure that $f(n)=n$ for all integers $n=0, \pm 1, \pm 2, \dots$. It follows trivially that as $|x| \rightarrow \infty$, $\lim f(x)/x=1$. When this is filtered back through the normalizations, it says precisely that $f'(p) \in (0, \infty)$.

As regards (2), the first observation is that an attracting fixpoint for A is repelling for A^{-1} . But A and A^{-1} have the same multiplier, as do $\Theta(A)$ and $\Theta(A^{-1})$. Therefore it suffices to treat an attracting fixpoint. We may then proceed as above, this time assuming that $p=0=f(p)$, and that $A, \Theta(A)$ are respectively h_α, h_β , where $\lambda(A)^{-1}=a<1$. But 0 is also attracting for h_β , and we now know that $\lambda(\Theta(A))^{-1}=\beta<1$. We also know $f(\infty)=\infty$ (repelling fixpoints for h_α, h_β) and therefore f is a homeomorphism of \mathbf{R} .

Now fix $x \in [-1, -\alpha] \cup [\alpha, 1] = E$. Then $y=f(x)$ also lies in a compact set $f(E)$ not containing 0. If we define x_n by $\alpha^n x$, we see that $f(x_n)/x_n = (\beta/\alpha)^n y/x$. In case (a), we have $\beta < \alpha$, and as $n \rightarrow \infty$, $\lim f(x_n)/x_n = 0$ uniformly for $x \in E$. It follows that $f'(0)=0$. Similarly $f'(0)=\infty$ if $\beta > \alpha$. Finally, if $\alpha=\beta$, then it follows that the collection of all limit values for $f(x)/x$ as $x \rightarrow 0$ coincides with the closed set $\{f(x)/x: x \in E\}$. This can be a single point only if $f(x) \equiv cx$, which is the linear fractional case. \square

Remarks. Teichmüller space theory assures that the three cases above for hyperbolic fixpoints can all occur in a single example. A question raised, however, is this: Can the case $\lambda(A) \equiv \lambda(\Theta(A))$ persist for all hyperbolic A ? The question of equality for all A has been much studied. For example, Sorvali [8] showed that equality holds for all A only if Θ is conjugation in $M(\mathbf{R})$. Tukia has vastly generalized his result [11], allowing arbitrary dimension, and even relaxing discreteness. I have borrowed Sorvali's method in dealing with hyperbolic fixpoints.

Finally, we can rephrase the conclusions of Theorem 2 in the language of Section 1:

Corollary. For discrete $\Gamma \subseteq M(\mathbf{R})$, all hyperbolic fixpoints lie in $M_r \setminus M_y$, whereas no parabolic fixpoints can belong to M_r .

3. *The case of compact \mathcal{D} .* It is first necessary to introduce a tool which Sullivan has shown to be very useful. This is the notion of *Bounded Distortion* (BD). We consider a family \mathcal{J} of real valued functions of a real variable, each $f \in \mathcal{J}$ defined on a closed interval \mathcal{D}_f . We say that \mathcal{J} is of bounded distortion if each f belongs to

\mathcal{C}' , is monotone increasing, $f'(x) > 0$, and there is some K with

$$\frac{f'(x)}{f'(y)} \leq K \quad (\text{all } x, y \in \mathcal{D}_f, \text{ all } f \in \mathcal{F}).$$

Following is a list of obvious properties:

(BD 1) A family \mathcal{F} is BD if and only if the family $\mathcal{F}^{-1} = \{f^{-1} : f \in \mathcal{F}\}$ is BD.

(BD 2) If the domains have comparable lengths, and the images likewise have comparable lengths, and if \mathcal{F} is BD, then it is also uniformly bilipschitzian, hence equicontinuous, normal with compact closure, and all limit functions are bilipschitzian.

(BD 3) If the images all lie in a finite interval \mathcal{E} , and if $\varphi : \mathcal{E} \rightarrow \mathcal{E}'$ ($\varphi \in \mathcal{C}'$) has its derivative bounded and bounded away from zero, then the composition $\varphi \circ \mathcal{F} = \{\varphi \circ f : f \in \mathcal{F}\}$ is BD whenever \mathcal{F} is BD.

(BD 4) If for $f \in \mathcal{F}$, $l = l_f$ is an affine rescaling $t \mapsto c_f t + d_f$ of $l_f^{-1}(\mathcal{D}_f)$ onto \mathcal{D}_f , then the rescaled family $\mathcal{F}_r = \{f \circ l_f : f \in \mathcal{F}\}$ is (BD) if and only if \mathcal{F} is (BD).

We shall prove that if \mathcal{R} is compact, then every point is a Mostow-rigidity point. For simplicity I will assume Γ is without elliptics. The assumption is inessential. One might appeal to Selberg's lemma and work with a torsion free subgroup of finite index (still with compact manifold and the same boundary map), or one may study the proof of Lemma 1 below, and observe that the requirement $B_n \in \mathbf{H}$ is inessential there. The lemma is only harder to state, not to prove. Thus, we now set out two lemmas on contracting and expanding maps, and indicate how the theorem is proved. The lemmas will be proved in their turn (Section 4).

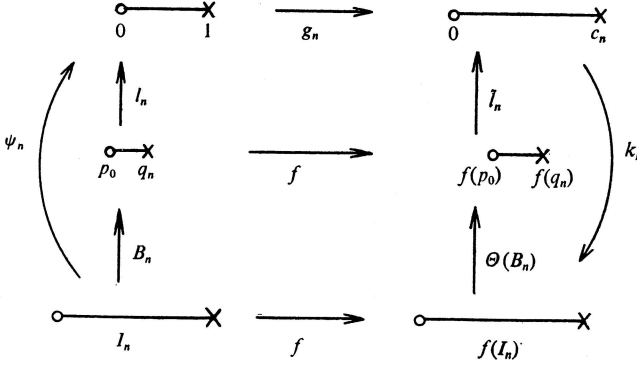
Lemma 1. If B_n is a discrete sequence in \mathbf{H} , and to each B_n is associated an interval $I_n \subset \mathbf{R}$ with $I_n \rightarrow I_0 = [a, b]$, and $B_n(I_n) \rightarrow p \in (a, b)$, then the family $\mathcal{F} = \{B_n|_{I_n}\}$ is BD if and only if the repelling fixpoints $N(B_n)$ cluster only outside I_0 .

Lemma 2. If \mathcal{R} is compact, and $p \in \mathbf{R}$, then there exists a sequence of intervals $J_n = [p, p_n]$ ($p_n \rightarrow p$), and a sequence $A_n \in \Gamma$, with $A_n(J_n)$ convergent to a nondegenerate interval I_0 , and such that $\mathcal{D} = \{A_n|_{J_n}\}$ is BD.

Theorem 3. Suppose \mathcal{R} is compact, $f'(p) \in (0, \infty)$. Then f is linear fractional.

Proof. Let us apply Lemma 2 at p . Because Γ is of the first kind, there is $A_0 \in \Gamma$ with $p_0 = A_0(p)$ interior to I_0 . Take $B_n = A_0 A_n^{-1}$, $I_n = A_n(J_n)$, and consider the families $\mathcal{F} = \{B_n|_{I_n}\}$, $\bar{\mathcal{F}} = \{\Theta(B_n)|_{\mathcal{F}(I_n)}\}$. It is clear that the set-up of Lemma 1 is in force for both families. However, \mathcal{F} is BD because \mathcal{D} is BD, in view of (BD 3). Therefore the repelling fixpoints of B_n are bounded away from I_0 . Hence in turn, and because f is a homeomorphism, the repelling fixpoints of $\Theta(B_n)$ are bounded away from $f(I_0)$, and therefore $\bar{\mathcal{F}}$ is also BD. We consider the following diagram, which is quite comparable to the one in Section 1, except that l_n is an affine rescaling

$B_n(I_n)=[p_0, q_n]$ ($q_n=A_0(p_n)\rightarrow A_0(p)=p_0$) to the unit interval $[0, 1]$, and \tilde{l}_n is an affine rescaling of $f(B_n(I_n))$ to an interval $[0, c_n]$ using the same scale factor as l_n .



In view of the properties (BD 1—4) the mappings ψ_n may be assumed convergent to bilipschitzian ψ . We also easily see that $c_n=(f(q_n)-f(p_0))/(q_n-p_0)$ has limit $f'(p_0)=\Theta(A_0)'(f(p))f'(p)/A_0'(p)\in(0, \infty)$, so that k_n also may be assumed to converge to bilipschitzian k . Of course g_n converges to the linear map $x\rightarrow f'(p_0)x$, and we arrive at the formula $f(x)=k(f'(p_0)\psi(x))$, which shows that f is absolutely continuous on I_0 . By the Kuusalo/Bowen theorem, f is linear fractional. \square

It is perhaps worth noting that this argument applies even if we only know that f has a positive, finite, one-sided derivative at p .

4. The Lemmas. A moment's reflection will convince the reader that if the lemmas can be proved in the context of the circle \mathcal{S} , that the results will transfer to \mathbf{R} , thanks to (BD 3). One interprets the derivative with respect to arc length, but this is the modulus of the ordinary complex derivative anyway. An interval J becomes a subarc of \mathcal{S} , with length $|J|$. The upper half plane U becomes the unit ball \mathbf{B} .

One advantage of course is that \mathcal{S} is compact, and each Möbius transformation is class \mathcal{C}' over all of \mathcal{S} , with derivative bounded and bounded away from zero, and second derivative bounded. A second advantage is that none of the mappings are Euclidean translations, and all have isometric circles. For hyperbolic $B\in M(\mathbf{B})$, the isometric circle $I(B)$ is the locus $\{z:|B'(z)|=1\}$, and its interior $I(B)^\circ$ contains the repelling fixpoint, and is known as the *expanding zone*. In the context of a discrete group $\Gamma\subseteq M(\mathbf{B})$, the intersection of the exteriors of the isometric circles is a fundamental region F for Γ . If $\mathcal{R}=\Gamma\setminus\mathbf{B}$ is compact, then F has compact closure in \mathbf{B} , which is a way of saying that the expanding zones form an open cover of \mathcal{S} . This observation is the key to Lemma 2.

Proof of Lemma 1. Each $B \in M(\mathcal{B})$ has the form $z \mapsto e^{i\alpha} f_r(e^{i\beta} z)$, where

$$r = |B(0)|, f_r(z) = \frac{z+r}{1+rz}, re^{-i\beta} = -B^{-1}(0), e^{i\alpha} = B(e^{-i\beta}).$$

It is a simple matter to calculate

$$|B'(e^{i(\varphi-\beta)})| = |f_r'(e^{i\varphi})| = \frac{1-r^2}{1+2r \cos \varphi + r^2} = g_r(\varphi),$$

where g_r has the following general appearance (Figure 1):

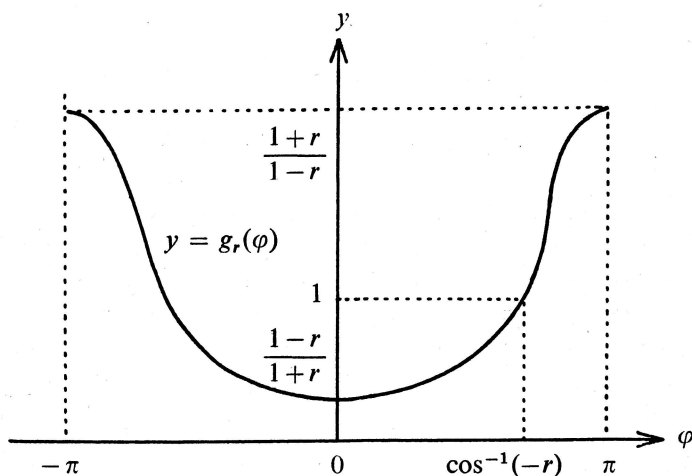


Figure 1

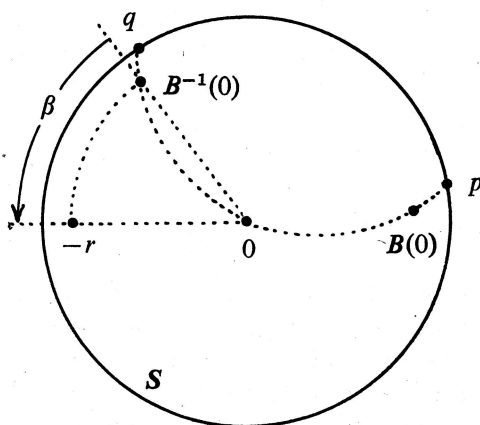


Figure 2

It follows that on intervals $\exp i[\varphi, \varphi + \delta]$ of length δ on \mathcal{S} , that $\max |f'_r|$ and $\min |f'_r|$ are assumed at opposite ends except (a) when the interval contains -1 ($\pi - \delta \leq \varphi \leq \pi$) in which case $\max = (1+r)/(1-r)$ and \min is at whichever end is farther from -1 , or (b) when the interval contains 1 ($-\delta \leq \varphi \leq 0$) in which case $\min = (1-r)/(1+r)$, and \max is at the end farther from 1 . The possibilities $-\delta/2 \leq \varphi \leq \pi - \delta/2$ are representative, and we find that the ratio

$$K(r, \varphi) = \frac{\max \{|f'_r(e^{i\theta})| : \theta \in [\varphi, \varphi + \delta]\}}{\min \{|f'_r(e^{i\psi})| : \psi \in [\varphi, \varphi + \delta]\}}$$

is increasing as $r \nearrow 1$, with limits

$$K(\varphi) = \sup_r K(r, \varphi) = \begin{cases} \frac{2}{1 + \cos(\varphi + \delta)} & \left(-\frac{\delta}{2} \leq \varphi \leq 0\right) \\ \frac{1 + \cos \varphi}{1 + \cos(\varphi + \delta)} & (0 \leq \varphi < \pi - \delta) \\ \infty & \left(\pi - \delta \leq \varphi \leq \pi - \frac{\delta}{2}\right). \end{cases}$$

It follows that $K(r, \varphi)$ is bounded as $r \nearrow 1$ if and only if φ is bounded away from $\pi - \delta$, which is really to say that $\exp i[\varphi, \varphi + \delta]$ must be bounded away from -1 .

Returning to our sequence B_n , the consequence of discreteness is that the parameters r_n must tend to 1. In the context of our $B_n|_{I_n} \in \mathcal{J}$, the interval $I_n = \exp i[\varphi_n, \varphi_n + \delta_n]$ ($\varphi_n \rightarrow \varphi_0, \delta_n \rightarrow \delta_0$) is rotated through the angle β_n onto the arc $J_n = e^{i\beta_n} I_n$ where f_{r_n} acts. In view of (BD 4) \mathcal{J} is BD if and only if $\{f_{r_n}|_{J_n}\}$ is BD, hence if and only if the intervals J_n are bounded away from -1 , or if and only if the numbers $-e^{-i\beta_n}$ cluster away from I_0 .

Now the parameters α_n, β_n, r_n for B_n , and as well the attractors $p_n = P(B_n)$, and the repellers $q_n = N(B_n)$ lie in compact sets, and we may assume them convergent with limits $\alpha_0, \beta_0, 1, p_0, q_0$. The relations are

$$e^{i\alpha_n} - e^{-i\beta_n} = r_n(p_n + q_n), \quad -e^{i(\alpha_n - \beta_n)} = p_n q_n,$$

and letting $n \rightarrow \infty$, we find

$$e^{i\alpha_0} - e^{-i\beta_0} = p_0 + q_0, \quad -e^{i(\alpha_0 - \beta_0)} = p_0 q_0,$$

from which it follows that p_0, q_0 are (in some order) $e^{i\alpha_0}$ and $-e^{-i\beta_0}$. One problem is to sort them out, but it is clear from Figure 2, that if $p_0 \neq q_0$, then $e^{i\beta_n} q_n \rightarrow -1$, and hence $q_0 = -e^{-i\beta_0}$.

If we now invoke the additional hypothesis that $B_n(I_n)$ contracts to $p \in \text{Int}(I_0)$, then $B_n(I_n)$, vastly smaller than I_n , includes some points in the expanding zone for B_n^{-1} , hence points within distance $\cos^{-1} r_n$ of p_n . It follows that $p_0 = p \in \text{Int}(I_0)$.

Now if \mathcal{J} is BD, the numbers $-e^{-i\beta_n}$ cluster outside I_0 , and hence $-e^{-i\beta_0} \notin I_0$. Hence p_0 must be $e^{i\alpha_0}$ and the repellers $q_n \rightarrow q_0 = -e^{-i\beta_0} \notin I_0$. Conversely, if the q_n

cluster outside I_0 , then q_n, p_n are bounded away from each other, and therefore $q_0 = \lim -e^{-i\beta_n}$, the latter numbers being therefore eventually bounded away from I_0 , and \mathcal{J} is BD. \square

Proof of Lemma 2. As we have observed, \mathcal{S} is covered by the expanding zones for $\{B \in \Gamma\}$. Extract a finite subcover, and by elementary continuity considerations, obtain a finite cover of \mathcal{S} by closed intervals $\{I_j: j \in F\}$, with associated mappings $A_j \in \Gamma$, such that there are numbers c, δ with the properties:

- (i) $|A'_j(z)| \geq c > 1$ for $z \in I_j$, and
- (ii) every interval of length δ lies entirely inside some I_j .

We now take an interval $J_0 = [p, p_1]$ of length $< \delta$ and find $j_1 \in F$ such that $J_0 \subset I_{j_1}$, and consider $J_1 = A_{j_1}(J_0)$. Either J_1 has length more than δ , or it lies in I_{j_2} (some $j_2 \in F$). This process eventually terminates (say at stage n : $|J_{n-1}| < \delta \leq |J_n|$) because lengths are expanded by a uniform factor c at every stage. Having reached the conclusion, we replace J_0 by a subinterval $[p, p_2]$ and relabel the successive intervals so that J_n has length *exactly* δ . We then have a uniform bound on the total lengths. Indeed, from $|J_{k+p}| \geq c^p |J_k|$, we find

$$\sum_{k=0}^n |J_k| \leq \sum_{p=0}^n \left(\frac{1}{c}\right)^p \delta < \frac{c\delta}{c-1}.$$

We shall denote the mapping $A_{j_n} A_{j_{n-1}} \dots A_{j_2} A_{j_1}$ simply by A , and the final interval $J_n = A(J_0)$ by I . Considering that our factors come from the finite pool $\{A_j: j \in F\}$, there is a uniform upper bound M for $\{|A'_j(z)|: j \in F, z \in \mathcal{S}\}$, and of course we still have $|A'_j(z)| \geq c$ for all z under consideration. Now suppose $z_1, w_1 \in J_0$. Define z_i, w_i inductively by $z_{i+1} = A_{j_i}(z_i)$, $w_{i+1} = A_{j_i}(w_i)$. We estimate:

$$\begin{aligned} \log \frac{|A'(z_1)|}{|A'(w_1)|} &= \sum_{i=1}^n \log \frac{|A'_{j_i}(z_i)|}{|A'_{j_i}(w_i)|} = \sum_{i=1}^n \operatorname{Re} \left\{ \int_{w_i}^{z_i} \frac{A''_{j_i}(z)}{A'_{j_i}(z)} dz \right\} \\ &\leq \sum_{i=1}^n \int_{J_i} \frac{|A''_{j_i}(z)|}{|A'_{j_i}(z)|} |dz| \leq \frac{M}{c} \sum_{i=0}^{n-1} |J_i| < \frac{M\delta}{c-1}. \end{aligned}$$

This argument is a close adaptation of Lemma 1 in [10]. It seemed that no economies were to be gained by the reference, since the set-up required explanation in any case.

If now in place of J_0 , we have a sequence of intervals $J_m = [p, p_m]$ ($p_m \rightarrow p$), we obtain by this process a sequence $A^{(m)} \in \Gamma$, with final interval I_m of uniform length δ . It is clear that we may require that I_m be convergent upon a nondegenerate interval of length δ . Thus it only remains to see that the sequence $\{A^{(m)}|_{J_m}\}$ is BD, but this follows from the estimate above, which is independent of the length of the original interval J_0 and the number of factors in the word A . \square

Added October 1983: Since the submission of this article several months ago, I have received a preprint "Rigidity theorems for Möbius groups" by Pekka Tukia. He has closed the gap between M_r and the complement of the parabolic fixpoints

by showing that M , quite generally includes the so-called "radial set", or limit set of conical approach; moreover, for $n=1$ and Γ of finite covolume, the radial set *excludes only* parabolic fixpoints. These results taken together make a theorem more general than the present Theorem 3.

References

- [1] AGARD, S.: A geometric proof of Mostow's rigidity theorem for groups of divergence type. - Acta Math. (to appear).
- [2] BOWEN, R.: Hausdorff dimension of quasicircles. - Inst. Hautes Études Sci. Publ. Math. 50, 1979, 11—25.
- [3] KUUSALO, T.: Boundary mappings of geometric isomorphisms of Fuchsian groups. - Ann. Acad. Sci. Fenn. Ser. A I Math. 545, 1973, 1—7.
- [4] MOSTOW, G. D.: Quasiconformal mappings in n -space and the rigidity of hyperbolic space forms. - Inst. Hautes Études Sci. Publ. Math. 34, 1968, 53—104.
- [5] MOSTOW, G. D.: Strong rigidity of locally symmetric spaces. - Annals of Mathematics Studies 78. Princeton University Press, Princeton, New Jersey, 1973.
- [6] MYRBERG, P. J.: Ein Approximationssatz für die Fuchsschen Gruppen. - Acta Math. 57, 1931, 389—409.
- [7] SHEINGORN, M.: Transitivity for the modular group. - Math. Proc. Cambridge Philos. Soc. 88, 1980, 409—423.
- [8] SORVALI, T.: The boundary mapping induced by an isomorphism of covering groups. - Ann. Acad. Sci. Fenn. Ser. A I Math. 526, 1972, 1—30.
- [9] SULLIVAN, D.: On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions. - Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference, edited by I. Kra and B. Maskit, Annals of Mathematics Studies 97. Princeton University Press, Princeton, New Jersey, 1981, 465—496.
- [10] SULLIVAN, D.: Conformal dynamical systems. - Preprint, 1982.
- [11] TUKIA, P.: Multiplier preserving isomorphisms between Möbius groups. - Ann. Acad. Sci. Fenn. Ser. A I Math. 1, 1975, 327—341.

University of Minnesota
School of Mathematics
Minneapolis, Minnesota 55455
USA

Received 24 March 1983