# ON HYPERBOLIC ONCE-PUNCTURED-TORUS BUNDLES

James W. Cannon* and Warren Dicks**

September 9, 2002

Abstract. For a hyperbolic once-punctured-torus bundle over a circle, a choice of normalization determines a family of arcs in the Riemann sphere. We show that, in each arc in the family, the set of cusps is dense and forms a single orbit of a finitely generated semigroup of Möbius transformations. This was previously known for the case of the complement of the figure-eight knot.

## 0. General summary.

Let $\mathbf{H}^3$ denote hyperbolic three-space, $\partial\mathbf{H}^3$ the boundary of $\mathbf{H}^3$, and $\overline{\mathbb{C}}$ the Riemann sphere, $\mathbb{C} \cup \{\infty\}$. Let $\mathrm{Iso}(\mathbf{H}^3)$ denote the group of orientation-preserving isometries of $\mathbf{H}^3$, acting on $\mathbf{H}^3 \cup \partial\mathbf{H}^3$. Let $\mathrm{PSL}_2(\mathbb{C})$ act on $\overline{\mathbb{C}}$ as the group of Möbius transformations. We identify $\partial\mathbf{H}^3 = \overline{\mathbb{C}}$ and $\mathrm{Iso}(\mathbf{H}^3) = \mathrm{PSL}_2(\mathbb{C})$ in a compatible way.

Let $M$ be a complete, finite-volume, hyperbolic, once-punctured-torus bundle over a circle. The fibration of $M$ over the circle lifts to the universal covers, giving rise to a fibration of $\mathbf{H}^3$ over $\mathbb{R}$. Each fiber in $\mathbf{H}^3$ has a spine which is a tree, and McMullen [16] has shown that the ends of this tree reach every point of $\partial\mathbf{H}^3 = \overline{\mathbb{C}}$. Deleting a single edge from this tree leaves two complementary subtrees lying in $\mathbf{H}^3$. We prove that the set of points in $\overline{\mathbb{C}}$ which are reached by the ends of one of these subtrees forms a closed disk embedded in $\overline{\mathbb{C}}$, so bounded by a Jordan curve; moreover, the two complementary subtrees determine two complementary disks, so determine the same Jordan curve. On such a curve, two of the points will be found to be distinguished in a natural way, so the curve can be viewed as the union of two arcs with common endpoints. Our main result is that the set of cusps in any such arc is dense and forms a single orbit of a finitely generated semigroup of Möbius transformations. For the case of the complement of the figure-eight knot, this was seen in [1].

The hyperbolic structure on $M$ is unique, by Mostow rigidity, and corresponds to a discrete faithful representation of $\pi_1(M)$ in $\mathrm{PSL}_2(\mathbb{C})$ which is unique up to group conjugation and complex conjugation, so we have an action of $\pi_1(M)$ on the Riemann sphere $\overline{\mathbb{C}}$. A construction given in [7] in a more general context, and described in our context algebraically in Section 3 below, produces a topological two-sphere with an action by $\pi_1(M)$ called "the model". Throughout, we work with the model from a highly algebraic viewpoint. Bowditch [3, Theorem 9.1] has recently shown that the model and $\overline{\mathbb{C}}$ are homeomorphic as $\pi_1(M)$-spaces. See also [6].

*Partially supported by the NSF (USA) through Grants DMS-9506725 and DMS-9803868

**Partially partially supported by the DGES and the DGI (Spain) through grants PB96-1152 and BFM2000-0354, respectively.

In Section 1, we describe the algebraic and topological results we obtain. Section 2 recalls some group theory, and Section 3 some topology. Sections 4, 5, and 7 contain the arguments, while Section 6 gives a detailed statement of the technical results obtained, and this may be useful for possible applications, such as programming. In Section 8, we observe that work of Nielsen, Jørgensen, Thurston, Bowditch, and others, can be used to translate our results into the geometric results stated above.

## 1. Summary of the algebraic and topological results.

"A journey of a thousand miles began with a single step."

- Lao-Tsze

Let $\langle A, B, C\rangle$ denote the group with presentation $\langle A, B, C \mid A^2 = B^2 = C^2 = 1\rangle$. Let $\mathcal{E}$ denote the set of ends of $\langle A, B, C\rangle$, that is, the set of right infinite, reduced words in $A$, $B$, $C$. For any $G \in \langle A, B, C\rangle$, we let $[G]_{\mathcal{E}}$ denote the set of elements of $\mathcal{E}$ which have $G$ as an initial segment, and let $[G^*]_{\mathcal{E}}$ denote the complement of $[G]_{\mathcal{E}}$ in $\mathcal{E}$. For example, $[A]_{\mathcal{E}}$ is the set of right infinite, reduced words which start with $A$ (or infinite journeys which begin with the single step $A$). Then $\{[G]_{\mathcal{E}} \mid G \in \langle A, B, C\rangle\}$ is the basis of clopen sets of a topology on $\mathcal{E}$, so $\mathcal{E}$ is a compact, Hausdorff, totally disconnected topological space. In a natural way, $\mathcal{E}$ is an $\langle A, B, C\rangle$-*space*, that is, $\langle A, B, C\rangle$ acts by homeomorphisms on $\mathcal{E}$, namely by multiplication; here, and throughout, actions are on the left. It is easy to see that every $\langle A, B, C\rangle$-orbit is dense in $\mathcal{E}$.

Let us briefly recall "how to make $ABC$ parabolic", where *parabolic* is taken to mean having a unique fixed point; details can be found in [1, Section 4], for example. Let $\mathcal{E}^{ABC}$ denote the set of points of $\mathcal{E}$ fixed by $ABC$, so $\mathcal{E}^{ABC} = \{(ABC)^\infty, (CBA)^\infty\}$. Let $\sim$ denote the (closed) equivalence relation on $\mathcal{E}$ generated by

$$GE_1 \sim GE_2 \text{ for all } G \in \langle A, B, C\rangle, E_1, E_2 \in \mathcal{E}^{ABC}.$$

Here "closed" has the obvious sense, that is, if $a_n \sim b_n$, and $a_n \to a$, and $b_n \to b$, then $a \sim b$. The quotient space $\mathcal{E}/{\sim}$ is homeomorphic to a circle, and we shall denote it by $S^1$; see Section 3 and Fig. 3, below.

Clearly $S^1$ is an $\langle A, B, C\rangle$-space, and $ABC$ is parabolic on $S^1$. Each element of $S^1$ is a subset of $\mathcal{E}$ consisting of either one or two elements; an element of $S^1$ which consists of two elements is called a *cusp* of $S^1$. The set of cusps is denoted $\text{cusps}(S^1)$; it is dense, and forms a single $\langle A, B, C\rangle$-orbit. In $S^1$, set

$$\begin{aligned} 0_{S^1} &:= \{(ACB)^\infty, (BCA)^\infty\}, \\ 1_{S^1} &:= \{(BAC)^\infty, (CAB)^\infty\}, \\ \infty_{S^1} &:= \{(CBA)^\infty, (ABC)^\infty\}, \end{aligned}$$

so $0_{S^1}$, $1_{S^1}$ and $\infty_{S^1}$ are three distinguished cusps in $S^1$. For any $G \in \langle A, B, C\rangle$, we let $[G]_{S^1}$ denote the image of $[G]_{\mathcal{E}}$ in $S^1$, and similarly for $[G^*]_{S^1}$. If $G \neq 1$ then $[G]_{S^1}$ is homeomorphic to a compact interval, and, if we write $G = HZ$ where $Z$ is the last letter of $G$, then $[G]_{S^1} = H([Z]_{S^1})$. We orient $S^1$ so that $[A]_{S^1}$ starts at $\infty_{S^1}$ and finishes at $0_{S^1}$, and hence $[B]_{S^1}$ starts at $0_{S^1}$ and finishes at $1_{S^1}$, while $[C]_{S^1}$ starts at $1_{S^1}$ and finishes at $\infty_{S^1}$. Moreover, the starting point of $[G]_{S^1}$ is obtained by applying $H$ to the starting point of $[Z]_{S^1}$, and similarly for the finishing point.

Let $\mathrm{Aut}\langle A,B,C\rangle$ denote the group of automorphisms of $\langle A,B,C\rangle$. In a natural way, $\mathcal{E}$ and $S^1$ and $\mathrm{cusps}(S^1)$ are $\mathrm{Aut}\langle A,B,C\rangle$-spaces.

For any $T \in \mathrm{Aut}\langle A,B,C\rangle$, we shall use the triple $(T(A),T(B),T(C))$ to denote $T$. Let

$$R = (A, BCB, B), \qquad L = (B, BAB, C),$$

and let $\langle R,L\rangle$ denote the subgroup of $\mathrm{Aut}\langle A,B,C\rangle$ generated by $R$ and $L$.

We are interested in the more sophisticated construction, implicit in [7], of "making a hyperbolic element of $\mathrm{Aut}\langle A,B,C\rangle$ parabolic". Let $p, a_1, b_1, \ldots, a_p, b_p$ be positive integers and let $F\!:= \prod_{i=1}^{p}(R^{a_i}L^{b_i})$, that is, $F = R^{a_1}L^{b_1}\cdots R^{a_p}L^{b_p} \in \langle\langle R,L\rangle\rangle$, where $\langle\langle R,L\rangle\rangle$ denotes the subsemigroup of $\langle R,L\rangle$ generated by $R$ and $L$. The subgroup $\langle A,B,C,F\rangle$ of $\mathrm{Aut}\langle A,B,C\rangle$ is a semidirect product, $\langle A,B,C\rangle \rtimes \langle F\rangle$.

In Section 2, we shall recall the definition of a hyperbolic automorphism of $\langle A,B,C\rangle$, and see that $F$ is hyperbolic, and that, moreover, up to squaring, conjugation in $\mathrm{Aut}\langle A,B,C\rangle$, and composition with an inner automorphism, every hyperbolic automorphism can be written in this form.

Let $\mathcal{E}^F$ denote the set of points of $\mathcal{E}$ fixed by $F$. The normalization of $F$ was chosen to ensure that $\mathcal{E}^F$ is large enough for our purposes. In Section 3, we define an equivalence relation $\approx$ on $\mathcal{E}$, and in Section 4 we see that $\approx$ is the *closed* equivalence relation generated by

$$GE_1 \approx GE_2 \text{ for all } G \in \langle A,B,C\rangle, E_1, E_2 \in \mathcal{E}^F.$$

In [7], as is recalled in Section 3 below, it is shown that the quotient space $\mathcal{E}/{\approx}$ is homeomorphic to a two-sphere; we shall denote it by $S^2$, or $S^2_F$ if there is a need to avoid ambiguity. Moreover, $S^2$ is an $\langle A,B,C,F\rangle$-space, and $F$ is parabolic on $S^2$. Also $ABC$ is parabolic on $S^2$, and there is an induced quotient map $S^1 \to S^2$ of $\langle A,B,C,F\rangle$-spaces; we will sometimes refer to this map as *the Peano curve associated to* $F$. The image of a cusp of $S^1$ will be called a *cusp* of $S^2$, and the set of cusps of $S^2$ will be denoted $\mathrm{cusps}(S^2)$; it is a dense $\langle A,B,C,F\rangle$-subspace of $S^2$, and forms a single $\langle A,B,C\rangle$-orbit. The map $\mathrm{cusps}(S^1) \to \mathrm{cusps}(S^2)$ is bijective. The images of $\infty_{S^1}$, $0_{S^1}$ and $1_{S^1}$ in $S^2$ will be denoted $\infty_{S^2}$, $0_{S^2}$ and $1_{S^2}$, respectively. We shall see that $\infty_{S^2} = \mathcal{E}^F$.

For any $G \in \langle A,B,C\rangle$, we let $[G]_{S^2}$ denote the image of $[G]_{\mathcal{E}}$ (or, equivalently, of $[G]_{S^1}$) in $S^2$, and similarly for $[G^*]_{S^2}$. In Section 7, we shall see that if $G \neq 1$ then $[G]_{S^2}$ is homeomorphic to a closed disk, and the boundary is a Jordan curve. We denote this boundary by $\partial[G]_{S^2}$.

Thus the Jordan disk $[G]_{S^2}$ is the image of the interval $[G]_{S^1}$ under the quotient map $S^1 \to S^2$. In general, intervals in $S^1$ have images which are unions of "consecutive" Jordan disks joined together at various points, and these need not be simply connected. For example, in the upper part of the fourth picture in [1, Fig. 17] there are three consecutive disks pairwise joined together at one point.

We shall see that $[A]_{S^2} \cap [B]_{S^2}$, $[B]_{S^2} \cap [C]_{S^2}$, $[C]_{S^2} \cap [A]_{S^2}$ are arcs with endpoints $1_{S^2}$ and $\infty_{S^2}$, forming an embedded graph with three edges and two vertices of valence three. See Fig. 5. The complement of this graph has three components, and these are the interiors of $[A]_{S^2}$, $[B]_{S^2}$, and $[C]_{S^2}$. The three regions $[A]_{S^2}$, $[B]_{S^2}$, and $[C]_{S^2}$ are successively broken up by the $\langle A,B,C\rangle$-images of the three arcs; thus, for example, if $G$ ends in $A$, then $GA$ applied to $[B]_{S^2} \cap [C]_{S^2}$ gives an arc dividing $[G]_{S^2}$ into $[GB]_{S^2}$ and $[GC]_{S^2}$, and so on.

It transpires that $0_{S^2}$, $\infty_{S^2}$, and $1_{S^2}$ all lie in $\partial[B]_{S^2}$. We orient $S^2$ so that when travelling along $\partial[B]_{S^2}$ with $[B]_{S^2}$ on the right, starting at $0_{S^2}$ one reaches $\infty_{S^2}$ before $1_{S^2}$.

The image in $[G]_{S^2}$ of the starting (resp. finishing) point of $[G]_{S^1}$ is called the *starting* (resp. *finishing*) point of $[G]_{S^2}$. It is easy to see that the starting and finishing points lie on the boundary, so they divide $\partial[G]_{S^2}$ into two oriented arcs joining the starting point to the finishing point, and the oriented arc that has $[G]_{S^2}$ to its right (resp. left) will be denoted $\partial^+[G]_{S^2}$ (resp. $\partial^-[G]_{S^2}$). For each such oriented arc, we find a finitely generated subsemigroup $\mathcal{M}$ of $\langle A, B, C, F\rangle$ acting on the arc, such that the cusps in the arc are dense and consist of the $\mathcal{M}$-orbit of each endpoint of the arc. Moreover we show that the action of $\mathcal{M}$ on the arc is modeled by an affine action of $\mathcal{M}$ on a closed bounded interval in the real line, obtained by restricting an affine action of $\langle A, B, C, F\rangle$ on the real line.

It suffices to do this for each of the six oriented arcs $\partial^+[A]_{S^2}$, $\partial^+[B]_{S^2}$, $\partial^+[C]_{S^2}$, $\partial^-[A]_{S^2}$, $\partial^-[B]_{S^2}$, $\partial^-[C]_{S^2}$, and then apply $\langle A, B, C\rangle$. We will find that $\partial^+[B]_{S^2}$ is obtained by concatenating $\partial^+[C]_{S^2}$ and $\partial^+[A]_{S^2}$, and reversing the orientation. Similarly, $\partial^-[A]_{S^2}$ is obtained by concatenating $\partial^-[B]_{S^2}$ and $\partial^-[C]_{S^2}$ and reversing the orientation. Also, $[B]_{S^2}\cap[C]_{S^2} = \partial^+[C]_{S^2}$, and $[C]_{S^2}\cap[A]_{S^2} = \partial^-[C]_{S^2}$, while $[A]_{S^2}\cap[B]_{S^2}$ is obtained by concatenating $\partial^+[A]_{S^2}$ and $\partial^-[B]_{S^2}$. See Fig. 5.

A detailed description of the six arcs and their associated semigroups can be found in Section 6.

## 2. Group theory.

For the next two subsections, we temporarily forget all the notation of Section 1.

**2.1 Definitions.** Let $\mathrm{F}_2$ denote the free group of rank two, with specified free generators denoted $X$, $Y$. We will refer to the element $XY\bar{X}\bar{Y}$ as "the commutator". Here, and throughout, an overline denotes the inverse.

Let $\mathrm{Aut}(\mathrm{F}_2)$ denote the automorphism group of $\mathrm{F}_2$. We identify $\mathrm{F}_2$ with the subgroup of $\mathrm{Aut}(\mathrm{F}_2)$ consisting of the inner automorphisms.

Any $T \in \mathrm{Aut}(\mathrm{F}_2)$ is said to be *positive with respect to $X$ and $Y$* if $T$ carries the semigroup $\langle\langle X, Y\rangle\rangle$ into itself; similar terminology applies with any free generating set of $\mathrm{F}_2$.

The three elements

$$R\colon (X, Y) \mapsto (X, YX), \quad L\colon (X, Y) \mapsto (XY, Y), \quad M\colon (X, Y) \mapsto (Y, X)$$

of $\mathrm{Aut}(\mathrm{F}_2)$ are positive with respect to $X$ and $Y$. Both $R$ and $L$ fix the commutator, while $M$ inverts it. Also $M$ has order two, and conjugates $R$ to $L$.

Let $\mathrm{F}_2'$ denote the derived subgroup of $\mathrm{F}_2$. We identify the abelianization $\mathrm{F}_2/\mathrm{F}_2'$ with $\mathbb{Z}^2$, with $X$ and $Y$ mapping to $(1, 0)$ and $(0, 1)$ respectively. Since $\mathrm{Aut}(\mathbb{Z}^2) = \mathrm{GL}_2(\mathbb{Z})$, we have a natural map

$$\mathrm{matrix}\colon \mathrm{Aut}(\mathrm{F}_2) \to \mathrm{GL}_2(\mathbb{Z}). \tag{2.1.1}$$

Here

$$\mathrm{matrix}(R) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \mathrm{matrix}(L) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \mathrm{matrix}(M) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

By the Euclidean algorithm, $\mathrm{matrix}(\langle R, L\rangle) = \mathrm{SL}_2(\mathbb{Z})$, and hence $\mathrm{matrix}(\langle R, M\rangle) = \mathrm{GL}_2(\mathbb{Z})$.

Let

$$Q\colon= L\bar{R}, \quad P\colon= L\bar{R}L,$$

so $\langle R, L\rangle = \langle P, Q\rangle$. One can check that $L\bar{R}L = \bar{R}L\bar{R}$, $P^2 = Q^3$, and $P^4 = Q^6 = XY\bar{X}\bar{Y}$. Here $P^2 = Q^3$ is central, and the map (2.1.1) kills its square. It is well-known that this now gives a complete set of defining relations for $\mathrm{SL}_2(\mathbb{Z})$. It follows that both the kernel of (2.1.1) and $\mathrm{F}_2$ meet $\langle R, L\rangle$ in the infinite cyclic, central subgroup $\langle XY\bar{X}\bar{Y}\rangle$. It follows also that $\langle R, L\rangle$ has presentations $\langle P, Q \mid P^2 = Q^3\rangle$ and $\langle R, L \mid L\bar{R}L = \bar{R}L\bar{R}\rangle$ (so is isomorphic to the braid group on three strings). Notice that the semidirect product $\mathrm{F}_2 \rtimes \langle R, M\rangle$ maps to $\mathrm{Aut}(\mathrm{F}_2)$ with infinite cyclic kernel.

Nielsen [21] showed that $\mathrm{Aut}(\mathrm{F}_2)$ is generated by what are now called the Nielsen transformations. Moreover, the three Nielsen transformations $M$, $\bar{X}\bar{P}M$ and $R$ suffice, and thus $X$, $Y$, $R$, $L$ and $M$ also form a generating set.

Notice that the kernel of (2.1.1) must be precisely $\mathrm{F}_2$, since no unexplained relations get imposed on $R$ and $L$. We have seen that (2.1.1) is surjective.

Other consequences are that every element of $\mathrm{Aut}(\mathrm{F}_2)$ sends the commutator to a conjugate of itself or its inverse, and that $\langle R, L\rangle$ is the subgroup of $\mathrm{Aut}(\mathrm{F}_2)$ which fixes the commutator. Hence, $\langle R, M\rangle$ is the subgroup of $\mathrm{Aut}(\mathrm{F}_2)$ which stabilizes $\langle XY\bar{X}\bar{Y}\rangle$.

In $\langle R, L\rangle$, since $P^2 = Q^3$ is central, conjugating by $P$ is the same as conjugating by $\bar{P}$, and, since $LPR = P$, so $PR\bar{P} = \bar{L}$ and $PL\bar{P} = \bar{R}$. For $T \in \langle R, L\rangle$, we define the *transpose* of $T$ to be $T^{\mathrm{tr}}\colon= P\bar{T}\bar{P} = \bar{P}\bar{T}P$.

Let $J$ be an element of $\mathrm{GL}_2(\mathbb{Z})$. Recall that $J$ is said to be *hyperbolic* if the eigenvalues of $J$ are not roots of unity. And $J$ is *parabolic* if the eigenvalues of $J$ are $\pm 1$ but $J^2$ is not the identity matrix, $I$. And $J$ is *elliptic* if some proper power of $J$ equals $I$. If $J \in \mathrm{SL}_2(\mathbb{Z}) - \{\pm I\}$, then $J$ is hyperbolic, parabolic or elliptic, depending as the absolute value of the trace of $J$ is greater than 2, equal to 2, or less than 2, respectively.

For any $T \in \mathrm{Aut}(\mathrm{F}_2)$, we apply all conjugate-invariant attributes of $\mathrm{matrix}(T)$ to $T$ itself, and speak of the *eigenvalues*, *trace*, *determinant*, *hyperbolicity*, etc. of $T$. Also, we let $T$ act on the real projective line $\overline{\mathbb{R}}\colon= \mathbb{R} \cup \{\infty\}$ as the Möbius transformation associated to $\mathrm{matrix}(T)$. Thus $R$ acts by $R(x) = x+1$, and $L$ acts by $L(x) = \frac{x}{x+1} = \frac{1}{1+\frac{1}{x}}$. Hence $L$ and $R$ map $[0, \infty]$ to its left and right 'halves', $[0, 1]$ and $[1, \infty]$, respectively, thus motivating the $RL$-notation. (Many authors take the view from the upper half of the complex plane and interchange 'left' and 'right'.)

In $\mathrm{Aut}(\mathrm{F}_2)$, define

$$B\colon (X, Y) \mapsto (\bar{X}, \bar{Y}), \quad A\colon (X, Y) \mapsto (Y\bar{X}\bar{Y}, \bar{Y}), \quad C\colon (X, Y) \mapsto (\bar{X}, X\bar{Y}\bar{X}).$$

Then $CB = X$, $AB = Y$, and the subgroup $\langle A, B, C\rangle$ of $\mathrm{Aut}(\mathrm{F}_2)$ has the presentation

$$\mathrm{F}_2 \rtimes \{1, B\} = \langle A, B, C \mid A^2 = B^2 = C^2 = 1\rangle.$$

Since $\mathrm{matrix}(\langle A, B, C\rangle)$ is the center of $\mathrm{GL}_2(\mathbb{Z})$, $\langle A, B, C\rangle$ is normal in $\mathrm{Aut}(\mathrm{F}_2)$. Thus $\mathrm{Aut}(\mathrm{F}_2)$ can be viewed as a subgroup of $\mathrm{Aut}\langle A, B, C\rangle$. But $\mathrm{F}_2$ is a characteristic subgroup of $\langle A, B, C\rangle$, so $\mathrm{Aut}(\mathrm{F}_2) = \mathrm{Aut}\langle A, B, C\rangle$.

One can check that $P^2 = CBA$ so $Q^3 = P^2 = CBA$ and $(CBA)^2 = XY\bar{X}\bar{Y}$. □

**2.2 Proposition.** *Let $T$ be a hyperbolic element of* $\mathrm{Aut}(\mathrm{F}_2)$.

*Then $T^2$ has positive trace and determinant.*

*If $T$ has positive trace and determinant, then there exist $G \in \mathrm{F}_2$, $\Phi \in \langle R, L\rangle$, and positive integers $p$, $a_1, b_1, \dots a_p, b_p$ such that $G\Phi T\bar{\Phi} = R^{a_1}L^{b_1}\cdots R^{a_p}L^{b_p}$.*

*Proof.* Clearly $\det(T^2) = (\det(T))^2 = (\pm 1)^2 = 1$. By the Cayley-Hamilton theorem,

$$\mathrm{trace}(T^2) = (\mathrm{trace}(T))^2 - 2(\det(T)),$$

and, since $T$ is hyperbolic, we see that $T^2$ has positive trace.

We now assume that $T$ has positive trace and determinant.

By the above-mentioned results of Nielsen's, by replacing $T$ with $T$ composed with a suitable element of $\mathrm{F}_2$, we may assume that $T$ fixes the commutator, and work in $\langle R, L\rangle = \langle P, Q\rangle$ for the remainder of the proof. Since $P^2 = Q^3 = CBA$ is central, $T$ is the product of an integral power of $CBA$ and an alternating product which alternates $\bar{Q}$'s and $\bar{Q}^2$'s with $P$'s.

Since $T$ is not elliptic, it does not lie in the group generated by $CBA$ and $P$, nor in the group generated by $CBA$ and $Q$. Hence, by conjugating $T$ by an element of $\langle P, Q\rangle = \langle R, L\rangle$, we may assume that the alternating product starts with $\bar{Q}$ or $\bar{Q}^2$ and ends with $P$. Thus $T$ is the product of an integral power of $CBA$ and an element of the semigroup generated by $R = \bar{Q}^2P$ and $L = \bar{Q}P$.

Since the trace of $T$ is positive, $T$ is the product of a power of $(CBA)^2 = XY\bar{X}\bar{Y}$ and an element of $\langle\langle R, L\rangle\rangle$. By composing with a power of $XY\bar{X}\bar{Y}$, we may assume that $T$ lies in $\langle\langle R, L\rangle\rangle$. Since $T$ is not parabolic, it is not a power of $R$ nor of $L$. By conjugating $T$ by a power of $R$ or $L$, we may assume $T = R^{a_1}L^{b_1}\cdots R^{a_p}L^{b_p}$, where $p$ and the $a_i$, $b_i$ are positive integers. □

We now restore the notation of Section 1, which we have seen is compatible with the notation of Definitions 2.1.

**2.3 Notation.** Henceforth $\mathrm{F}_2$ will be viewed as the unique torsion-free subgroup of index two in $\langle A, B, C\rangle$, consisting of words of even length. The preferred basis is $X = CB$, $Y = AB$.

We identify $\mathrm{Aut}(\mathrm{F}_2)$ with $\mathrm{Aut}\langle A, B, C\rangle$, and, hence, all of the terminology for elements of the former apply to elements of the latter.

We can view $\mathcal{E}$, the set of ends of $\langle A, B, C\rangle$, as the set of ends of $\mathrm{F}_2$, that is, right infinite words in $X^{\pm 1}$, $Y^{\pm 1}$, which are reduced in the sense that no symbol is followed by its inverse. By bounded cancellation, $\mathrm{Aut}(\mathrm{F}_2) = \mathrm{Aut}\langle A, B, C\rangle$ acts continuously on $\mathcal{E}$; see [8].

We have

$$R = (A, BCB, B), \qquad L = (B, BAB, C), \qquad M = (C, B, A).$$

It follows from the results of Nielsen recalled in Definition 2.1 that the stabilizer of $ABC$ in $\mathrm{Aut}\langle A, B, C\rangle$ is $\langle R, L\rangle$. (Steve Humphries has pointed out that this can be proved directly by a standard argument used for braid groups; see, for example, the proof of Proposition 10.7 in [4].) Hence the stabilizer of $\langle ABC\rangle$ is $\langle R, M\rangle$.

Let $p, a_1, b_1, \dots, a_p, b_p$ be positive integers and let $F\colon= \prod_{i=1}^{p}(R^{a_i}L^{b_i})$, that is, $F = R^{a_1}L^{b_1}\cdots R^{a_p}L^{b_p}$.

Let $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ denote matrix($F$), so $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \prod_{i=1}^{p}(\begin{pmatrix} 1 & a_i \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ b_i & 1 \end{pmatrix})$.

In $\mathbb{R}$, let

$$\lambda := \frac{a+d+\sqrt{(a+d)^2-4}}{2}, \quad \mu_+ := \frac{\lambda - d}{c}, \quad \mu_- := \frac{a-\lambda}{c},$$

so $\mu_+$ and $\mu_-$ are fixed by $F$ acting as a Möbius transformation on $\overline{\mathbb{R}}$. In Proposition 2.5, we shall show that $\mu_+ > 1$ and $-1 < \mu_- < 0$.

For $x \in \mathbb{R}$ and $n \in \mathbb{Z}$, let $[x]$ denote the greatest integer in the interval $(-\infty, x]$, and let $[x]_n := [nx] - [(n-1)x]$, so we have a sequence $\mathbb{N} - \{0\} \to \{[x], [x]+1\}$, $n \mapsto [x]_n$.

In $\mathcal{E}$, let

$$E^+ := \prod_{n=1}^{\infty}((BC)^{[\mu_+]_n}BA), \quad E^- := \prod_{n=1}^{\infty}(CA(BA)^{[\frac{1+\mu_-}{-\mu_-}]_n}). \quad \square$$

In Theorem 4.1, we shall see that $E^+ = F^\infty(B)$ and $E^- = \bar{F}^\infty(C)$ are fixed by $F$, and in Theorem 4.2, we shall show that

$$\mathcal{E}^F = \{(ABC)^\infty, (CBA)^\infty\} \cup \langle A, B, C\rangle E^+ \cup \langle A, B, C\rangle E^-.$$

**2.4 Example.** Here, and intermittently throughout, for concreteness, we consider the case

$$F = RL^3 = (BCBABCBABCB, BCBABCBABCBABCB, B).$$

It is straightforward to calculate matrix($F$) $= \begin{pmatrix} 4 & 1 \\ 3 & 1 \end{pmatrix}$, $\lambda = \frac{5+\sqrt{21}}{2}$, $\mu_+ = \frac{3+\sqrt{21}}{6}$,

$$F : (BC, BA) \mapsto (BCBABCBABCBABC, BCBA),$$

$([\frac{3+\sqrt{21}}{6}]_n \mid n \geq 1) = (1,1,1,2,1,1,1,2,1,1,1,2,1,1,1,2,1,1,2,1,\dots)$, and

$$E^+ = F^\infty(B) = BCBABCBABCBABCBCBABCBABCBABCBABCBCBA\cdots$$

Also $\bar{F} = (ABABABA, C, CABABAC)$, $\mu_- = \frac{3-\sqrt{21}}{6}$, $\frac{1+\mu_-}{-\mu_-} = \frac{1+\sqrt{21}}{2}$,

$$\bar{F} : (CA, BA) \mapsto (CABABACABABABA, CABABABA),$$

$([\frac{1+\sqrt{21}}{2}]_n \mid n \geq 1) = (2,3,3,3,2,3,3,3,3,2,3,3,3,3,2,3,3,3,3,2,\dots)$, and

$$E^- = \bar{F}^\infty(C) = CABABACABABABACABABABACABABABACABABAC\cdots$$

We will use this information when we refer to the case $F = RL^3$ in the sequel. $\square$

We record the following.

**2.5 Proposition.** *With Notation 2.3,*

$$\mu_+ = F(\mu_+) \in F[0,\infty] \subseteq R[0,\infty] = [1,\infty],$$
$$\mu_- = \bar{F}(\mu_-) \in \bar{F}[-\infty,0] \subseteq \bar{L}[-\infty,0] = [-1,0]. \quad \square$$

**2.6 Digression.** Let $\mathrm{IrrEnds}\langle\langle R, L\rangle\rangle$ denote the set of right infinite "irrational" words in $R$ and $L$, that is, right infinite words of the form $\prod_{i=1}^{\infty}(R^{c_i}L^{d_i})$ where the $c_i$, $d_i$ are positive integers, except for $c_1$ which is a non-negative integer. Consider the map

$$\begin{aligned}&\mathrm{IrrEnds}\langle\langle R, L\rangle\rangle \to [0,\infty],\\ &\prod_{i=1}^{\infty} R^{c_i}L^{d_i} \mapsto [c_1, d_1, c_2, d_2, \dots] := c_1 + \frac{1}{d_1 + \frac{1}{c_2+\dots}}.\end{aligned}$$

It is straightforward to show that the above map respects the action of $\langle\langle R, L\rangle\rangle$ on $[0,\infty]$ by Möbius transformations. Hence $F$ fixes $F^\infty \in \mathrm{IrrEnds}\langle\langle R, L\rangle\rangle$. Applying the map, we get an element $[a_1, b_1, \dots, a_p, b_p, a_1, b_1, \dots, a_p, b_p, \dots]$ of $[0,\infty]$ which is fixed by $F$, so must be $\mu_+$ since this is the only element of $[0,\infty]$ fixed by $F$. This implies that the sequence $a_1, \dots, b_p$ which codifies $F$ in the $RL$-notation, also codifies $\mu_+$ in the continued fraction notation. Hence $\mu_+$ contains sufficient information to recover $F^\infty$, and hence, up to a power, to recover $F$. We will not make use of this fact, but it explains why $F$ will occasionally be less important than $\mu_+$ in the calculations. □

For $F = RL^3$, we have $\mu_+ = \frac{3+\sqrt{21}}{6} = [1, 3, 1, 3, \dots]$.

## 3. $S^1$ AND $S^2$.

In this section we review some topology. We emphasize that the main ideas used here come from hyperbolic geometry, but we restrict the exposition to a topological viewpoint.

We retain Notation 2.3.

**3.1 Definitions.** The semidirect product $\mathbb{Z}^2 \rtimes \mathrm{GL}_2(\mathbb{Z}) = \mathbb{Z}^2 \rtimes \mathrm{Aut}(\mathbb{Z}^2)$ is called the *integral affine plane group*, denoted $\mathrm{Af}_2(\mathbb{Z})$. This group acts in a natural way on the set $\mathbb{Z}^2$, and this action extends to an action on $\mathbb{R}^2$ as the group of affine transformations of $\mathbb{R}^2$ which carry $\mathbb{Z}^2$ to itself.

There is a natural map $\mathrm{F}_2 \rtimes \langle R, M\rangle \to \mathrm{Af}_2(\mathbb{Z})$, and it vanishes on the (infinite cyclic) kernel of the surjective map $\mathrm{F}_2 \rtimes \langle R, M\rangle \to \mathrm{Aut}(\mathrm{F}_2)$. Thus there is an induced map $\mathrm{Aut}(\mathrm{F}_2) \to \mathrm{Af}_2(\mathbb{Z})$. Moreover, it is surjective, with kernel $\mathrm{F}_2'$.

It follows that there is a specified action of $\mathrm{Aut}(\mathrm{F}_2) = \mathrm{Aut}\langle A, B, C\rangle$ on $\mathbb{R}^2$ with $X(x,y) = (x+1, y)$, $Y(x,y) = (x, y+1)$. We find

$$\begin{aligned}&A(x,y) = (-1-x, -y), \quad B(x,y) = (-1-x, -1-y), \quad C(x,y) = (-x, -1-y),\\ &R(x,y) = (x+y, y), \quad M(x,y) = (y, x), \quad L(x,y) = (x, x+y). \quad \square\end{aligned}$$

Notice that, in $\mathbb{R}^2$, $\mathrm{Aut}\langle A, B, C\rangle$ sends parallel straight lines to parallel straight lines, and the induced action on the inverses of their slopes gives the previously defined action of $\mathrm{Aut}\langle A, B, C\rangle$ on $\overline{\mathbb{R}}$ by Möbius transformations.

**3.2 Digression.** Recall that *the lower central series* $(\gamma_n \mathrm{F}_2 \mid n \geq 1)$ is defined recursively by $\gamma_1\mathrm{F}_2 := \mathrm{F}_2$, and $\gamma_{n+1}\mathrm{F}_2 := [\gamma_n\mathrm{F}_2, \mathrm{F}_2]$, that is, the subgroup of $\mathrm{F}_2$ generated by $\{gf\bar{g}\bar{f} \mid g \in \gamma_n\mathrm{F}_2, f \in \mathrm{F}_2\}$. Thus, for $n = 1, 2, 3$, $\mathrm{F}_2/\gamma_n\mathrm{F}_2$ is, respectively, 1, $\mathbb{Z}^2$, and the integral Heisenberg group $H$, which is obtained by making the commutator central.

For any characteristic subgroup $N$ of $\mathrm{F}_2$, there is a natural homomorphism $\mathrm{Aut}(\mathrm{F}_2) \to \mathrm{Aut}(\mathrm{F}_2/N)$, and the kernel, $\hat{N}$, is called the "congruence subgroup" of $\mathrm{Aut}(\mathrm{F}_2)$ arising from $N$.

Now let $G$ be a characteristic subgroup of $\mathrm{F}_2$, and let $N = [G, \mathrm{F}_2]$. Then $\hat{N}$ contains $G$, and we have a natural map $\alpha_G\colon \mathrm{Aut}(\mathrm{F}_2)/G \to \mathrm{Aut}(\mathrm{F}_2/[G, \mathrm{F}_2])$.

For example, for $G = \mathrm{F}_2$, we saw in Definitions 2.1 that Nielsen showed that $\alpha_{\mathrm{F}_2}$ is an isomorphism.

For $G = \mathrm{F}_2'$, Andreadakis found that $\alpha_{\mathrm{F}_2'}$ is an isomorphism; see [22, Proof of Theorem 2]. One sees this by checking that the kernel of the natural map

$$\mathrm{Aut}(H) \to \mathrm{Aut}(H/H') = \mathrm{GL}_2(\mathbb{Z})$$

consists of the inner automorphisms, and applying the five-lemma.

The above results show that there are natural identifications

$$\begin{aligned} \mathrm{Aut}(\mathrm{F}_2)/\gamma_1\mathrm{F}_2 &= \mathrm{Aut}(\mathrm{F}_2/\gamma_2\mathrm{F}_2) = \mathrm{Aut}(\mathbb{Z}^2) = \mathrm{GL}_2(\mathbb{Z}), \\ \mathrm{Af}_2(\mathbb{Z}) = \mathrm{Aut}(\mathrm{F}_2)/\gamma_2\mathrm{F}_2 &= \mathrm{Aut}(\mathrm{F}_2/\gamma_3\mathrm{F}_2) = \mathrm{Aut}(H). \end{aligned}$$

In particular, both $\mathrm{F}_2 = \gamma_1\mathrm{F}_2$ and $\mathrm{F}_2' = \gamma_2\mathrm{F}_2$ are congruence subgroups of $\mathrm{Aut}(\mathrm{F}_2)$, arising from the subgroups $\gamma_2\mathrm{F}_2$ and $\gamma_3\mathrm{F}_2$, respectively. □

We now consider the action of $\langle A, B, C\rangle$ on $\mathbb{R}^2 - \mathbb{Z}^2$.

**3.3 Definitions.** In Definitions 3.1, the action of $\mathbb{Z}^2$ on $\mathbb{R}^2$ is that of a subgroup, and the quotient space $\mathbb{R}^2/\mathbb{Z}^2 = \mathbb{R}^2/\mathrm{F}_2$ is a torus, while $(\mathbb{R}^2 - \mathbb{Z}^2)/\mathrm{F}_2$ is a once-punctured torus.

The quotient space $\mathbb{R}^2/\langle A, B, C\rangle$ is a sphere with four double points, and $A$, $B$ and $C$ all act in the same way on the double branched cover $\mathbb{R}^2/\mathrm{F}_2$, as "180° rotation of a donut on a skewer". Also, $(\mathbb{R}^2 - \mathbb{Z}^2)/\langle A, B, C\rangle$ is a once-punctured sphere with three double points.

Consider the action of $\langle A, B, C\rangle$ on $\mathbb{R}^2$. Its kernel is $\mathrm{F}_2'$. We take the triangle with vertices $(-1, 0)$, $(0, 0)$ and $(0, -1)$ as a fundamental domain. The edge joining $(-1, 0)$ and $(0, 0)$ is rotated 180° by $A$, and we label this edge with the letter $A$. We label the other two edges $B$ and $C$ correspondingly. Notice that $ABC$ acts as multiplication by $-1$, so is parabolic with unique fixed point $(0, 0)$, while $BAC$ fixes $(0, -1)$, and $ACB$ fixes $(-1, 0)$. We now tessellate $\mathbb{R}^2$ with the images under $\langle A, B, C\rangle$ of this labelled triangle. Thus the horizontal lines are labelled with $A$s, the diagonal lines with $B$s, and the vertical lines with $C$s. See Fig. 1. Our real interest is in the action of $\langle A, B, C\rangle$ on $\mathbb{R}^2 - \mathbb{Z}^2$, so we delete $\mathbb{Z}^2$, the set of all the vertices, throughout the foregoing description.

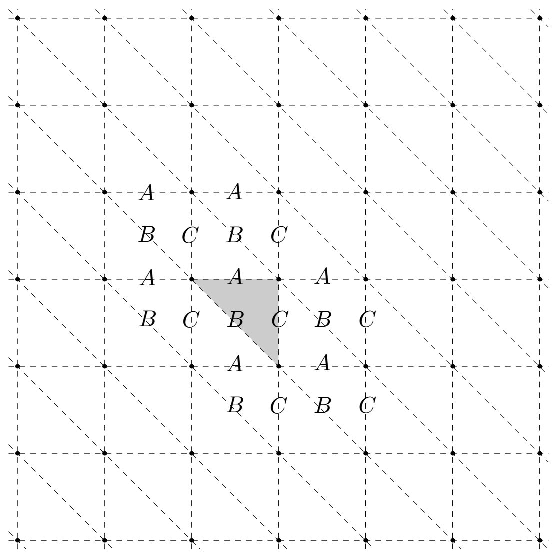


FIG. 1. $\mathbb{R}^2 - \mathbb{Z}^2$ tessellated with ideal triangles.

Now $\mathbb{R}^2 - \mathbb{Z}^2$ is a branched cover of the once-punctured sphere with three double points $(\mathbb{R}^2 - \mathbb{Z}^2)/\langle A, B, C\rangle$. The universal cover of $\mathbb{R}^2 - \mathbb{Z}^2$ is a topological open disk, $\mathbb{D}$, tessellated with ideal triangles; see Fig. 2. One can also think of $\mathbb{D}$ as an infinite-sheeted cover of $\mathbb{R}^2 - \mathbb{Z}^2$, with infinite spirals around the missing points of $\mathbb{Z}^2$; see Fig. 1.

The dual of the tessellation of $\mathbb{D}$ is an $\langle A, B, C\rangle$-tree; see Fig. 3. The ends $(ABC)^\infty$ and $(CBA)^\infty$ determine a doubly infinite path through triangles which can be thought of as being incident to a single cusp, or ideal vertex. We can compactify the path with this cusp. More generally, we can attach the circle $S^1 = \mathcal{E}/{\sim}$ described in Section 1, to compactify the $\langle A, B, C\rangle$-tree and $\mathbb{D}$ so that cusps$(S^1)$ becomes the set of cusps, or vertices at infinity, of the ideal triangles. See, for example, [1, Section 4] and Fig. 3. (For $\mathbb{R}^2 - \mathbb{Z}^2$, $(0,0)$ corresponds to the $\mathrm{F}_2'$-orbit of $\infty_{S^1}$, $(-1,0)$ corresponds to $0_{S^1}$, and $(0,-1)$ corresponds to $1_{S^1}$.) The compactification $\overline{\mathbb{D}} := \mathbb{D} \cup S^1$ is an $\langle A, B, C\rangle$-space which is a topological closed disk; see Fig. 2.

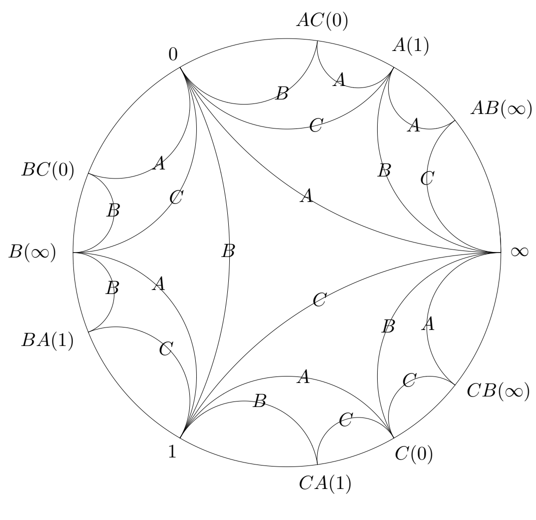


FIG. 2. Disc tessellated with ideal triangles.

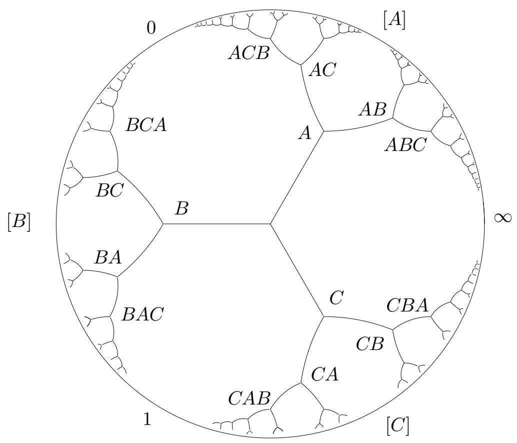


FIG. 3. $S^1 := \mathcal{E}/\!\sim$, oriented counter-clockwise.

Let $\mathbb{S}^2$ be the $\langle A, B, C\rangle$-space obtained by taking two copies of $\overline{\mathbb{D}}$ and identifying the two copies of $S^1$. We then orient $\mathbb{S}^2$. Thus $\mathbb{S}^2$ is an oriented topological two-sphere consisting of two copies of $\mathbb{D}$ and one copy of $S^1$, which we call the *northern* and *southern* hemispheres, and the *equator*. □

**3.4 Definitions.** By Notation 2.3, $F$ acts on $\mathbb{R}^2$ (and $\mathbb{R}^2 - \mathbb{Z}^2$) by

$$F(x, y) = (ax + by, cx + dy).$$

Since $F$ is hyperbolic, it has an attracting and a repelling real eigenspace, with irrational slopes $\frac{1}{\mu_+}$ and $\frac{1}{\mu_-}$, respectively.

It is not difficult to show that reading off the labels of the edges in the third quadrant crossed by the attracting eigenspace gives $E^+$. Reading the labels in the first quadrant (starting from the interior of our chosen fundamental domain) gives $ABCE^+$ (resp. $CBAE^+$) if we read the labels of those edges which touch the attracting eigenspace from above (resp. below). The part of the repelling eigenspace for $F$ in the fourth (resp. second) quadrant, read from below, gives $E^-$ (resp. $CBAE^-$). See Fig. 4.

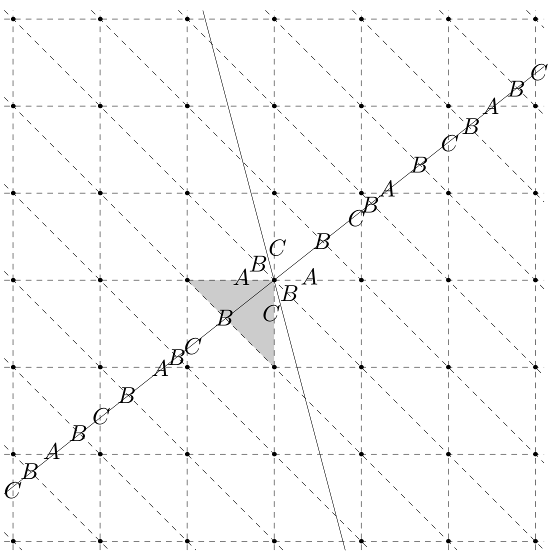

FIG. 4. Eigenspaces.

In Example 2.4, we saw that for $F = RL^3$ we have $\mu_{+} = \frac{3+\sqrt{21}}{6}$ and $E^{+} = BCBABCBABC\cdots$, which can be read in the third quadrant in Fig. 4. Similarly, $\mu_{-} = \frac{3-\sqrt{21}}{6}$ and $E^{-} = CABABA\cdots$ can be read from below in the fourth quadrant in Fig. 4.

The lines of slope $\frac{1}{\mu_{+}}$ in $\mathbb{R}^2$ are of two types. Each line of slope $\frac{1}{\mu_{+}}$ in $\mathbb{R}^2$ which contains an element of $\mathbb{Z}^2$ breaks into two lines on deleting $\mathbb{Z}^2$, and so gives rise to two lines, informally called "half lines", of slope $\frac{1}{\mu_{+}}$ in $\mathbb{R}^2 - \mathbb{Z}^2$. Each line of slope $\frac{1}{\mu_{+}}$ in $\mathbb{R}^2$ which does not meet $\mathbb{Z}^2$ is unaffected by deleting $\mathbb{Z}^2$, and so gives rise to one line, informally called a "whole line", of slope $\frac{1}{\mu_{+}}$ in $\mathbb{R}^2 - \mathbb{Z}^2$. These (whole and half) lines partition, or foliate, $\mathbb{R}^2 - \mathbb{Z}^2$. Each line lifts to an $\mathrm{F}_2'$-orbit of disjoint topological lines in $\mathbb{D}$. The union of the lifted lines is all of $\mathbb{D}$; we denote this partition, or foliation, of $\mathbb{D}$ by $\mathcal{F}_{+}$. The closure of each lifted line in $\overline{\mathbb{D}}$ consists of the lifted line in $\mathbb{D}$ together with two points in $S^1$, the *ends* of the lifted line. When the line that is to be lifted is a whole line, neither of the ends of the lifted line are cusps and we call the lifted line an "irrational" foliation line. When the line that is to be lifted is a half line, one of the ends of the lifted line is a cusp, and the other is not, and we call the lifted line a "rational" foliation line.

Similarly, the lines of slope $\frac{1}{\mu_{-}}$ in $\mathbb{R}^2$ determine a foliation $\mathcal{F}_{-}$ of $\mathbb{D}$.

We endow the northern hemisphere of $\mathbb{S}^2$ with the foliation $\mathcal{F}_{+}$, and the southern hemisphere of $\mathbb{S}^2$ with the foliation $\mathcal{F}_{-}$. As we will see shortly, each non-cusp on the equator, $S^1$, is the end of at most one foliation line.

Now consider a cusp $e$ in $S^1$. There are a countable set of (rational) foliation lines incident to $e$, and, in each hemisphere, they are cyclically permuted by the (infinite cyclic) $\langle A, B, C\rangle$-stabilizer of $e$. Informally, this collection of foliation lines and their ends is called a "spider".

Let $\approx$ be the smallest equivalence relation on $S^1$ which relates two elements of $S^1$ if they are joined by a foliation line in $\mathbb{S}^2$. Each $\approx$-equivalence class in $S^1$ consists of one of the following: a single element not incident to any foliation lines; two elements joined by an irrational foliation line; the countable set of extremities of a spider. This equivalence relation on $S^1 = \mathcal{E}/{\sim}$ induces an equivalence relation on $\mathcal{E}$ again denoted $\approx$. There is a natural homeomorphism of quotient spaces $\mathcal{E}/{\approx} \;\simeq\; S^1/{\approx}$.

In $\mathbb{S}^2$, collapsing each foliation line and its two ends to a point gives a quotient space of $\mathbb{S}^2$, and, by a theorem of R. L. Moore, this quotient is a topological two-sphere; see Theorem A.13 in the Appendix. Since every element of $\mathbb{S}^2$ off the equator lies in a foliation line, we see that the quotient two-sphere can be naturally identified with $S^1/{\approx}$.

Thus $\mathcal{E}/{\approx}$ is a two-sphere. We denote it by $S^2$. □

**3.5 Notation.** Consider the triangularly tessellated plane described in Definitions 3.3; see Fig. 1. For any $x \in [-1, \mu_{+}]$, consider the line through $(x, 0)$ of slope $\frac{1}{\mu_{+}}$. Thus we are considering a line that is parallel to the attracting eigenspace and passes through the fundamental domain; see Fig. 4. The portion of the line that lies in the third quadrant, read from below, gives $E(x, \mu_{+})$, where we define

$$E(x,y) = (BC)^{[-x+y]}BA\prod_{n=1}^{\infty}((BC)^{[-x+y(n+1)]-[-x+yn]}BA).$$

Notice that $E(0,\mu_+) = E^+$. The remaining (upper) portion, read from above, gives

$$(CB)^{[x+1]}AB\prod_{n=1}^{\infty}((CB)^{[x+1+\mu_+ n]-[x+1+\mu_+(n-1)]}AB). \tag{3.5.1}$$

It can be checked that this is $B\cdot E(\mu_+ - 1 - x, \mu_+)$. For $x = -1$ and $x = \mu_+$, this is $AE^+$ and $CE^+$, respectively. Hence, for all $G \in \langle A, B, C\rangle$, and all $x \in [-1, \mu_+]$,

$$GB\cdot E(\mu_+ - 1 - x, \mu_+) \approx G\cdot E(x, \mu_+). \tag{3.5.2}$$

Also, for all $G \in \langle A, B, C\rangle$, $G(ABC)^\infty \approx G(CBA)^\infty \approx GE^+$.

Notice that each end involved in one of these expressions has a tail of the form $(ABC)^\infty$, or $(CBA)^\infty$, or in which every other letter is $B$ and there are infinitely many $A$'s and $C$'s.

We have described the relations given by $\mathcal{F}_+$, and there are similar relations arising from $\mathcal{F}_-$. Here each end involved has a tail of the form $(ABC)^\infty$, or $(CBA)^\infty$, or in which every other letter is $A$ and there are infinitely many $B$'s and $C$'s.

Together, these relations determine $\approx$.

We now see that, in $S^1$, the set of elements $\approx$-related to other elements consists of the cusps together with a complementary set; moreover, the complementary sets contributed by $\mathcal{F}_+$ and $\mathcal{F}_-$ are disjoint. We will consider these sets in more detail in Section 9.

We note also that $\infty_{S^2} = \{(ABC)^\infty, (CBA)^\infty\} \cup \langle ABC\rangle E^+ \cup \langle ABC\rangle E^-$. □

## 4. Fixed ends.

We next want to verify the claim that $F$ fixes $E^+$ and $E^-$. These ends come from labellings on fixed half-eigenspaces in the third and fourth quadrants, respectively. It is not difficult to show that $F$ respects the labellings, but there are some details to check, and we find it more convenient, or at least more algebraic, to proceed indirectly, by considering a map assigning ends to irrational real numbers. This contrasts with Digression 2.6 where we described a classical map assigning real numbers to irrational ends.

**4.1 Theorem.** *With Notation 2.3, $F$ fixes $E^+$ and $E^-$.*

*Proof.* Define a map

$$\Phi\colon \mathbb{R} - \mathbb{Q} \to \mathcal{E}, \quad x \mapsto \prod_{n=1}^{\infty}((BC)^{[x]_n}BA).$$

In terms of Notation 3.5, $\Phi(x) = E(0, x)$. We will examine how $\Phi$ behaves with respect to the actions of $\langle R, M\rangle$, which acts on $\mathbb{R} - \mathbb{Q}$ by Möbius transformations, as well as acting on $\mathcal{E}$.

Let $x \in \mathbb{R} - \mathbb{Q}$.

We see

$$R(\Phi(x)) = R(\prod_{n=1}^{\infty}((BC)^{[x]_n}BA)) = \prod_{n=1}^{\infty}((BC)^{[x]_n}BCBA)$$
$$= \prod_{n=1}^{\infty}((BC)^{[x+1]_n}BA) = \Phi(x+1) = \Phi(R(x)).$$

Thus $\Phi$ and $R$ commute (even on rational $x$).

We next examine what happens for $M$. Consider first the case where $x > 1$. Here

$$M(\Phi(x)) = M(\prod_{m=1}^{\infty}((BC)^{[x]_m}BA))$$
$$= \prod_{m=1}^{\infty}((BA)^{[x]_m}BC) \tag{4.1.1}$$
$$= \prod_{n=1}^{\infty}((BC)^{[\frac{1}{x}]_n}BA) \tag{4.1.2}$$
$$= \Phi(\frac{1}{x}) = \Phi(M(x)),$$

where the claim that (4.1.1) = (4.1.2) is justified as follows. Let $m \geq 1$. Consider the $m$th occurrence of $BC$ in (4.1.1), and let $n$ denote the number of $BA$'s which precede it. Since $x > 1$, the $n+1$st $BA$ in (4.1.1) is immediately preceded by the $m$th $BC$. Also, $[mx] = n$, so $n < mx < n+1$. Hence $\frac{n}{x} < m < \frac{n+1}{x}$, and, since $\frac{1}{x} < 1$, we see $[\frac{1}{x}]_{n+1} = 1$ and $[\frac{n}{x}] = m-1$. Hence the $n+1$st occurrence of $BA$ in (4.1.2) has $m$ $BC$'s preceding it, and one immediately preceding it, which is the $m$th occurrence of $BC$ in (4.1.2). Thus the $m$th $BC$ is immediately followed by the $n+1$st $BA$ in both (4.1.1) and (4.1.2). Since this happens for all $m$, we see (4.1.1) = (4.1.2). Hence $M(\Phi(x)) = \Phi(M(x))$. Since $M^2$ is the identity, we see $\Phi(M(\frac{1}{x})) = \Phi(x) = M(M(\Phi(x))) = M(\Phi(M(x))) = M(\Phi(\frac{1}{x}))$.

Now suppose that $x > 0$. The previous paragraph shows that $M(\Phi(x)) = \Phi(M(x))$. Also $[-x] = -[x] - 1$, so we see $[-x]_1 = -[x]_1 - 1$, and $[-x]_n = -[x]_n$ for $n \geq 2$. Thus

$$M\Phi(-x) = \prod_{n=1}^{\infty}((BA)^{[-x]_n}BC) = AB\prod_{n=1}^{\infty}((AB)^{[x]_n}BC)$$
$$= AB\prod_{n=1}^{\infty}((BC)^{[\frac{1}{x}]_n}AB) = ABCB\prod_{n=1}^{\infty}((CB)^{[-\frac{1}{x}]_n}AB)$$
$$= ABC\prod_{n=1}^{\infty}((BC)^{[\frac{1}{-x}]_n}BA) = ABC\Phi(M(-x)).$$

Thus, if $x > 0$, then $M(\Phi(x)) = \Phi(M(x))$, while if $x < 0$, then $M(\Phi(x)) = ABC\Phi(M(x))$. So the actions commute modulo a power of $ABC$.

Now $\langle\langle R, L\rangle\rangle$ acts on the set of positive irrational numbers by Möbius transformations and, since $L = MRM$, it follows that, if $T \in \langle\langle R, L\rangle\rangle$, then $T(\Phi(x)) = \Phi(T(x))$. Hence, if $T$ fixes $x$, then $T$ fixes $\Phi(x)$ in $\mathcal{E}$. It follows that $F$ fixes

$$\Phi(\mu_{+}) = \prod_{n=1}^{\infty}((BC)^{[\mu_{+}]_n}BA) = E^{+}.$$

Similarly, on the negative irrational real numbers, $\Phi$ commutes with the action of $\langle\langle \bar{R}, \bar{L}\rangle\rangle$, so $\bar{F}$, and hence $F$, fixes

$$\begin{aligned}\Phi(\mu_{-}) &= \prod_{n=1}^{\infty}((BC)^{[\mu_{-}]_n}BA) = CB\prod_{n=1}^{\infty}((CB)^{[-\mu_{-}]_n}BA)\\ &= CB\prod_{n=1}^{\infty}((BA)^{[\frac{1}{-\mu_{-}}]_n}CB) = C\prod_{n=1}^{\infty}((AB)^{[\frac{1}{-\mu_{-}}]_n}BC)\\ &= C\prod_{n=1}^{\infty}((AB)^{[\frac{1}{-\mu_{-}}-1]_n}AC) = C\prod_{n=1}^{\infty}((AB)^{[\frac{1+\mu_{-}}{-\mu_{-}}]_n}AC) = E^{-}. \quad \square\end{aligned}$$

We now describe $\mathcal{E}^F$.

**4.2 Theorem.** *With Notation 2.3,*

$$\mathcal{E}^F = \{(ABC)^{\infty}, (CBA)^{\infty}\} \cup \langle ABC\rangle E^{+} \cup \langle ABC\rangle E^{-} = \infty_{S^2}.$$

*Hence $\approx$ is the smallest closed equivalence relation on $\mathcal{E}$ such that $GE_1 \approx GE_2$ for all $G \in \langle A, B, C\rangle$ and all $E_1, E_2 \in \mathcal{E}^F$.*

*Proof.* We have seen that each of the four orbits lies in the set of fixed ends. See [13] or [15] for proofs that there can be no other fixed points. The second part is implied by the fact that $\mathcal{E}/\approx$ is a topological two-sphere, and hence Hausdorff; see Lemma A.12 in the Appendix. $\square$

## 5. arc($F$).

In this section we prove the existence of an arc in $S^2$ which is then seen to be $\partial^{+}[B]_{S^2}$.

**5.1 Notation.** Throughout this section we use a modified version of Notation 2.3, in which we abbreviate

$$\mu = \mu_{+} = \frac{\lambda - d}{c}, \quad E = E^{+} = \prod_{n=1}^{\infty}((BC)^{[\mu]_n}BA).$$

For $n \in \mathbb{N}$, we let $E_n = E^{+}_n$, the initial segment of $E$ of length $n$ (as a word in $A$, $B$ and $C$), and $\bar{E}_n$ denotes its inverse. Similar notation applies for any element of $\mathcal{E}$.

Set

$$\mathcal{M} = \mathcal{M}^{+}_B = \langle\langle B, \bar{E}_{2n}F \mid 0 \leq n \leq a+b+c+d-2\rangle\rangle.$$

where we understand this to mean that $\mathcal{M}^{+}_B$ is the semigroup generated in the group $\langle A, B, C, F\rangle$ by the indicated set.

From Definitions 3.1, we have an action of $\mathrm{Aut}\langle A, B, C\rangle$ on $\mathbb{R}^2$, and $\langle A, B, C, F\rangle$ acts on the lines of slope $\frac{1}{\mu}$, that is, parallel to the attracting eigenspace. By collapsing these lines to the $x$-axis, we can give $\mathbb{R}$ the structure of an $\langle A, B, C, F\rangle$-space with

$$A(x) := -x-1, \quad B(x) := -x+\mu-1, \quad C(x) := -x+\mu, \quad F(x) := \lambda^{-1}x.$$

Here $X = CB$ acts by $X(x) = x+1$, and $Y = AB$ acts by $Y(x) = x-\mu$.

We set $\mathrm{cusps}(\mathbb{R}) = \mathbb{Z} + \mathbb{Z}\mu$, the image of $\mathbb{Z}^2$ under the collapsing map, and an $\langle A, B, C, F\rangle$-subspace of $\mathbb{R}$. $\square$

**5.2 Notation.** Let $\phi\colon [-1, \mu] \to \mathcal{E}$ be given by $x \mapsto B \cdot E(\mu-1-x, \mu)$.

An explicit expression for $\phi(x)$ is given in (3.5.1), and we can see from this expression that $\phi$ is one-sided continuous, from the right.

It is straightforward to show that $B\phi B\colon [-1, \mu] \to \mathcal{E}$ is given by $x \mapsto E(x, \mu)$, so $B\phi B$ is one-sided continuous, from the left.

Let $\psi$ denote the induced map from $[-1, \mu]$ to $S^2$, which assigns to each $x$ the $\approx$-equivalence class of $\phi(x)$.

By (3.5.2), $\phi(x) \approx B\phi B(x)$, so $\psi = B\psi B$, and this is continuous from the right and from the left, so is continuous.

It is not difficult to show that two different elements of $\mathcal{E}$ cannot be $\approx$-related if they both start with $B$ and have $B$ in every other position. It then follows easily that $\psi$ is injective.

The image of $\psi$ is thus an arc in $S^2_F$, and we denote it by $\mathrm{arc}(F)$.

By considering $\bar{F}$ rather than $F$, we get an arc in $S^2_{\bar{F}}$ denoted $\mathrm{arc}(\bar{F})$. But there is a natural identification $S^2_{\bar{F}} = S^2_F$, and, in $S^2_F$, $\mathrm{arc}(\bar{F})$ arises from the set of lines of slope $\frac{1}{\mu_-}$ which pass through the fundamental domain; see Fig. 4. $\square$

Next we analyse the set of cusps which lie in $\mathrm{arc}(F)$.

**5.3 Lemma.** *With Notation 5.1, if $n \in \mathbb{N}$, and $i = [\frac{(n+1)\mu}{\mu+1}]$ and $j = [\frac{n+1}{\mu+1}]$, then the following hold.*

(i). *In $\langle A, B, C\rangle$, $\bar{E}_{2n}$ is a product of $i$ $X$'s and $j$ $Y$'s in some order, and $\bar{E}_{2n+1} = B\bar{E}_{2n}$.*

(ii). *In $\mathbb{R}$, $\bar{E}_{2n}(0)$ is $n-(\mu+1)[\frac{n+1}{\mu+1}]$, the unique element of $n+(\mu+1)\mathbb{Z}$ which lies in the interval $(-1, \mu)$.*

(iii). *In $\mathbb{R}$, $\bar{E}_{2n+1}(0)$ is $-n-1+\mu+(\mu+1)[\frac{n+1}{\mu+1}]$, the unique element of $-n-2+(\mu+1)\mathbb{Z}$ which lies in the interval $(-1, \mu)$.*

*Proof.* Let $m$ denote the number of occurrences of $BA$ in $E_{2n}$. From the definition of $E$ we find that $[m\mu]+m \le n \le [(m+1)\mu]+m$, from which it follows that $m = [\frac{n+1}{\mu+1}] = j$, and $n-m = i$. Hence (i) holds.

Since $\bar{E}_{2n}(0) = i - j\mu$, we see that (ii) holds.

Applying $B$ to (ii) we get (iii). $\square$

**5.4 Remark.** If we identify $\mathbb{R}/((\mu+1)\mathbb{Z})$ with $[-1, \mu)$, then the foregoing shows that, for all $n \in \mathbb{N}$, $n+(\mu+1)\mathbb{Z}$ becomes $\bar{E}_{2n}(0)$, $-(n+2)+(\mu+1)\mathbb{Z}$ becomes $\bar{E}_{2n+1}(0)$, and $-1+(\mu+1)\mathbb{Z}$ becomes the endpoint $-1$. $\square$

**5.5 Theorem.** *With Notation 5.1 and 5.2, the following hold.*

(i). $\mathrm{cusps}([-1,\mu]) = \{-1, \mu, \bar{E}_n(0) \mid n \in \mathbb{N}\}$.
(ii). $\mathrm{cusps}([-1,\mu])$ *is dense in* $[-1,\mu]$.
(iii). *The sequence* $(\bar{E}_{2n}(0) \mid n \in \mathbb{N})$ *starts at* $E_0(0) = 0$*, and each successive term is given by adding 1 or subtracting* $\mu$ *so as to stay in the interval* $(-1,\mu)$*. The sequence* $(\bar{E}_{2n+1}(0) \mid n \in \mathbb{N})$ *starts at* $E_1(0) = B(0) = \mu - 1$*, and each successive term is given by subtracting 1 or adding* $\mu$ *so as to stay in the interval* $(-1,\mu)$*.*
(iv). $\mathrm{cusps}(\mathrm{arc}(F)) = \{0_{S^2}, 1_S^2, \bar{E}_n(\infty_{S^2}) \mid n \in \mathbb{N}\}$.
(v). $\mathrm{cusps}(\mathrm{arc}(F))$ *is dense in* $\mathrm{arc}(F)$.
(vi). *Under the homeomorphism* $\psi\colon [-1,\mu] \to \mathrm{arc}(F)$, $\psi(-1) = 0_{S^2}$, $\psi(\mu) = 1_{S^2}$*, and, for all* $n \in \mathbb{N}$, $\psi(\bar{E}_n(0)) = \bar{E}_n(\infty_{S^2})$*.*

*Proof.* Notice that $\mathrm{cusps}([-1,\mu]) = (\mathbb{Z}+\mu\mathbb{Z}) \cap [-1,\mu]$ contains exactly one element from each coset $n + (1+\mu)\mathbb{Z}$, except for $n = -1$ which contributes two elements. Lemma 5.3 now shows that (i) holds. Here (ii) and (iii) are clear.

The sets $\mathrm{cusps}(\mathrm{arc}(F))$ and $\mathrm{cusps}([-1,\mu])$ both arise from the lines of slope $\frac{1}{\mu}$ in $\mathbb{R}^2$ which meet both $\mathbb{Z}^2$ and the fundamental domain for $\langle A, B, C\rangle$. If a line of slope $\frac{1}{\mu}$ contains $(i,j) \in \mathbb{Z}^2$, then it contains $(i - j\mu, 0)$; moreover, the line meets the fundamental domain for $\langle A, B, C\rangle$ if and only if $i - j\mu \in [-1,\mu]$. We now see that $\mathrm{cusps}(\mathrm{arc}(F)) = \psi(\mathrm{cusps}([-1,\mu]))$, so (v) holds. The endpoints of $\mathrm{arc}(F)$ are given by the two ends $\phi(-1) = AE$ and $\phi(\mu) = CE$, so are $A(\infty_{S^2}) = 0_{S^2}$ and $C(\infty_{S^2}) = 1_{S^2}$, respectively. Moreover, it is straightforward to calculate that, for all $n \in \mathbb{N}$, in $\mathcal{E}$, $\phi(\bar{E}_{2n}(0)) = \bar{E}_{2n}CBAE$, so, in $S^2$, $\psi(\bar{E}_{2n}(0)) = \bar{E}_{2n}(\infty_{S^2})$. Similarly, $\phi(\bar{E}_{2n+1}(0)) = \bar{E}_{2n+1}E$ and $\psi(\bar{E}_{2n+1}(0)) = \bar{E}_{2n+1}(\infty_{S^2})$. Thus (iv) and (vi) hold. □

**5.6 Theorem.** *With Notation 5.1 and 5.2, the following hold.*

(i). $\mathrm{arc}(F) \cup \mathrm{arc}(\bar{F}) = ([A]_{S^2} \cap [B]_{S^2}) \cup ([B]_{S^2} \cap [C]_{S^2}) \cup ([C]_{S^2} \cap [A]_{S^2})$.
(ii). $\mathrm{arc}(F) \cap \mathrm{arc}(\bar{F}) = \{0_{S^2}, 1_{S^2}, \infty_{S^2}\}$.
(iii). $\mathrm{arc}(F) \cup \mathrm{arc}(\bar{F})$ *forms an embedded graph in* $S^2$ *with three edges and two branch points* $1_{S^2}$ *and* $\infty_{S^2}$*. The complement of the graph consists of three regions in* $S^2$ *which are the interiors of* $[A]_{S^2}$, $[B]_{S^2}$ *and* $[C]_{S^2}$*.*
(iv). $\mathrm{arc}(F)$ *is* $\partial^{+}[B]_{S^2}$*, and, with reversed orientation, is the concatenation of* $\partial^{+}[C]_{S^2}$ *and* $\partial^{+}[A]_{S^2}$*.*
(v). $\mathrm{arc}(\bar{F})$ *is* $\partial^{-}[A]_{S^2}$*, and, with reversed orientation, is the concatenation of* $\partial^{-}[B]_{S^2}$ *and* $\partial^{-}[C]_{S^2}$*.*

*Proof.* (i). Notice that $\mathrm{arc}(F)$ lies in $[B]_{S^2} \cap [B^*]_{S^2}$, since $B\phi B(x) \in [B]_{\mathcal{E}}$, for all $x \in [-1,\mu]$, and $\phi(x) \in [A]_{\mathcal{E}}$ for all $x \in [-1,0)$, and $\phi(x) \in [C]_{\mathcal{E}}$ for all $x \in [0,\mu]$.

Similarly, $\mathrm{arc}(\bar{F})$ lies in $[A]_{S^2} \cap [A^*]_{S^2}$.

The elements of $([A]_{S^2} \cap [B]_{S^2}) \cup ([B]_{S^2} \cap [C]_{S^2}) \cup ([C]_{S^2} \cap [A]_{S^2})$ are given by relations $E_1 \approx E_2$ where $E_1$ and $E_2$ are elements of $\mathcal{E}$ which begin with different letters.

Consider first the relations $G\phi(x) \approx GB\phi B(x)$, as in (3.5.2), where $x \in [-1,\mu]$, and $G \in \langle A, B, C\rangle$. Taking $G = B$ gives the same relation as taking $G = 1$ and replacing $x$ with $B(x)$. If $0 \le x \le \mu$, then taking $G = C$ gives the same relation as taking $G = 1$ and replacing $x$ with $C(x)$. Similarly, if $-1 \le x < 0$, then taking $G = A$ gives the same relation as taking $G = 1$ and replacing $x$ with $A(x)$. This means that each relation $G\phi(x) \approx GB\phi B(x)$, where $G$ is the inverse of an initial segment of $\phi(x)$ or of $B\phi B(x)$, is already of the form $\phi(x') \approx B\phi B(x')$ for some $x' \in [-1,\mu]$. Thus all points of $([A]_{S^2} \cap [B]_{S^2}) \cup ([B]_{S^2} \cap [C]_{S^2}) \cup ([C]_{S^2} \cap [A]_{S^2})$ arising in this way already lie in $\operatorname{arc}(F)$.

A similar result holds for $\operatorname{arc}(\bar{F})$.

It remains to consider the points $G(\infty_{S^2})$ where $G$ is the inverse of an initial segment of some element of $\infty_{S^2} = \mathcal{E}^F$. If $G = \bar{E}_n$ for some $n$, we have seen that $G(\infty_{S^2})$ lies in $\operatorname{arc}(F)$. Similarly, if $G = \bar{E}_n^-$ then $G(\infty_{S^2})$ lies in $\operatorname{arc}(\bar{F})$. If $G$ is the inverse of an initial segment of $(ABC)^\infty$ or $(CBA)^\infty$, then $G(\infty_{S^2}) \in \{0_{S^2}, 1_{S^2}, \infty_{S^2}\}$.

(ii). It follows from the foregoing, and the fact that the tails involve alternating $B$'s or alternating $A$'s, that $0_{S^2}$, $1_{S^2}$, and $\infty_{S^2}$ are the only points in common for $\operatorname{arc}(F)$ and $\operatorname{arc}(\bar{F})$, and $0_{S^2}$ is an endpoint of both.

(iii) is now clear.

(iv) and (v). Thus $\partial[B]_{S^2}$ is given by concatenating two of the edges joining $\infty_{S^2}$ to $1_{S^2}$ in the embedded graph, and it is clear we must take all of $\operatorname{arc}(F)$ and part of $\operatorname{arc}(\bar{F})$. In Section 1, we specified that $S^2$ was to be oriented so that, in travelling along $\partial[B]_{S^2}$ from $0_{S^2}$ to $1_{S^2}$ via $\infty_{S^2}$, we have $[B]_{S^2}$ on the right. Hence in travelling along $\operatorname{arc}(F)$ we have $[B]_{S^2}$ on the right, so $\operatorname{arc}(F)$ is $\partial^+[B]_{S^2}$. The remaining claims are straightforward to check. □

We now examine the action of $\mathcal{M}$, and we begin by studying the action of $F$.

**5.7 Lemma.** *With Notation 5.1, for any $n \in \mathbb{N}$, the following hold.*

(i). *There exists $m_n \in \mathbb{N}$ such that $F(E_n) = E_{m_n}$.*
(ii). $m_{2n} < m_{2n+2} < m_{2n+1} < m_{2n+4}$.
(iii). *$F(E_{2n+1}) = F(E_{2n})F(B)$ with no cancellation.*
(iv). $F(A) = E_{2(a+c)-3}$, $F(A)A = E_{2(a+c)-2}$.
(v). $F(B)B = E_{2(a+b+c+d)-4}$, $F(B) = E_{2(a+b+c+d)-3}$,
(vi). $F(C) = E_{2(b+d)-3}$, $F(C)C = E_{2(b+d)-2}$.
(vii). *$AFA$, $FA$, $BFB$, $FB$, $CFC$ and $FC$ all lie in $\mathcal{M}$.*

*Proof.* Recall that $R = (A, BCB, B)$ and $L = (B, BAB, C)$ are positive with respect to $BC = \bar{X}$ and $BA = \bar{Y}$. In fact, $R(BC) = BC$, $R(BA) = BCBA$, $L(BC) = BABC$, $L(BA) = BA$. Thus $F$ is positive with respect to $BC$ and $BA$, so $F(BC)$ is a product of $a$ $BC$'s and $c$ $BA$'s in some order, and $F(BA)$ is a product of $b$ $BC$'s and $d$ $BA$'s in some order.

Since $E$ is a right-infinite word in $BC$ and $BA$, $F(E_{2n})$ is an initial segment of $F(E) = E$, so (i) holds for even $n$. Also, $m_{2n} < m_{2n+2} < m_{2n+4}$.

Taking $n = 1$, we see that $F(BC) = E_{2(a+c)}$. Since $L(BA)$ is an initial segment of $L(BC)$, it follows that $F(BA)$ is an initial segment of $F(BC)$. Hence $F(BA) = E_{2(b+d)}$.

For all $i \in \mathbb{N}$, $R^iL(BC)$ terminates in $BABC$, and it follows that $F(BC)$ terminates in $BABC$. Hence $F(A) = F(BC)CBA$ terminates in $B$, and $F(A) = E_{2(a+c)-3}$, $F(A)A = E_{2(a+c)-2}$. Also, $F(A)B = E_{2(a+c)-4}$, so two of the given generators of $\mathcal{M}$ are $\bar{E}_{2(a+c)-2}F = AF(A)F = AFA$ and $\bar{E}_{2(a+c)-4}F = BF(A)F = BFA$. Since $B \in \mathcal{M}$, we see $FA \in \mathcal{M}$.

For all $i \in \mathbb{N}$, $RL^i(BA)$ terminates in $BCBA$, and it follows that $F(BA)$ terminates in $BCBA$. Hence $F(C) = F(BA)ABC$ terminates in $B$, and $F(C) = E_{2(b+d)-3}$, $F(C)C = E_{2(b+d)-2}$. As in the preceding paragraph, $CFC$ and $FC$ lie in $\mathcal{M}$.

Now $F(B) = F(BABC)CBA = F(BCBA)ABC$ is an initial segment of $F(BABC)$, of $F(BCBA)$, and (hence) of $F(BCBC)$. It follows that $F(B) = E_{2(a+b+c+d)-3}$, and that $F(B)$ terminates in $B$, and has $F(BC)$ and (hence) $F(BA)$ as initial segments. Thus $F(E_{2n+1}) = F(E_{2n})F(B)$ has $F(E_{2n+2})$ as an initial segment.

Since $E_{2n+4}$ is $E_{2n}$ followed by $BCBA$ or $BABC$ or $BCBC$, we see that $F(E_{2n+4})$ is $F(E_{2n})$ followed by $F(BCBA)$ or $F(BABC)$ or $F(BCBC)$, all three of which begin with $F(B)$. Thus $F(E_{2n+4})$ begins with $F(E_{2n})F(B) = F(E_{2n+1})$.

This completes the proof of all the statements. □

For instance, in Example 2.4, $a + c = 7$ and $F(A)$ has length 11.

**5.8 Lemma.** *With Notation 5.1, let* $\mathrm{K} = \langle F, ABC\rangle$*, and let* $\langle A, B, C, F\rangle$*, and hence* $\mathcal{M}$*, act on the set of cosets* $\langle A, B, C, F\rangle/\mathrm{K}$ *by multiplication. Then the* $\mathcal{M}$*-orbit of the coset* $\mathrm{K}$ *is* $\mathcal{M}\mathrm{K} = \{\bar{E}_n\mathrm{K} \mid n \in \mathbb{N}\}$*, and the* $\mathcal{M}$*-orbit of the cosets* $A\mathrm{K}$ *and* $C\mathrm{K}$ *is*

$$\mathcal{M}A\mathrm{K} = \mathcal{M}C\mathrm{K} = \{A\mathrm{K}, C\mathrm{K}, \bar{E}_n\mathrm{K} \mid n \in \mathbb{N}\}.$$

*Proof.* Let $Q = \{\bar{E}_n\mathrm{K} \mid n \in \mathbb{N}\}$.

We show first that $Q$ is closed under the action of $\mathcal{M}$.

Consider any $r, n \in \mathbb{N}$, with $0 \le r \le a + b + c + d - 2$.

Since $E_{2n+1} = E_{2n}B$, on taking inverses we see that $B$ interchanges the elements of $Q$ in pairs.

Notice that $E_{2r}$ is an initial segment of $F(B)$, and $F(B)$ its own inverse, so $F(B)$ terminates in $\bar{E}_{2r}$, so $F(B) = E_{2s+1}\bar{E}_{2r}$ where $s = a + b + c + d - 2 - r$, so $0 \le s \le a + b + c + d - 2$.

By Lemma 5.7, $F(E_{2n})E_{2r} = E_{2m+2r}$ for some $m$, and, on taking inverses, we see that all the generators of $\mathcal{M}$ carry $\bar{E}_{2n}\mathrm{K}$ into $Q$. Also

$$F(E_{2n+1})E_{2r} = F(E_{2n}B)E_{2r} = E_{2m}F(B)E_{2r} = E_{2m}E_{2s+1} = E_{2m+2s+1}.$$

On taking inverses, we see that the generators of $\mathcal{M}$ carry $\bar{E}_{2n+1}\mathrm{K}$ into $Q$.

Hence $Q$ is closed under the action of $\mathcal{M}$. We next show that each element of $Q$ lies in the $\mathcal{M}$-orbit of $\mathrm{K}$.

Suppose that, for some $n \in \mathbb{N}$, $\bar{E}_{2n}\mathrm{K}$ does not lie in $\mathcal{M}\mathrm{K}$. We may assume that $n$ is smallest possible, and we shall obtain a contradiction. There exists a unique $q \in \mathbb{N}$ such that $F(E_{2q})$ is an initial segment of $E_{2n}$, and $E_{2n}$ is an initial segment of $F(E_{2q+2})$ and (hence) of $F(E_{2q+1})$. In particular $q < n$, so, by minimality, $\bar{E}_{2q}\mathrm{K} \in \mathcal{M}\mathrm{K}$. Also, $E_{2n} = F(E_{2q})E_{2r}$ for some $r$ such that $0 \le r \le a+b+c+d-2$. On taking inverses, we see that $\bar{E}_{2r}F$ carries $\bar{E}_{2q}\mathrm{K}$ to $\bar{E}_{2n}\mathrm{K}$, which shows that $\bar{E}_{2n}$ does lie in $\mathcal{M}\mathrm{K}$, a contradiction. Thus all the $\bar{E}_{2n}\mathrm{K}$ lie in $\mathcal{M}\mathrm{K}$. Since $B \in \mathcal{M}$, we see that $\bar{E}_{2n+1}\mathrm{K} = B\bar{E}_{2n}\mathrm{K}$ also lies in $\mathcal{M}\mathrm{K}$, as desired. We have actually shown that a smaller semigroup than $\mathcal{M}$ has the same orbit.

Since $\mathcal{M} = \mathcal{M}B$, and $BA\mathrm{K} = C\mathrm{K}$, we see that $\mathcal{M}A\mathrm{K}$ contains $C\mathrm{K}$, and $\mathcal{M}A\mathrm{K} = \mathcal{M}C\mathrm{K}$.

Since $FA$ lies in $\mathcal{M}$, and $FAA\mathrm{K} = \mathrm{K}$, we see that $\mathcal{M}A\mathrm{K}$ contains $\mathcal{M}\mathrm{K}$, so, by the first part, $\mathcal{M}\mathrm{K} \supseteq Q$.

Suppose that $0 \le n \le a + b + c + d - 2$. To prove the result, it suffices to show that $\bar{E}_{2n}FA\mathrm{K}$ and $\bar{E}_{2n}FC\mathrm{K}$ lie in $Q \cup \{A\mathrm{K}, C\mathrm{K}\}$. We consider only the former, since the argument is similar for the latter.

Consider first the case where $n \le a+c-2$. Let $r = a+c-2-n$. Since $F(A) = E_{2(a+c)-3}$, it begins with $E_{2n}$ and with $E_{2r+1}$, and since it is of order two, it terminates in $\bar{E}_{2r+1}$. Thus $F(A) = E_{2n}\bar{E}_{2r+1}$, so $\bar{E}_{2n}FA\mathrm{K} = \bar{E}_{2n}F(A)\mathrm{K} = \bar{E}_{2r+1}\mathrm{K} \in Q$.

Consider next the case where $n = a+c-1$, so $E_{2n} = F(A)A$, so $\bar{E}_{2n}FA\mathrm{K} = \bar{E}_{2n}F(A)\mathrm{K} = A\mathrm{K}$.

Finally, consider the case were $a+c \le n \le a+b+c+d-2$, and let $r = n-a-c$. Then $F(BC) = E_{2(a+c)}$ is an initial segment of $E_{2n}$, and $E_{2n}$ is an initial segment of $F(B) = F(BC)F(C) = F(BC)E_{2(b+d)-3}$. Thus $E_{2n} = F(BC)E_{2r}$. Hence $\bar{E}_{2n}FA\mathrm{K} = \bar{E}_{2n}F(A)\mathrm{K} = \bar{E}_{2r}F(CB)F(A)\mathrm{K} = \bar{E}_{2r}CBA\mathrm{K} = \bar{E}_{2r}\mathrm{K} \in Q$. □

**5.9 Theorem.** *With Notation 5.1 and 5.2, the following hold.*

(i). *In* $\mathbb{R}$, $\mathcal{M}(0) = \mathrm{cusps}((-1,\mu))$ *and* $\mathcal{M}(-1) = \mathcal{M}(\mu) = \mathrm{cusps}([-1,\mu])$.

(ii). $\mathrm{cusps}([-1,\mu])$ *and* $[-1,\mu]$ *are* $\mathcal{M}$*-subspaces of* $\mathbb{R}$.

(iii). *In* $S^2$, $\mathcal{M}(\infty_{S^2}) = \mathrm{cusps}(\mathrm{arc}(F)) - \{0_{S^2}, 1_{S^2}\}$ *and* $\mathcal{M}(0_{S^2}) = \mathcal{M}(1_{S^2}) = \mathrm{cusps}(\mathrm{arc}(F))$.

(iv). $\mathrm{cusps}(\mathrm{arc}(F))$ *and* $\mathrm{arc}(F)$ *are* $\mathcal{M}$*-subspaces of* $S^2$.

(v). $\psi\colon [-1,\mu] \to \mathrm{arc}(F)$ *is a homeomorphism of* $\mathcal{M}$*-spaces.*

*Proof.* In $\mathbb{R}$, $F$ and $ABC$ fix 0, so, by Lemma 5.8,

$$\mathcal{M}(0) = \{\bar{E}_n(0) \mid n \in \mathbb{N}\} = \mathrm{cusps}((-1,\mu)).$$

By continuity, $\mathcal{M}$ acts on the closure $[-1,\mu]$. Thus (i) and (ii) hold.

In $S^2$, $F$ and $ABC$ fix $\infty_{S^2}$, so (iii) and (iv) hold.

Since $\psi(\bar{E}_n(0)) = \bar{E}_n(\infty_{S^2})$, we see that $\psi$ commutes with $\mathcal{M}$ on $\mathrm{cusps}((-1,\mu))$, so, by continuity, $\psi$ commutes with $\mathcal{M}$ on $[-1,\mu]$. Hence (v) holds. □

**5.10 Remarks.** (i). All the foregoing results have straightforward analogues for any hyperbolic $T$ in $\langle\langle R, L\rangle\rangle$. For example, if $T$ ends in $R$, then $\mu_- < -1$ and every other letter of $E^-$ is a $C$.

(ii). It is natural to ask for which $F \in \langle\langle R, L\rangle\rangle$ there exists a finite set of elements $F_1, \ldots, F_q$ in the normalizer of $\langle A, B, C, F\rangle$ in $\mathrm{Aut}\langle A, B, C\rangle$ such that each $F_i$ carries $\mathrm{arc}(F)$ to a proper subset of itself, and such that the resulting $q$ subsets cover all of $\mathrm{arc}(F)$ and overlap pairwise in at most a single point.

It was seen in [1] that this holds with $q = 2$ if $F = RL$; it is not difficult to show it holds with $q = 3$ if $F = R^2L^2$.

Notice that $\{F_1, \ldots, F_q\}$ is not a subset of $\langle A, B, C, F\rangle$, because each element of $\langle A, B, C, F\rangle$ acts on $\mathbb{R}$ as an isometry followed by scaling by a power of $\lambda^{-1}$, but a proper sum of powers of $\lambda^{-1}$ cannot equal 1.

(iii). Apart from 1 and $B$, each element of $\mathcal{M}$ acts on $[-1,\mu]$ as an isometry on $\mathbb{R}$ followed by scaling by a positive power of $\lambda^{-1}$, so has a unique fixed point, and has infinite order. Clearly $B$ has a unique fixed point in $[-1,\mu]$. Thus each element of $\mathcal{M} - \{1\}$ has a unique fixed point in $\mathrm{arc}(F)$.

We consider an example. By Lemma 5.7(vii), $FC \in \mathcal{M}$. We shall see in Section 6.3 that the inverse $C\bar{F} = (C\bar{F}C)C$ lies in $\mathcal{M}_A^-$ so has a fixed point $x$ in $\partial^-[A]_{S^2} = \mathrm{arc}(\bar{F})$. Thus $FC$ fixes $x$. Because $\mu$ is irrational, $FC$ does not fix $-1$, 0 or $\mu$ in $[-1,\mu]$, so $x$ is not $0_{S^2}$, $1_{S^2}$ or $\infty_{S^2}$. By Theorem 5.6(ii), $x$ does not lie in $\mathrm{arc}(F)$. Since $FC$ has infinite order, we see that $\mathcal{M}$ does *not* act discontinuously on the complement of $\mathrm{arc}(F)$ in $S^2$.

(iv). Consider any cusp $x$ in $\mathrm{arc}(F)$. There exist three parabolic elements $AFA$, $CFC$, and $F$, of $\mathcal{M}$ fixing $0_{S^2}$, $1_{S^2}$, and $\infty_{S^2}$, respectively. Also, we have proved there exist three elements of $\mathcal{M}$ carrying $0_{S^2}$, $1_{S^2}$, and $\infty_{S^2}$ to $x$, and these three elements carry $\mathrm{arc}(F)$ to one- or two-sided neighbourhoods of $x$ in $\mathrm{arc}(F)$. Hence there exist three parabolic elements of $\langle A, B, C, F\rangle$ fixing $x$ which respectively map these three neighbourhoods in $\mathrm{arc}(F)$ to themselves. □

## 6. Profiles of the six arcs.

Let us summarize some of the results of Section 5.

**6.1 Profile of $\partial^+[B]_{S^2}$.** Let $G = F$, $\mu = \mu_+$, and $E = E^+ = \prod_{n=1}^{\infty}((BC)^{[\mu]_n}BA)$. Let

$$\mathcal{M}_B^+ := \langle\langle B, \bar{E}_{2n}G \mid 0 \leq n \leq a+b+c+d-2\rangle\rangle.$$

Then

$$\mathrm{cusps}(\partial^+[B]_{S^2}) = \{0_{S^2}, 1_{S^2}, \bar{E}_n(\infty_{S^2}) \mid n \in \mathbb{N}\} = \mathcal{M}_B^+(0_{S^2}) = \mathcal{M}_B^+(1_{S^2}).$$

In words, the $\mathcal{M}_B^+$-orbit of each endpoint is the dense set of cusps in $\partial^+[B]_{S^2}$. In particular, $\partial^+[B]_{S^2}$ is an $\mathcal{M}_B^+$-subspace of $S^2$.

Let $\mathbb{R}$ have the structure of an $\langle A, B, C, F\rangle$-space with

$$A(x) := -x-1, \quad B(x) := -x+\mu-1, \quad C(x) := -x+\mu, \quad G(x) := \lambda^{-1}x.$$

We set $\mathrm{cusps}(\mathbb{R}) = \mathbb{Z}+\mathbb{Z}\mu$, an $\langle A, B, C, F\rangle$-subspace of $\mathbb{R}$. The interval $[-1, \mu]$ forms an $\mathcal{M}_B^+$-subspace of $\mathbb{R}$. Also, $\partial^+[B]_{S^2}$ and $[-1, \mu]$ are homeomorphic as $\mathcal{M}_B^+$-spaces, with a bijective correspondence between the cusps, and $0_{S^2}$, $\infty_{S^2}$ and $1_{S^2}$ correspond to $-1$, $0$, and $\mu$, respectively. □

We remark that the value $a+b+c+d-2$ can be replaced with $\max\{a+c, b+d\}-1$, if one wishes to have a smaller semigroup.

The other five arcs are formally similar. We record the descriptions to facilitate possible applications.

**6.2 Profile of $\partial^+[C]_{S^2}$.** Let $A' = BAB$, $B' = C$, $C' = B$. It is clear that $E^+$ can be partitioned into an infinite reduced word in $A'$, $B'$, $C'$. Also $C'A'B' = ABC$, so, if we identify $C'$ with the element of $\mathrm{Aut}\langle A, B, C\rangle$ which acts as conjugation by $C'$, and define $H = C'(A', B', C')$, then $H$ fixes $ABC$, so $H$ lies in $\langle R, L\rangle$. In fact $H = R$. Let

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \mathrm{matrix}(\bar{H}FH) = \begin{pmatrix} a-c & b-d \\ c & d \end{pmatrix},$$

$$\mu = \bar{H}(\mu_+) = \mu_+ - 1 = \frac{\lambda - c - d}{c} = \frac{\lambda - d'}{c'}.$$

We shall see that

$$\prod_{n=1}^{\infty}((B'C')^{[\mu]_n}B'A') = C'E^+.$$

We denote this end by $E$, and, for $n \in \mathbb{N}$, we define $E_n$ to be the initial segment of $E$ of length $n$ with respect to $A'$, $B'$ and $C'$. Here $a' + b' + c' + d' = a + b$, and we let $G = C'FC'$ and

$$\mathcal{M}^{+}_{C} := \langle\langle B', \bar{E}_{2n}G \mid 0 \le n \le a+b-2 \rangle\rangle.$$

Then $\bar{H}FH = \bar{R}FR \in \langle\langle R, L\rangle\rangle$, $H[B]_{\mathcal{E}} = C'[C]_{\mathcal{E}}$, and

$$\operatorname{cusps}(\partial^{+}[C]_{S^2}) = \{1_{S^2}, \infty_{S^2}, \bar{E}_n C'(\infty_{S^2}) \mid n \in \mathbb{N}\} = \mathcal{M}^{+}_{C}(1_{S^2}) = \mathcal{M}^{+}_{C}(\infty_{S^2}).$$

Now let $\mathbb{R}$ have the structure of an $\langle A, B, C, F\rangle$-space with

$$A'(x) := -x - 1, \quad B'(x) := -x + \mu - 1, \quad C'(x) := -x + \mu, \quad G(x) := \lambda^{-1}x.$$

The interval $[-1, \mu]$ forms an $\mathcal{M}^{+}_{C}$-subspace of $\mathbb{R}$, and $\partial^{+}[C]_{S^2}$ and $[-1, \mu]$ are homeomorphic as $\mathcal{M}^{+}_{C}$-spaces, with $1_{S^2}$, $C'(\infty_{S^2})$ and $\infty_{S^2}$ corresponding to $-1$, $0$, and $\mu$, respectively. $\square$

The information in the last paragraph is similar for the remaining four arcs and we shall omit it, and condense the remaining information.

**6.3 Profile of $\partial^{+}[A]_{S^2}$.** Let

$$A' = (BC)^{a_1}B, \quad B' = A, \quad C' = (BC)^{a_1+1}B,$$
$$H = A'(A', B', C') = R^{a_1}L, \quad \mu = \bar{H}(\mu_{+}) = \frac{\mu_{+} - a_1}{a_1 + 1 - \mu_{+}},$$
$$G = A'FA', \quad E = \prod_{n=1}^{\infty}((B'C')^{[\mu]_n}B'A') = A'E^{+},$$
$$\mathcal{M}^{+}_{A} := \langle\langle B', \bar{E}_{2n}G \mid 0 \le n \le c + 2a_1c + 2d - 2\rangle\rangle.$$

Then $\bar{H}FH = \bar{L}\bar{R}^{a_1}FR^{a_1}L \in \langle\langle R, L\rangle\rangle$, $H[B]_{\mathcal{E}} = A'[A]_{\mathcal{E}}$, and

$$\operatorname{cusps}(\partial^{+}[A]_{S^2}) = \{0_{S^2}, \infty_{S^2}, \bar{E}_n A'(\infty_{S^2}) \mid n \in \mathbb{N}\} = \mathcal{M}^{+}_{A}(\infty_{S^2}) = \mathcal{M}^{+}_{A}(0_{S^2}). \quad \square$$

We now consider $\bar{F}$.

By Definitions 2.1, the transpose of $F = \prod_{i=1}^{p}(R^{a_i}L^{b_i})$ is $F^{\mathrm{tr}} := \prod_{i=0}^{p-1}(R^{b_{p-i}}L^{a_{p-i}}) \in \langle\langle R, L\rangle\rangle$. Also, conjugation by $P = \bar{R}L\bar{R}$, and by $\bar{P} = R\bar{L}R$, interchanges $\bar{F}$ and $F^{\mathrm{tr}}$.

**6.4 Profile of $\partial^{-}[A]_{S^2}$.** Let

$$A' = C, \quad B' = A, \quad C' = B,$$
$$H = A'(A', B', C') = \bar{R}L, \quad \mu = \bar{H}(\mu_{-}) = \frac{1 + \mu_{-}}{-\mu_{-}},$$
$$G = A'\bar{F}A', \quad E = \prod_{n=1}^{\infty}((B'C')^{[\mu]_n}B'A') = A'E^{-},$$
$$\mathcal{M}^{-}_{A} := \langle\langle B', \bar{E}_{2n}G \mid 0 \le n \le 2a + c - 1\rangle\rangle.$$

Then $\bar{H}\bar{F}H = \bar{R}F^{\mathrm{tr}}R \in \langle\langle R, L\rangle\rangle$, $H[B]_{\mathcal{E}} = A'[A]_{\mathcal{E}}$, and

$$\operatorname{cusps}(\partial^{-}[A]_{S^2}) = \{0_{S^2}, \infty_{S^2}, \bar{E}_n A'(\infty_{S^2}) \mid n \in \mathbb{N}\} = \mathcal{M}^{-}_{A}(\infty_{S^2}) = \mathcal{M}^{-}_{A}(0_{S^2}). \quad \square$$

**6.5 Profile of $\partial^{-}[C]_{S^2}$.** Let

$$A' = (AB)^{b_p-1}A, \quad B' = C, \quad C' = (AB)^{b_p}A,$$
$$H = C'(A', B', C') = R\bar{L}R^{b_p+1}L, \quad \mu = \bar{H}(\mu_{-}) = \frac{b_p(-\mu_{-}) - 1}{1 + (b_p+1)\mu_{-}}$$
$$G = C'FC, \quad E = \prod_{n=1}^{\infty}((B'C')^{[\mu]_n}B'A') = C'(ABC)E^{-},$$
$$\mathcal{M}_C^{-} := \langle\langle B', \bar{E}_{2n}G \mid 0 \le n \le (2b_p+1)b + 2d - 2\rangle\rangle.$$

Then $\bar{H}\bar{F}H = (\bar{L}\bar{R}^{b_p})F^{\mathrm{tr}}R^{b_p}L \in \langle\langle R, L\rangle\rangle$, $H[B]_{\mathcal{E}} = C'[C]_{\mathcal{E}}$, and

$$\mathrm{cusps}(\partial^{-}[C]_{S^2}) = \{1_{S^2}, \infty_{S^2}, \bar{E}_n C'(\infty_{S^2}) \mid n \in \mathbb{N}\} = \mathcal{M}_C^{-}(1_{S^2}) = \mathcal{M}_C^{-}(\infty_{S^2}). \quad \square$$

**6.6 Profile of $\partial^{-}[B]_{S^2}$.** Let

$$m = \begin{cases} a_p & \text{if } b_p = 1 \\ 0 & \text{if } b_p \ge 2 \end{cases}, \quad A' = (AC)^{m+1}A, \quad B' = B, \quad C' = (AC)^m A,$$
$$H = A'A(A', B', C') = \bar{R}L^{m+1}R, \quad \mu = \bar{H}(\mu_{-}) = \frac{m+1+(m+2)\mu_{-}}{(m+1)(-\mu_{-}) - m},$$
$$G = A'A\bar{F}AA', \quad E = \prod_{n=1}^{\infty}((B'C')^{[\mu]_n}B'A' = A'AE^{-},$$
$$\mathcal{M}_B^{-} := \langle\langle B', \bar{E}_{2n}G \mid 0 \le n \le (2m+3)(a-b) + (m+2)(c-d) - 2\rangle\rangle.$$

Then $\bar{H}\bar{F}H = \bar{R}\bar{L}^m\bar{R}F^{\mathrm{tr}}RL^mR \in \langle\langle R, L\rangle\rangle$, $H[B]_{\mathcal{E}} = A'A[B]_{\mathcal{E}}$, and

$$\mathrm{cusps}(\partial^{-}[B]_{S^2}) = \{0_{S^2}, 1_{S^2}, \bar{E}_n A'A(\infty_{S^2}) \mid n \in \mathbb{N}\} = \mathcal{M}_B^{-}(0_{S^2}) = \mathcal{M}_B^{-}(1_{S^2}). \quad \square$$

## 7. Proofs for the six arcs.

In this section, we return to Notation 2.3 and verify the claims of Section 6. We need a preliminary result.

**7.1 Lemma.** *With Notation 2.3, the following hold.*

(i). $R[A]_{\mathcal{E}} = [A]_{\mathcal{E}}$, $R[B]_{\mathcal{E}} = [BC]_{\mathcal{E}}$, *and* $R[C]_{\mathcal{E}} = [BA]_{\mathcal{E}} \cup [C]_{\mathcal{E}}$.
(ii). $\bar{R}[A]_{\mathcal{E}} = [A]_{\mathcal{E}}$, $\bar{R}[B]_{\mathcal{E}} = [B]_{\mathcal{E}} \cup [CA]_{\mathcal{E}}$, *and* $\bar{R}[C]_{\mathcal{E}} = [CB]_{\mathcal{E}}$.
(iii). $L[A]_{\mathcal{E}} = [A]_{\mathcal{E}} \cup [BC]_{\mathcal{E}}$, $L[B]_{\mathcal{E}} = [BA]_{\mathcal{E}}$, *and* $L[C]_{\mathcal{E}} = [C]_{\mathcal{E}}$.
(iv). $\bar{L}[A]_{\mathcal{E}} = [AB]_{\mathcal{E}}$, $\bar{L}[B]_{\mathcal{E}} = [AC]_{\mathcal{E}} \cup [B]_{\mathcal{E}}$, *and* $\bar{L}[C]_{\mathcal{E}} = [C]_{\mathcal{E}}$.

*Proof.* (i). Since $R = (A, BCB, B)$, we see that

$$R((CBA)^{\infty}) = (CBA)^{\infty}, \; R((ACB)^{\infty}) = (ACB)^{\infty}, \text{ and } R((BAC)^{\infty}) = B(CBA)^{\infty},$$

so $R(\infty_{S^1}) = \infty_{S^1}$, $R(0_{S^1}) = 0_{S^1} = B(1_{S^1})$, and $R(1_{S^1}) = B(\infty_{S^1})$.

Thus $R$ fixes the endpoints of $[A]_{S^1}$, so $R[A]_{S^1} = [A]_{S^1}$, and hence $R[A]_{\mathcal{E}} = [A]_{\mathcal{E}}$.

Similarly, $R$ carries the endpoints of $[B]_{S^1}$ to the endpoints of $B[C]_{S^1}$ so $R[B]_{S^1} = B[C]_{S^1}$, and hence $R[B]_{\mathcal{E}} = B[C]_{\mathcal{E}}$.

Similarly, or by considering complements, we see $R[C]_{\mathcal{E}} = [BA]_{\mathcal{E}} \cup [C]_{\mathcal{E}}$.

(ii). Since $\bar{R} = (A, C, CBC)$, we see that $\bar{R}((CBA)^\infty) = (CBA)^\infty$, $\bar{R}((ACB)^\infty) = (ACB)^\infty$, and $\bar{R}((BAC)^\infty) = C(ACB)^\infty$. Thus $\bar{R}(\infty_{S^1}) = \infty_{S^1} = C(0_{S^1})$, $\bar{R}(0_{S^1}) = 0_{S^1}$, and $\bar{R}(1_{S^1}) = C(0_{S^1})$. Now the result follows as before.

(iii). Since $L = (B, BAB, C)$, we see that $L(\infty_{S^1}) = \infty_{S^1}$, $L(0_{S^1}) = B(\infty_{S^1})$, and $L(1_{S^1}) = 1_{S^1} = B(0_{S^1})$, and the result follows.

(iv). Since $\bar{L} = (ABA, A, C)$, we see that $\bar{L}(\infty_{S^1}) = \infty_{S^1} = A(0_{S^1})$, $\bar{L}(0_{S^1}) = A(1_{S^1})$, and $\bar{L}(1_{S^1}) = 1_{S^1}$, and the result follows. $\square$

**7.2 Theorem.** *With Notation 2.3, the descriptions given in Sections 6.1–6.6 hold.*

*Proof.* By Section 5, for any hyperbolic $F' \in \langle\langle R, L\rangle\rangle$, we have an arc in $S^2_{F'}$ denoted $\mathrm{arc}(F')$.

If $H$ is 1, $R$ or $R^{a_1}L$, and we take $F' = \bar{H}FH$, then it is clear that $F' \in \langle\langle R, L\rangle\rangle$ and $F'$ is hyperbolic. Notice that $H$ carries $\mathcal{E}^{F'}$ to $\mathcal{E}^F$, and determines a homeomorphism $S^2_{F'} \simeq S^2_F$ which carries $\mathrm{arc}(F')$ to an arc in $S^2_F$ which we denote $H(\mathrm{arc}(\bar{H}FH))$.

Now $\mathrm{arc}(F') \subseteq [B]_{S^2_{F'}} \cap [B^*]_{S^2_{F'}}$. Hence $H(\mathrm{arc}(\bar{H}FH)) \subseteq (H[B])_{S^2_F} \cap (H[B^*])_{S^2_F}$. This is an arc joining $(H(0))_{S^2_F}$ to $(H(1))_{S^2_F}$, that is, an arc joining the image in $S^2_F$ of $H(0_{S^1})$ to the image in $S^2_F$ of $H(1_{S^1})$.

For $H = 1$, as we know, we have $\mathrm{arc}(F) \subseteq [B]_{S^2_F} \cap [B^*]_{S^2_F}$, an arc joining $0_{S^2_F}$ to $1_{S^2_F}$.

For $H = R$, we have, by Lemma 7.1(i), $R[B]_{\mathcal{E}} = B[C]_{\mathcal{E}}$, so

$$R(\mathrm{arc}(\bar{R}FR)) \subseteq [BC]_{S^2_F} \cap [(BC)^*]_{S^2_F}.$$

On applying $B$, we get $BR(\mathrm{arc}(\bar{R}FR)) \subseteq [C]_{S^2_F} \cap [C^*]_{S^2_F}$, an arc joining $1_{S^2_F}$ to $\infty_{S^2_F}$.

For $H = R^{a_1}L$, we have, by Lemma 7.1(i) and (ii), $R^{a_1}L[B]_{\mathcal{E}} = R^{a_1}B[A]_{\mathcal{E}} = A'[A]_{\mathcal{E}}$, where $A' = R^{a_1}(B) = (BC)^{a_1}B$, so $R^{a_1}L(\mathrm{arc}(\bar{L}\bar{R}^{a_1}FR^{a_1}L)) \subseteq [A'A]_{S^2_F} \cap [(A'A)^*]_{S^2_F}$. On applying $A'$ we get $A'R^{a_1}L(\mathrm{arc}(\bar{L}\bar{R}^{a_1}FR^{a_1}L)) \subseteq [A]_{S^2_F} \cap [A^*]_{S^2_F}$, an arc joining $\infty_{S^2_F}$ to $0_{S^2_F}$.

By Definitions 2.1, $F^{\mathrm{tr}} = P\bar{F}\bar{P} = \bar{P}\bar{F}P$ lies in $\langle\langle R, L\rangle\rangle$ and is hyperbolic. Let $m = \begin{cases} a_p & \text{if } b_p = 1 \\ 0 & \text{if } b_p \geq 2 \end{cases}$, so $F^{\mathrm{tr}}RL^mR$ starts with $RL^mR$. If $H$ is $PR$, $\bar{P}R^{b_p}L$ or $PRL^mR$, and we take $F' = \bar{H}\bar{F}H$, then $F' \in \langle\langle R, L\rangle\rangle$ and $F'$ is hyperbolic. As before, $H$ carries $\mathcal{E}^{F'}$ to $\mathcal{E}^{\bar{F}} = \mathcal{E}^F$, and $H$ determines a homeomorphism $S^2_{F'} \simeq S^2_F$ which carries $\mathrm{arc}(F')$ to an arc in $S^2_F$ which we denote $H(\mathrm{arc}(\bar{H}\bar{F}H))$, and which is an arc joining $(H(0))_{S^2_F}$ to $(H(1))_{S^2_F}$ which lies in $(H[B])_{S^2_F} \cap (H[B^*])_{S^2_F}$.

For $H = PR = \bar{R}L$, we have

$$H[B]_{\mathcal{E}} = \bar{R}L[B]_{\mathcal{E}} = \bar{R}[BA]_{\mathcal{E}} = [CA]_{\mathcal{E}},$$

so $H(\mathrm{arc}(\bar{H}\bar{F}H)) \subseteq [CA]_{S^2_F} \cap [(CA)^*]_{S^2_F}$, an arc joining $1_{S^2_F}$ to $C(0_{S^2_F})$. On applying $C$, we get $CH(\mathrm{arc}(\bar{H}\bar{F}H)) \subseteq [A]_{S^2_F} \cap [A^*]_{S^2_F}$, an arc joining $\infty_{S^2_F}$ to $0_{S^2_F}$.

For $H = \bar{P}R^{b_p}L = R\bar{L}R^{b_p+1}L$, we have

$$\begin{aligned} H[B]_{\mathcal{E}} &= R\bar{L}R^{b_p+1}L[B]_{\mathcal{E}} = R\bar{L}R^{b_p+1}[BA]_{\mathcal{E}} = R\bar{L}(BC)^{b_p+1}B[A]_{\mathcal{E}} \\ &= R(AC)^{b_p+1}A[AB]_{\mathcal{E}} = (AB)^{b_p+1}[BC]_{\mathcal{E}} = C'[C]_{\mathcal{E}}, \end{aligned}$$

where $C' = (AB)^{b_p} A$. Thus $C'H(\mathrm{arc}(\bar{H}\bar{F}H)) \subseteq [C]_{S^2_F} \cap [C^*]_{S^2_F}$, an arc joining $1_{S^2_F}$ to $\infty_{S^2_F}$.

For $H = PRL^m R = \bar{R}L^{m+1}R$, we have

$$\begin{aligned} H[B]_{\mathcal{E}} &= \bar{R}L^{m+1}R[B]_{\mathcal{E}} = \bar{R}L^{m+1}[BC]_{\mathcal{E}} \\ &= \bar{R}(BA)^{m+1}B[C]_{\mathcal{E}} = (CA)^{m+1}C[CB]_{\mathcal{E}} = A'A[B]_{\mathcal{E}}, \end{aligned}$$

where $A' = (CA)^m C$. Thus $AA'H(\mathrm{arc}(\bar{H}\bar{F}H)) \subseteq [B]_{S^2_F} \cap [B^*]_{S^2_F}$, an arc joining $0_{S^2_F}$ to $1_{S^2_F}$.

It remains to check that the semigroups behave as claimed in Sections 6.2-6.6, and this follows from the fact that, in each case, we chose $H$ so as to respect initial segments of fixed ends. $\square$

## 8. The relationship with hyperbolic geometry.

We continue using Notation 2.3.

**8.1 Background.** The action of $F$ on $\mathbb{R}^2 - \mathbb{Z}^2$ of Definitions 3.1 determines a pseudo-Anosov automorphism of the once-punctured torus $(\mathbb{R}^2 - \mathbb{Z}^2)/\mathbb{Z}^2$, and every pseudo-Anosov automorphism of the once-punctured torus is of this form, or at least its square is.

Jørgensen showed that the mapping torus of this automorphism has a complete finite-volume hyperbolic structure as the union of finitely many ideal tetrahedra; see [14], [12, Appendix], or [2, Section 1]. Hence the fundamental group of the mapping torus, $\mathrm{F}_2 \rtimes \langle F \rangle = \langle X, Y, F \rangle$ has a discrete faithful representation in $\mathrm{PSL}_2(\mathbb{C})$ with $F$ and the commutator being parabolic (with the same fixed point).

Multiplication by $-1$ induces an involution on $(\mathbb{R}^2 - \mathbb{Z}^2)/\mathbb{Z}^2$, and dividing out by the action of this involution gives a once-punctured sphere with three double points. The pseudo-Anosov automorphism commutes with the involution, so the pseudo-Anosov automorphism induces a pseudo-Anosov automorphism on the once-punctured sphere with three double points, and, at the same time, the involution extends to the hyperbolic mapping torus. By the Mostow Rigidity Theorem, the extended involution is an isometry, so we get a hyperbolic quotient space which is clearly the mapping torus for the pseudo-Anosov automorphism of the once-punctured sphere with three double points. This latter space has $\langle A, B, C, F \rangle$ as orbifold group. Hence the boundary of hyperbolic three-space $\partial\mathbf{H}^3$, identified with the Riemann sphere $\overline{\mathbb{C}}$, is an $\langle A, B, C, F \rangle$-space, and we have a discrete faithful representation of $\langle A, B, C, F \rangle$ in $\mathrm{PSL}_2(\mathbb{C})$ with $F$ and $ABC$ parabolic (with the same fixed point). The universal cover of the hyperbolic mapping torus has cusps in $\overline{\mathbb{C}}$ which are the cusps of the ideal tetrahedra.

We normalize so that the fixed points of the parabolic elements $ABC$, $BCA$ and $CAB$ are $\infty_{\overline{\mathbb{C}}}$, $0_{\overline{\mathbb{C}}}$ and $1_{\overline{\mathbb{C}}}$, respectively. Thus the cusps form the $\langle A, B, C \rangle$-orbit of $\infty_{\overline{\mathbb{C}}}$.

The Mostow Rigidity Theorem implies that, with this normalization, the representation of $\langle A, B, C, F \rangle$ in $\mathrm{PSL}_2(\mathbb{C})$ is unique up to complex conjugation.

Thus we have two $\langle A, B, C, F \rangle$-spaces, $S^2$ and $\overline{\mathbb{C}}$, which are two-spheres.

Bowditch [3, Theorem 9.1] has recently shown there exists a (unique) homeomorphism of $\langle A, B, C, F \rangle$-spaces $S^2 \simeq \overline{\mathbb{C}}$. See also [6].

(We remark that the case $F = RL$ of this result can be proved by using results of [1]. The details are somewhat tedious and we will not go into them, but the idea is elementary and we will sketch it. Here it is important that the square of $RM$ is $RL$. The normalization used in [1] can be transformed into the normalization used here. Then [1, Section 8] shows that there is a two-generator semigroup acting on an arc in $\overline{\mathbb{C}}$ and on $\partial^{+}[B]_{S^2}$ in such a way that both these arcs can be identified with the set of ends of the semigroup modulo a certain equivalence relation. It follows that the natural map from $\partial^{+}[B]_{S^2}$ to $\overline{\mathbb{C}}$ is an embedding. Moreover, it can be shown that the identifications which take place under $\mathcal{E} \to \overline{\mathbb{C}}$ and $\mathcal{E} \to S^2$ are determined by these arcs, together with the $\langle A, B, C\rangle$-action. It then follows that $S^2$ and $\overline{\mathbb{C}}$ are quotient spaces of $\mathcal{E}$ with the same identifications, as desired.)

The homeomorphism $S^2 \simeq \overline{\mathbb{C}}$ carries $\infty_{S^2}$, to $\infty_{\overline{\mathbb{C}}}$ since these are the (unique) points fixed by $ABC$. Similar statements hold for 0 and for 1.

We may assume that $S^2$ and $\overline{\mathbb{C}}$ are homeomorphic as *oriented* $\langle A, B, C, F\rangle$-spaces, since we can replace the representation of $\langle A, B, C, F\rangle$ with its complex conjugate, if necessary.

We remark that both $ABC$ and $F$ fix $\infty_{\overline{\mathbb{C}}}$, so $F(z) = z + s$ for some $s \in \mathbb{C}$, and it can be shown that the imaginary part of $s$ is non-negative.

The composite map $\mathcal{E} \to S^2 \to \overline{\mathbb{C}}$ will be denoted $E \mapsto E_{\overline{\mathbb{C}}}$. For any $G \in \langle A, B, C\rangle$, we define $[G]_{\overline{\mathbb{C}}}$ similarly. □

**8.2 Consequences.** By the foregoing, we have a graph with two vertices, 1 and $\infty$, and three edges, embedded in $\overline{\mathbb{C}}$, determining three topological closed disks $[A]_{\overline{\mathbb{C}}}$, $[B]_{\overline{\mathbb{C}}}$, $[C]_{\overline{\mathbb{C}}}$ whose boundaries are Jordan curves.

The interpretation of these disks in the boundary of hyperbolic three-space is the following. We have a fibration $\mathbf{H}^3 \to \mathbb{R}$. Each fiber is a topological disk on which $\langle A, B, C\rangle$ acts with quotient a once-punctured sphere with three double points. And $F$ acts on $\mathbb{R}$ with quotient a circle. For any $v \in \mathbf{H}^3$, the $\langle A, B, C\rangle$-orbit of $v$ is the vertex set of a copy of the usual $\langle A, B, C\rangle$-tree forming the spine of a fiber, and the ends of the tree reach every point of the boundary $\overline{\mathbb{C}}$. Then $[A]_{\overline{\mathbb{C}}}$ is the set of points reached by infinite reduced paths in the tree which start with the step from $v$ to $Av$.

As in $S^2$, the two distinguished points 0 and $\infty$ divide $\partial[A]_{\overline{\mathbb{C}}}$ into two arcs $\partial^{+}[A]_{\overline{\mathbb{C}}}$, $\partial^{-}[A]_{\overline{\mathbb{C}}}$.

There are four other arcs obtained similarly. Each arc has a dense set of cusps, and these cusps form a single orbit of a semigroup of Möbius transformations lying in $\langle A, B, C, F\rangle$.

To depict an approximation of $\partial^{+}[B]_{\overline{\mathbb{C}}}$, we choose a large even number $N$, and generate $(x_n, z_n) := (\bar{E}^{+}_n(0_{\mathbb{R}}), \bar{E}^{+}_n(\infty_{\overline{\mathbb{C}}}))$, for $0 \le n \le N$, using Theorem 5.5(iii), as follows. We have $(x_0, z_0) = (0, \infty)$, and, given $(x_{2m}, z_{2m})$ with $0 \le 2m < N$, we have

$$(x_{2m+1}, z_{2m+1}) = (-x_{2m} + \mu_{+} - 1, B(z_{2m})),$$
$$(x_{2m+2}, z_{2m+2}) = \begin{cases} (x_{2m} + 1, CB(z_{2m})) & \text{if } x_{2m} < \mu_{+} - 1, \\ (x_{2m} - \mu_{+}, AB(z_{2m})) & \text{if } x_{2m} > \mu_{+} - 1. \end{cases}$$

We can now order the $(x_n, z_n)$ by the first coordinate, list the second coordinates, add $0_{\mathbb{C}}$ at the beginning and $1_{\mathbb{C}}$ at the end, and join the dots.

To depict an approximation of $\partial^{-}[A]_{\overline{\mathbb{C}}}$, we proceed as above, but replace $0_{\infty}$, $\infty_{\mathbb{C}}$, $1_{\infty}$, $\mu_{+}$, $B$, $CB$ and $AB$ with $\infty_{\mathbb{C}}$, $1_{\mathbb{C}}$, $0_{\mathbb{C}}$, $\frac{1+\mu_{-}}{-\mu_{-}}$, $A$, $BA$ and $CA$, respectively. □

**8.3 Example.** In Example 2.4, we noted that for

$$F = RL^3 = (BCBABCBABCB, BCBABCBABCBABCB, B),$$

$\mu_+ = \frac{3+\sqrt{21}}{6}$, $\mu_- = \frac{3-\sqrt{21}}{6}$. It is straightforward to calculate that the normalized representation is

$$A(z) = -\frac{s^2+2s+1}{z}, \quad B(z) = \frac{z-1}{\frac{z}{s+1}-1}, \qquad C(z) = \frac{z+s^2+s-1}{z-1}, \quad F(z) = z+s,$$

where $s^4+4s^3+9s^2+9s+4=0$ and $s$ is approximately $-0.8+0.6i$. (Recall that we want a non-negative imaginary part.)

We view the real projective line $\overline{\mathbb{R}}$ as the equator of the Riemann sphere $\overline{\mathbb{C}}$, and we make the convention that $i$ is in the northern hemisphere, and $-i$ in the southern hemisphere. Thus the northern hemisphere is the upper half of the complex plane, and the southern hemisphere is the lower half of the complex plane. We can represent the northern hemisphere conformally as a disk, and the southern hemisphere anti-conformally as a disk with the same boundary as the northern hemisphere. Consider Fig. 5. On the left is the northern hemisphere and the join-the-dots for the first 30,000 cusps in $\partial^+_{RL^3}[B]_{\mathbb{C}}$; on the right is the southern hemisphere and $\partial^-_{RL^3}[A]_{\mathbb{C}}$.

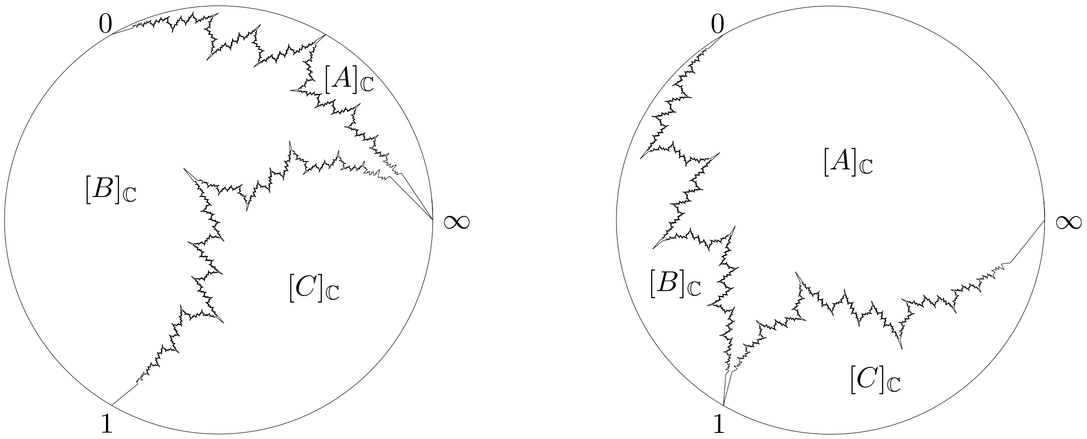


FIG. 5. The northern and southern hemispheres with the arcs for $RL^3$.

Let us digress to remark that, for this example, there exists a (faithful) non-discrete representation which is the same as the discrete representation algebraically, but now with $s$ near $-1.2+1.7i$. Fig. 6 shows the result of applying the join-the-dots procedure with a mere ten cusps. This is a rather bizarre way of testing for discreteness. □

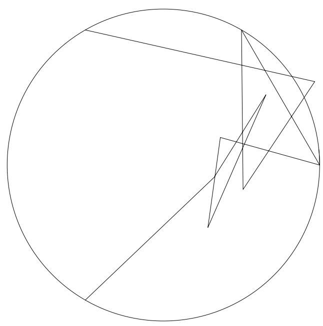

FIG. 6. Consequences of non-discreteness.

**8.4 Questions.** Does $\partial^+_F[B]_{\overline{\mathbb{C}}}$ lie in the northern hemisphere except for finitely many points on the equator? (*Added in proof*: For $F = RLRL^2$ the answer is 'no'.)

Does $\partial^-_F[A]_{\overline{\mathbb{C}}}$ lie in the southern hemisphere except for finitely many points on the equator?

The referee has observed the following.

**8.5 Remarks.** For any two distinct irrational numbers $\mu_+$ and $\mu_-$, one can define $E^+$ and $E^-$ by reading off half-lines of slopes $\frac{1}{\mu_+}$, $\frac{1}{\mu_-}$ starting at the origin, as in Fig. 4, and construct a two-sphere $S^2 = S^2_{\mu_+,\mu_-}$ with an $\langle A, B, C\rangle$-action as in Definitions 3.4 and the Appendix. Results analogous to Lemma 5.3, Theorem 5.5 and Theorem 5.6 can be obtained for this construction. Thus we find a distinguished graph on $S^2$. If $\mu_+\mu_- < 0$ it has three edges and two vertices (each of valence three).

By results of Bonahon, Minsky, McMullen and Thurston, there exists a discrete faithful representation $\rho\colon \langle A, B, C\rangle \to \mathrm{PSL}_2(\mathbb{C})$ with $\rho(ABC)$ parabolic such that, in a natural sense, $\rho(E^+)$ and $\rho(E^-)$ converge to the fixed point of $\rho(ABC)$; if $\rho$ is normalized so that the fixed points of $ABC$, $BCA$ and $CAB$ are $\infty_{\overline{\mathbb{C}}}$, $0_{\overline{\mathbb{C}}}$ and $1_{\overline{\mathbb{C}}}$, respectively, then $\rho$ is unique up to complex conjugation.

In detail, by results of Thurston and Bonahon, there exists a discrete faithful representation $\rho\colon \langle A, B, C\rangle \to \mathrm{PSL}_2(\mathbb{C})$ with $\rho(ABC)$ parabolic such that the "ending laminations" are the ordered pair $(\mu_+, \mu_-)$. Minsky [17] showed that $\rho$, when normalized, is unique. McMullen [16] showed that there exists a surjective continuous map of $\langle A, B, C\rangle$-spaces $S^2 \to \overline{\mathbb{C}}$.

One can conjecture that $S^2 \to \overline{\mathbb{C}}$ is a homeomorphism, and Bowditch [3, Theorem 9.1] has proved this in certain cases. Where it is true, one can depict subdivisions of $\overline{\mathbb{C}}$ into smaller and smaller Jordan domains, provided one has a sufficiently precise description of $\rho$. □

## 9. The exceptional sets.

**9.1 Definitions.** From Definitions 3.4, we have quotient maps of $\langle A, B, C, F\rangle$-spaces $\mathcal{E} \to S^1 \to S^2$, the latter being a Peano curve.

Let $X_{S^2}$ denote the set of points of $S^2$ which are the images of two or more points of $S^1$. Let $X_{S^1}$ denote the preimage in $S^1$ of $X_{S^2}$. Let $X_{\mathcal{E}}$ denote the preimage in $\mathcal{E}$ of $X_{S^1}$, or equivalently $X_{S^2}$. We call $X_{\mathcal{E}}$, $X_{S^1}$ and $X_{S^2}$ the *exceptional* sets. □

In this section we shall make some deductions about these exceptional sets.

**9.2 Observations.** The exceptional sets are closed under the actions of $\langle A, B, C, F\rangle$.

Notice that the quotient maps $\mathcal{E} \to S^1 \to S^2$ then decompose into maps of $\langle A, B, C, F\rangle$-spaces

$$X_{\mathcal{E}} \to X_{S^1} \to X_{S^2}, \quad \mathcal{E} - X_{\mathcal{E}} \to S^1 - X_{S^1} \to S^2 - X_{S^2}.$$

The latter two maps are both homeomorphisms. The map $X_{S^1} \to X_{S^2}$ is many-to-one.

We have seen that each cusp of $S^2$ is the image of countably many elements of $S^1$, so all the cusps lie in the exceptional sets for both $S^1$ and $S^2$.

The Peano curve $S^1 \to S^2$ fills in $[A]_{S^2}$ and then fills in the closure of the complement, so every point of $\partial[A]_{S^2}$ lies in $X_{S^2}$, since both the starting and finishing points are cusps. Thus we see that each of the six arcs described in Section 6 lies in $X_{S^2}$. It follows that $X_{S^2}$ contains the union of the $\langle A, B, C\rangle$-orbits of the points of the six arcs.

To see that this union is all of $X_{S^2}$, consider any point $x$ of $S^2$ which does not lie in this union. As $S^2$ successively decomposes into subsets of the form $[G]_{S^2}$ for longer and longer $G \in \langle A, B, C\rangle$, we see that $x$ never lies on a boundary, so determines an increasing sequence of terms in $\langle A, B, C\rangle$, which means that $x$ is the image of a unique element of $\mathcal{E}$, as claimed.

From Notation 3.5, we know that each element of $X_{\mathcal{E}}$ has a tail which is $(ABC)^{\pm\infty}$, or has $B$ as every alternate letter, or has $A$ as every alternate latter; this allows us to construct a corresponding partition

$$X_{\mathcal{E}} = X^0_{\mathcal{E}} \cup X^{+}_{\mathcal{E}} \cup X^{-}_{\mathcal{E}}.$$

Explicitly,

$$X^0_{\mathcal{E}} = \{G \cdot (ABC)^\infty, G \cdot (CBA)^\infty \mid G \in \langle A, B, C\rangle\},$$

and

$$X^{+}_{\mathcal{E}} = \{G \cdot E(x, \mu_{+}) \mid G \in \langle A, B, C\rangle, x \in [-1, \mu_{+}]\}.$$

The structure of $X^{-}_{\mathcal{E}}$ is formally similar to that of $X^{+}_{\mathcal{E}}$; from Section 6.4, we see that we have to consider $C\bar{F}C$ in place of $F$, and let $A$, $B$, $C$ take the roles of $B$, $C$, $A$, respectively.

For any elements $W_1$, $W_2$ of $\langle A, B, C\rangle$ which generate the free subgroup of index two, let $\mathcal{E}\langle\langle W_1, W_2\rangle\rangle$ denote the set of right infinite words in $W_1$ and $W_2$, viewed as a subset of $\mathcal{E}$. We emphasize that we are considering positive words only, and that $W_1^{-1}$ and $W_2^{-1}$ do not occur. □

We want to show that $X_{\mathcal{E}}$ is small in some sense. It suffices to consider only $X^{+}_{\mathcal{E}}$, since $X^0_{\mathcal{E}}$ is countable, and any results about $X^{+}_{\mathcal{E}}$ give analogous results about $X^{-}_{\mathcal{E}}$.

**9.3 Lemma.** *With Notations 2.3 and 9.1, for each $n \geq 0$,*

$$X^{+}_{\mathcal{E}} \subseteq \langle A, B, C\rangle \cdot \mathcal{E}\langle\langle F^n(BA), F^n(BC)\rangle\rangle.$$

*Proof.* It is not difficult to see that the elements of $X_{\mathcal{E}}$, $X^0_{\mathcal{E}}$, and $X^{+}_{\mathcal{E}}$ are permuted by $F$.

Since every element of $X^{+}_{\mathcal{E}}$ has a tail in which every alternate letter is $B$, we see that $X^{+}_{\mathcal{E}} \subseteq \langle A, B, C\rangle \cdot \mathcal{E}\langle\langle BA, BC\rangle\rangle$.

Hence

$$X^{+}_{\mathcal{E}} = F^n(X^{+}_{\mathcal{E}}) \subseteq F^n(\langle A, B, C\rangle \cdot \mathcal{E}\langle\langle BA, BC\rangle\rangle) = \langle A, B, C\rangle \cdot \mathcal{E}\langle\langle F^n(BA), F^n(BC)\rangle\rangle. \quad \square$$

Section 8.1 describes the relationship between the $\langle A, B, C, F\rangle$-spaces $S^2$ and $\overline{\mathbb{C}}$. In contrast, $S^1$ has the structure of an $\mathrm{Aut}\langle A, B, C\rangle$-space, but there is no natural $\mathrm{Aut}\langle A, B, C\rangle$-space structure on $\overline{\mathbb{R}}$. However, there is a natural *family* of $\mathrm{Aut}\langle A, B, C\rangle$-space structures on $\overline{\mathbb{R}}$, which we now describe.

**9.4 Definitions.** Let $\mathbb{S}^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 = 1\}$, with the (great circle) arc-length metric induced from $\mathbb{R}^2$, so $\mathbb{S}^1$ is geodesically connected.

Let us identify $\overline{\mathbb{R}}$ with $\mathbb{S}^1$ by identifying $t \in \overline{\mathbb{R}}$ with $(\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1}) \in \mathbb{S}^1$. Let dist denote the metric induced on $\overline{\mathbb{R}}$ by the arc-length metric on $\mathbb{S}^1$, so $\overline{\mathbb{R}}$ is geodesically connected.

Any $h \in \mathrm{PSL}_2(\mathbb{R})$ acts on $\overline{\mathbb{R}}$, and we define the *local dilatation* of $h$ by

$$h'(x) = \limsup_{t \to x} \frac{\mathrm{dist}(h(x), h(t))}{\mathrm{dist}(x,t)}.$$

If $h = \pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, with $ad - bc = 1$, then it is not difficult to show that

$$h'(x) = \lim_{t \to x} \frac{\mathrm{dist}(h(x), h(t))}{\mathrm{dist}(x,t)} = \frac{1 + x^2}{(ax+b)^2 + (cx+d)^2}.$$

Throughout the remainder of this section, we fix a discrete faithful representation $\rho\colon \langle A, B, C\rangle \to \mathrm{PSL}_2(\mathbb{R})$ such that $\rho(ABC)$ is parabolic. Moreover, we assume that $\rho$ has been normalized so that the fixed points in $\overline{\mathbb{R}}$ of $\rho(ABC)$, $\rho(BCA)$, and $\rho(CAB)$ are $\infty_{\overline{\mathbb{R}}}$, $0_{\overline{\mathbb{R}}}$, and $1_{\overline{\mathbb{R}}}$ respectively.

This gives $\overline{\mathbb{R}}$ the structure of an $\langle A, B, C\rangle$-space. It is not difficult to show that there is then a natural map $\mathcal{E} \to \overline{\mathbb{R}}$ which induces a homeomorphism of $\langle A, B, C\rangle$-spaces $S^1 \simeq \overline{\mathbb{R}}$; see Fig. 3. Moreover, $\infty_{S^1}$, $0_{S^1}$ and $1_{S^1}$ map to $\infty_{\overline{\mathbb{R}}}$, $0_{\overline{\mathbb{R}}}$ and $1_{\overline{\mathbb{R}}}$, respectively.

Clearly, we can use this homeomorphism to make $\overline{\mathbb{R}}$ into an $\mathrm{Aut}\langle A, B, C\rangle$-space and $S^1$ into a metric space. We identify $S^1 = \overline{\mathbb{R}} = \mathbb{S}^1$. Thus $[A]_{S^1}$, $[B]_{S^1}$ and $[C]_{S^1}$ can be thought of as the intervals $[-\infty, 0]$, $[0, 1]$ and $[1, \infty]$, respectively. Henceforth, we view $S^1$ as a metric space. □

**9.5 Theorem.** *The Hausdorff dimension of $X_{S^1}$ is zero, and hence the one-dimensional Lebesgue measure of $X_{S^1}$ is zero.*

*Proof.* Let $m$ and $n$ denote integers greater than 3.

Recall that $E^{+}_{n}$ denotes the initial segment of $E^{+} = F^{\infty}(B)$ of length $n$, and $\bar{E}^{+}_{n}$ denotes its inverse. Thus $E^{+}_{2n} \in \langle\langle BA, BC\rangle\rangle$, and $\lim\limits_{n\to\infty} E^{+}_{2n} = F^{\infty}(B)$.

We can decompose $S^1$ into the sequence

$$\begin{gathered}
[ABCA]_{S^1}, [ABCB]_{S^1}, [ABA]_{S^1}, [AC]_{S^1}, \\
[B]_{S^1}, [CA]_{S^1}, [CBC]_{S^1}, [CBAB]_{S^1}, [CBAC]_{S^1},
\end{gathered}$$

of nine intervals of positive length, each one overlapping in one endpoint with the next, cyclically; see Fig. 3.

Every second letter of $\bar{E}^{+}_{2n}$ is a $B$, so there is no cancellation in forming the product of $\bar{E}^{+}_{2n}$ with $(ABC)^{\infty}$ and $(ACB)^{\infty}$, so

$$\bar{E}^{+}_{2n} \cdot (ABC)^{\infty}, \bar{E}^{+}_{2n} \cdot (ACB)^{\infty} \in [ABCB]_{\mathcal{E}} \cup [ABAB]_{\mathcal{E}} \cup [CBCB]_{\mathcal{E}} \cup [CBAB]_{\mathcal{E}}.$$

Hence, in $\mathcal{E}$, both $\bar{E}^{+}_{2n}(ABC)^{\infty}$ and $\bar{E}^{+}_{2n}(ACB)^{\infty}$ are bounded away from both $[B]_{\mathcal{E}}$ and $(ABC)^{\pm\infty}$ (independently of $n$).

Hence, in $S^1$, both $\bar{E}^{+}_{2n}(\infty)$ and $\bar{E}^{+}_{2n}(0)$ are bounded away from both $[0, 1]$ and $\infty$.

Let $\pm\begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix} \in \mathrm{PSL}_2(\mathbb{R})$ denote $\rho(E^+_{2n})$. Hence, both $\bar{E}^+_{2n}(\infty) = -\frac{d_n}{c_n}$ and $\bar{E}^+_{2n}(0) = -\frac{b_n}{a_n}$ are bounded away from both $[0,1]$ and $\infty$. Thus there exist positive constants $k$, $K$ such that, for all $x \in [0,1]$, and all $n \geq 1$,

$$(x + \frac{d_n}{c_n})^2, (x + \frac{b_n}{a_n})^2 \in (k, \infty) \text{ and } (\frac{d_n}{c_n})^2, (\frac{b_n}{a_n})^2 \in [0, K).$$

The $\rho(E^+_{2n})$ are all distinct elements of $\mathrm{PSL}_2(\mathbb{R})$, and form a discrete subset, so $\lim\limits_{n\to\infty} a_n^2 + b_n^2 + c_n^2 + d_n^2 = \infty$. Now

$$a_n^2 + b_n^2 + c_n^2 + d_n^2 = a_n^2(1 + \frac{b_n^2}{a_n^2}) + c_n^2(1 + \frac{d_n^2}{c_n^2}) \leq (a_n^2 + c_n^2)(1 + K),$$

so $\lim\limits_{n\to\infty} a_n^2 + c_n^2 = \infty$.

Now consider any $x \in [0,1]$. For the action of $E^+_{2n}$ on $S^1$,

$$(E^+_{2n})'(x) = \frac{1 + x^2}{(a_n x + b_n)^2 + (c_n x + d_n)^2} = \frac{1 + x^2}{a_n^2(x + \frac{b_n}{a_n})^2 + c_n^2(x + \frac{d_n}{c_n})^2} < \frac{2}{k(a_n^2 + c_n^2)}.$$

So, for any $\epsilon \in (0,1)$, there exists $n$ such that, for all $m \geq n$ and all $x \in [0,1]$, $(E^+_{2m})'(x) < \epsilon$. By Lemma 5.7, it follows that, for all $x \in [0,1]$, $(F^n(BA))'(x) < \epsilon$ and $(F^n(BC))'(x) < \epsilon$. Since $[0,1]$ is geodesically connected, $F^n(BA)$ and $F^n(BC)$ are Lipschitz contractions on $[0,1]$ with constant $\epsilon$. It follows, by a standard argument, that the limit set of $F^n(BA)$ and $F^n(BC)$ acting on $[0,1]$ has Hausdorff dimension at most $-\frac{\log 2}{\log \epsilon}$; see, for example [11, Proposition 9.6].

Hence the image of

$$\mathcal{E}\langle\langle F^n(BA), F^n(BC)\rangle\rangle$$

in $S^1$ has Hausdorff dimension at most $-\frac{\log 2}{\log \epsilon}$.

By Lemma 9.3, the image of $X^+_{\mathcal{E}}$ in $S^1$ has Hausdorff dimension at most $-\frac{\log 2}{\log \epsilon}$, for all $\epsilon > 0$, so has Hausdorff dimension zero.

For purely formal reasons, the image of $X^-_{\mathcal{E}}$ in $S^1$ must also have Hausdorff dimension zero.

Hence the image, $X_{S^1}$, of $X_{\mathcal{E}}$ in $S^1$ has Hausdorff dimension zero, as desired. □

**9.6 Remarks.** Let $\mathbb{S}^2 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$, with the arc-length metric induced from $\mathbb{R}^3$.

Let us identify $\overline{\mathbb{C}}$ with $\mathbb{S}^2$ by identifying $x + iy \in \overline{\mathbb{C}}$ with

$$(\frac{2x}{x^2 + y^2 + 1}, \frac{2y}{x^2 + y^2 + 1}, \frac{x^2 + y^2 - 1}{x^2 + y^2 + 1}) \in \mathbb{S}^2.$$

Let dist denote the metric induced on $\overline{\mathbb{C}}$ by the arc-length metric on $\mathbb{S}^2$.

By [3, Theorem 9.1], there is a unique identification $S^2 = \overline{\mathbb{C}}$ of $\langle A, B, C, F\rangle$-spaces, and we identify $S^2 = \overline{\mathbb{C}} = \mathbb{S}^2$. Henceforth, $S^2$ is a metric space.

One can ask if the Hausdorff dimension of $X_{S^2}$ lies strictly between 1 and 2. One can also ask if each of the six arcs of Sections 6.1–6.6 have the same Hausdorff dimension as each other, and hence, as $X_{S^2}$.

In [10], for the case where $F = RL$, the answers were seen to be affirmative, and evidence was adduced that the Hausdorff dimension involved is approximately 1.2971. □

**Acknowledgements.** All the figures were produced using the Wolfram Research, Inc. computer software system *Mathematica*®, Version 2.2.

We thank Brian Bowditch, Yair Minsky and the referee for various interesting, helpful comments. This work was carried out while the first-named author was visiting the Centre de Recerca Matemàtica of the Institut d'Estudis Catalans, and he is grateful for the hospitality he received.

## Appendix. The Moore Decomposition Theorem.

In 1919, R. L. Moore [18] characterized the Euclidean plane topologically. Then, in 1925, Moore [19] noted that his axioms were also satisfied by a large class of quotient spaces of the plane, so that those identification spaces were also planes.

Since Moore's articles are somewhat inaccessible to today's readers because of evolving terminology and background, we will give a fairly straightforward statement of his theorem, and outline a proof. We will then show how to apply the theorem in a slightly more general situation than that considered in Section 3.

**A.1 The Moore Decomposition Theorem.** *Suppose that $f\colon \mathbb{S}^2 \to X$ is a continuous map from the two-sphere $\mathbb{S}^2$ onto a Hausdorff space $X$ such that, for each $x \in X$, the subspace $\mathbb{S}^2 - f^{-1}(x)$ is homeomorphic with the plane $\mathbb{R}^2$. Then $X$ is a two-sphere.*

**A.2 Remarks.** (i). The requirement that $\mathbb{S}^2 - f^{-1}(x)$ be homeomorphic with $\mathbb{R}^2$ is equivalent to the requirement that both $f^{-1}(x)$ and $\mathbb{S}^2 - f^{-1}(x)$ be nonempty and connected.

(ii). We shall be interested in the case where $X$ is a quotient space of $\mathbb{S}^2$ obtained by partitioning $\mathbb{S}^2$ into subspaces all of which are closed trees, and then collapsing the trees to points. The theorem implies that here $X$ is a two-sphere if and only if $X$ is Hausdorff.

(iii). Theorem A.1 has the following generalization to higher dimensions: *Suppose that $f\colon \mathbb{S}^n \to X$ is a continuous map from the $n$-sphere $\mathbb{S}^n$ onto a Hausdorff space $X$ such that, for each $x \in X$, the subspace $\mathbb{S}^n - f^{-1}(x)$ is homeomorphic with the Euclidean space $\mathbb{R}^n$. Then $X$ is an $n$-sphere provided that, in addition, $n \geq 5$, and $X$ satisfies the condition that every continuous map from the closed (two-dimensional) disk $\overline{\mathbb{D}}$ into $X$ can be approximated by an embedding.* This generalization was conjectured, and proved in many special cases, by Cannon [5], and then proved in general by R. D. Edwards; see Daverman's book [9]. The situation in dimensions 3 and 4 has not been completely resolved. □

The proof we shall give relies on a more intuitive theorem, called the Zippin Characterization Theorem; see, for example, [23, p.88]. Recall that a *Peano continuum* is a metrizable space which is a continuous image of $[0, 1]$; it is *degenerate* if it has only one point.

**A.3 The Zippin Characterization Theorem.** *A space $X$ is a two-sphere if the following four conditions are satisfied:*

(a). *$X$ is a nondegenerate Peano continuum.*
(b). *No point in $X$ separates $X$ (so that, in particular, $X$ contains at least one simple closed curve).*
(c). *Each simple closed curve in $X$ separates $X$.*
(d). *No arc in $X$ separates $X$.* □

*Proof of Theorem A.1 using Theorem A.3.* We assume the hypotheses of the Moore Decomposition Theorem and verify the four conditions of the Zippin Characterization Theorem in turn. Note that conditions (c) and (d) are true in the two-sphere by standard homological arguments; we shall use those same arguments here.

(a). Since $X$ is Hausdorff, the map $f$ is a closed surjection; hence it is easy to verify the conditions of the Urysohn metrization theorem, so $X$ is metrizable; see [20, Theorem 34.1]. Since $\mathbb{S}^2$ is a Peano continuum, so also is $X$. For each $x \in X$, both $f^{-1}(x)$ and $\mathbb{S}^2 - f^{-1}(x)$ are nonempty, so $X$ has more than one point.

(b). By hypothesis, if $x \in X$ then $\mathbb{S}^2 - f^{-1}(x)$ is connected. Hence $X - \{x\} = f(\mathbb{S}^2 - f^{-1}(x))$ is also connected.

(c). Let $J$ be a simple closed curve in $X$. Let $p_1, p_2 \in J$ cut $J$ into two arcs $A_1$ and $A_2$. Then $f^{-1}(A_1)$ and $f^{-1}(A_2)$ are compact, connected, and have nonconnected intersection $f^{-1}(p_1) \cup f^{-1}(p_2)$. Consider the segment

$$H_1(U) \oplus H_1(V) \to H_1(U \cup V) \to \tilde{H}_0(U \cap V)$$

of the Mayer-Vietoris reduced-homology sequence in the case where $U = \mathbb{S}^2 - f^{-1}(A_1)$ and $V = \mathbb{S}^2 - f^{-1}(A_2)$. These are simply connected since $f^{-1}(A_1)$ and $f^{-1}(A_2)$ are connected, so $H_1(U) \oplus H_1(V) = 0$. Also, $U \cup V = \mathbb{S}^2 - (f^{-1}(p_1) \cup f^{-1}(p_2))$ is connected but not simply connected, since $f^{-1}(p_1) \cup f^{-1}(p_2)$ is not connected. Thus $H_1(U \cup V) \neq 0$, and hence $\tilde{H}_0(U \cap V) \neq 0$, so $U \cap V = \mathbb{S}^2 - f^{-1}(J)$ is not connected. Thus $f^{-1}(J)$ separates $\mathbb{S}^2$, so $J$ separates $X$.

(d). We suppose that some arc $A$ in $X$ separates $x$ and $y$ in $X$, and we shall derive a contradiction.

Choose $x' \in f^{-1}(x)$ and $y' \in f^{-1}(y)$. Notice that $f^{-1}(A)$ separates $x'$ and $y'$ in $\mathbb{S}^2$.

Consider any $p \in A$. Then $p$ separates $A$ into arcs $A_1$ and $A_2$. We claim that one of $A_1$ and $A_2$ also separates $x$ and $y$ in $X$, and it suffices to show that one of $f^{-1}(A_1)$ and $f^{-1}(A_2)$ separates $x'$ and $y'$ in $\mathbb{S}^2$. Now consider the segment

$$H_1(U \cup V) \to \tilde{H}_0(U \cap V) \to \tilde{H}_0(U) \oplus \tilde{H}_0(V)$$

of the Mayer-Vietoris reduced-homology sequence for the pair $U = \mathbb{S}^2 - f^{-1}(A_1)$ and $V = \mathbb{S}^2 - f^{-1}(A_2)$. Here $U \cup V = \mathbb{S}^2 - f^{-1}(p)$ is simply connected since $f^{-1}(p)$ is connected, so $H_1(U \cup V) = 0$. But $x' - y'$ represents a nonzero element of $\tilde{H}_0(\mathbb{S}^2 - f^{-1}(A)) = \tilde{H}_0(U \cap V)$, hence maps to a nonzero element of $\tilde{H}_0(U) \oplus \tilde{H}_0(V)$, so $x'$ and $y'$ lie in different components of either $U = \mathbb{S}^2 - f^{-1}(A_1)$ or $V = \mathbb{S}^2 - f^{-1}(A_2)$, as claimed.

By induction, one obtains arcs $A = I_0 \supset I_1 \supset \cdots$ which separate $x$ and $y$ in $X$ such that $\cap_{n=1}^{\infty} I_n$ contains a single point $q$ which does not separate $x$ from $y$. But an arc from $x$ to $y$ in the path-connected open set $X - \{q\}$ misses some $I_n$, a contradiction. We conclude that $A$ cannot separate $X$.

The proof of the Moore Decomposition Theorem is complete. □

We now describe the Moore decomposition associated with two transverse “irrational” foliations of the plane.

**A.4 Notation.** Let $T$ be a flat torus. Puncture $T$ at a point $x_0$ to obtain the punctured torus $T_0 = T - \{x_0\}$, with the induced flat conformal structure. The torus $T$ may be cellulated by two "triangles" $\Delta_1$ and $\Delta_2$, each having its three vertices identified at the single point $x_0$. The torus $T$ has fundamental group $\pi_1(T) = \mathbb{Z}^2$. The fundamental group of the punctured torus $T_0$ is a free group of rank two. The inclusion map $T_0 \subset T$ induces a homomorphism $\pi_1(T_0) \to \pi_1(T)$ which is Abelianization.

The torus has universal cover $p\colon \mathbb{R}^2 \to T$, with $p$ conformal and covering translations conformal. The punctured torus $T_0$ has universal cover $p_0\colon \mathbb{D} \to T_0$, where $\mathbb{D}$ is the open unit disk in $\mathbb{R}^2$, with $p_0$ and the covering translations being conformal. The map $p_0$ factors through $p$, so that we obtain a conformal covering map $q_0\colon \mathbb{D} \to P_0$, where $P_0$ is the plane $\mathbb{R}^2$ punctured at the points of the $\mathbb{Z}^2$-lattice $p^{-1}(x_0)$.

$$\begin{array}{ccccc} \mathbb{D} & \xrightarrow{q_0} & P_0 & \subset & \mathbb{R}^2 \\ p_0 \downarrow & & \downarrow p|P_0 & & \downarrow p \\ T_0 & \equiv & T_0 & \subset & T \end{array}$$

Since $T_0$ is punctured at $x_0$, $T_0$ is a finite-area torus, and we may assume that the triangles $\Delta_1$ and $\Delta_2$ lift to ideal geodesic triangles in $\mathbb{D}$ which tile $\mathbb{D}$. Their images in $\mathbb{R}^2$ are topological disks which we do not know to have straight sides. For our purposes, however, we may straighten the sides equivariantly so that the images are Euclidean triangles with the vertices removed.

Any ideal triangle which is a lift of $\Delta_1$ or $\Delta_2$ has three accumulation points in the circle $\mathbb{S}^1 = \partial\mathbb{D}$ at infinity, and these are called the *cusps* of the triangle. They will be treated as lifts of $x_0$. The point $x_0$, and its lifts in $\mathbb{R}^2$, and the punctures of $P_0$ and $T_0$ will all be called *cusps* also. □

**A.5 Discussion.** Each point $x$ of $\mathbb{S}^1$ may be described as the intersection of a sequence of small disks in $\mathbb{S}^1 \cup \mathbb{D}$ each of which has a closed boundary arc in $\mathbb{S}^1$ and an open boundary arc in $\mathbb{D}$. We call such a family of open arcs *cuts in* $\mathbb{D}$ *defining* $x$.

Since the cusps in $\mathbb{S}^1$ are dense, we may define any point $x$ of $\mathbb{S}^1$ by cuts $C_i$ in $\mathbb{D}$ which join cusps.

Unless $x$ is itself a cusp, we may take the $C_i$ to be edges of the lifts of the ideal triangles $\Delta_1$ and $\Delta_2$. If $x$ is a cusp, then we may assume that the $C_i$ join cusps of ideal triangles which have $x$ as a cusp.

Thus the projections of the $C_i$ to $\mathbb{R}^2$ have a simple form: they are either the (straightened) edges of Euclidean triangles in our tessellation, or they are paths which start at one cusp, circle a second cusp (which is adjacent to the first cusp) a finite number of times, then end at a third cusp adjacent to the second. □

**A.6 Notation.** Let $\mathcal{F}_1$ and $\mathcal{F}_2$ denote distinct foliations of $\mathbb{R}^2$ by parallel lines of (distinct) slopes chosen so that no foliation line, or leaf, meets two cusps, that is, points of the given $\mathbb{Z}^2$-lattice. These foliations should behave like foliations with irrational slope, in the sense that each leaf comes arbitrarily close to the $\mathbb{Z}^2$-lattice at each end.

These foliations, when restricted to $P_0$, lift to foliations $\tilde{\mathcal{F}}_1$ and $\tilde{\mathcal{F}}_2$ of $\mathbb{D}$. □

We mention as a matter of interest the fact that the foliations $\tilde{\mathcal{F}}_1$ and $\tilde{\mathcal{F}}_2$ cannot possibly consist of hyperbolic geodesics. The reader will be able to construct a proof of this fact once he or she understands the "spiders" constructed below.

**A.7 Lemma.** *With Notation A.6, if $\tilde{\ell}$ is a leaf of $\tilde{\mathcal{F}}_i$, then the ends of $\tilde{\ell}$ converge to distinct points of $\mathbb{S}^1$.*

Hence we may compactify $\tilde{\ell}$ by adding two points at infinity. We denote this compactification by $cl(\tilde{\ell})$, and again call it a *leaf*.

*Proof of Lemma A.7.* Let $\ell$ be the projection of $\tilde{\ell}$ to $P_0$. Then $\ell$ is either a ray with one end at a cusp and the second end at infinity, or $\ell$ is a line (infinite in both directions). For an end which converges to a cusp, one can easily construct paths which start at one cusp adjacent to the end cusp of $\ell$, circle the end cusp of $\ell$ a finite number of times, then end at a third cusp adjacent to the end cusp of $\ell$. These lift to cuts in $\mathbb{D}$ which define the end as a point at infinity. For an end of $\ell$ which converges to infinity, the triangle edges which cross $\ell$ near that end lift to cuts in $\mathbb{D}$ which define a single point at infinity. □

**A.8 Lemma.** *With Notation A.6, no leaf $cl(\tilde{\ell})$ has two cusps as endpoints.*

*Proof.* The leaf $\ell$ does not meet two cusps. □

**A.9 Lemma.** *If two distinct leaves $cl(\tilde{\ell})$ and $cl(\tilde{m})$ have common endpoints, then the intersection of $cl(\tilde{\ell})$ and $cl(\tilde{m})$ consists of a single cusp.*

*Proof.* The common endpoint must project to a cusp, for all other ends can easily be separated by cross cuts. But no leaf has two cusps as endpoints. Therefore $cl(\tilde{\ell}) \cap cl(\tilde{m})$ can contain only that one cusp. □

**A.10 Notation.** We consider the upper open hemisphere $\mathbb{D}_1$ of $\mathbb{S}^2$ to be a copy of the hyperbolic unit disk $\mathbb{D}$. We consider the lower open hemisphere $\mathbb{D}_2$ of $\mathbb{S}^2$ to be another copy of the hyperbolic unit disk $\mathbb{D}$. These two copies have common boundary $\mathbb{S}^1 = \partial(\mathbb{D})$.

Applying Notation A.6, we put one copy of $\tilde{\mathcal{F}}_1$ in the upper hemisphere $\mathbb{D}_1$ of $\mathbb{S}^2$, one copy of $\tilde{\mathcal{F}}_2$ in the lower hemisphere $\mathbb{D}_2$. We define two points of $\mathbb{S}^2$ to be equivalent if they lie in the closure of the same leaf of $\tilde{\mathcal{F}}_1$ or $\tilde{\mathcal{F}}_2$. Taking the reflexive, transitive closure of this relation we get an equivalence relation $\approx$. We shall see later that transitivity is achieved after only two steps. Let $\mathbb{S}^2/\!\approx$ denote the set of equivalences classes, let $X$ denote this set endowed with the quotient topology, and let $\pi\colon \mathbb{S}^2 \to X = \mathbb{S}^2/\!\approx$ be the identification map. □

We shall show that $\pi\colon \mathbb{S}^2 \to X$ satisfies the conditions of the Moore Decomposition Theorem.

**A.11 Lemma.** *With Notation A.10, each equivalence class $g \in \mathbb{S}^2/\!\approx$ is a compact, connected, proper subset of $\mathbb{S}^2$ that does not separate $\mathbb{S}^2$, and consequently, its complement is homeomorphic with $\mathbb{R}^2$.*

*Proof.* By Lemmas A.8 and A.9, if $g$ contains more than one point, it is of one of two types. The first is the closure $cl(\tilde{\ell})$ of a leaf whose endpoints are not cusps; such an element is an arc and does not separate $\mathbb{S}^2$. The second is the union of countably many arcs meeting at a cusp, half of them from $\tilde{\mathcal{F}}_1$ and lying in the upper hemisphere $cl(\mathbb{D}_1)$ of $\mathbb{S}^2$, the other half lying in the lower hemisphere $cl(\mathbb{D}_2)$ of $\mathbb{S}^2$ and coming from $\tilde{\mathcal{F}}_2$. These arcs all project to rays, namely the two rays in $\mathcal{F}_1$ and the two rays in $\mathcal{F}_2$ emanating from a single cusp, that is a point of the $\mathbb{Z}^2$-lattice in $\mathbb{R}^2$. To show that such a $g$ is compact, it suffices to prove that the arcs exiting from a cusp $v$ form a null sequence as measured in the Euclidean metric.

There are exactly six triangles $T_1, T_2, \dots, T_6$ in $P_0$ containing a given cusp. The lifts of these six triangles at a point $v$ of $\mathbb{S}^1 = \partial(\mathbb{D})$ above this cusp form an infinite cycle at $v$ which consists of lifts of the six triangles repeated cyclically, infinitely often.

If the four rays of $\mathcal{F}_1$ and $\mathcal{F}_2$ beginning at $q_0(v) \in \mathbb{R}^2$ exit into, say $T_i$ and $T_{i+3}$, $T_j$ and $T_{j+3}$, then the corresponding leaves of $\tilde{\mathcal{F}}_1$ and $\tilde{\mathcal{F}}_2$ are trapped in successively smaller regions cut off by the lifts of those triangles. Hence these leaves have Euclidean diameter going to 0. Hence $g$ is compact.

Obviously $g$ does not separate $\mathbb{S}^2$. It has the appearance of a spider with infinitely many legs. □

**A.12 Lemma.** *With Notation A.10, $X = \mathbb{S}^2/{\approx}$ is Hausdorff.*

*Proof.* Let $g_1, g_2, \dots$ be a sequence of elements of $\mathbb{S}^2/{\approx}$ containing points $x_1, x_2, \dots \to x \in g \in \mathbb{S}^2/{\approx}$ and $y_1, y_2, \dots \to y \in h \in \mathbb{S}^2/{\approx}$. It suffices to prove that $g = h$.

Let $cl(\tilde{\ell}_1), cl(\tilde{\ell}_2), \dots$ be leaves containing $x_1, x_2, \dots$ and $cl(\tilde{m}_1), cl(\tilde{m}_2), \dots$ be leaves containing $y_1, y_2, \dots$. Then $cl(\tilde{\ell}_i)$ and $cl(\tilde{m}_i)$ share a common endpoint $v_i$. If $v_i$ is not a cusp, then $\tilde{\ell}_i = \tilde{m}_i$. We may assume that $v_i \to v$. We shall show that the equivalence class of $v$ contains both $x$ and $y$.

We may assume that the leaves $cl(\tilde{\ell}_i)$ converge to a continuum $\ell$ containing $v$ and $x$, and that the leaves $cl(\tilde{m}_i)$ converge to a continuum $m$ containing $v$ and $y$; recall that a *continuum* is a compact, connected, Hausdorff space.

We shall show that, if $v \neq x$, then $\ell$ lies in an element of $\mathbb{S}^2/{\approx}$, and similarly for $m$. This will show that $x \approx v \approx y$.

Suppose first that $\ell \subseteq \mathbb{S}^1$ and $v \neq x$. Choose a cusp $v'$ between $v$ and $x$ in $\ell$ such that each triangle containing $v'$ is small. Pick a triangle $\Delta$ at $v'$ whose other cusps $a$, $b$ say, are on either side of $v'$; by the smallness of $\Delta$ they lie in $\ell$, between $v$ and $x$. Let $\Delta'$ be the other triangle having $a$, $v'$ as cusps, and let $\Delta''$ be the other triangle having $b$, $v'$ as cusps. Then, for $i$ large, $cl(\tilde{\ell}_i)$ must intersect $\Delta'$, $\Delta$, and $\Delta''$. Hence $cl(\tilde{\ell}_i)$ cannot be close to $x$ and $v$, a contradiction.

Therefore we may assume $\ell$ contains a point $z \in \mathbb{D}_1$ or $\mathbb{D}_2$, say $\mathbb{D}_1$. Then the $\ell_i$'s must eventually be leaves of $\tilde{\mathcal{F}}_1$ that contain points converging to $z$. It follows that the $\ell_i$'s converge to the lift of a single line in $\tilde{\mathcal{F}}_1$ which may be the lift of one line missing $\mathbb{Z}^2$ or the lift of two collinear rays from a single element of $\mathbb{Z}^2$. Since those two lifted rays lie in the same element of $\mathbb{S}^2/{\approx}$, we see that $\ell$ lies in an element of $\mathbb{S}^2/{\approx}$, so $v \approx x$. Similarly, $v \approx y$, and the desired result follows. □

**A.13 Theorem.** *With Notation A2.7, the quotient map $\pi\colon \mathbb{S}^2 \to X = \mathbb{S}^2/{\approx}$ satisfies the conditions of the Moore Decomposition Theorem A.1, so $X$ is a two-sphere.* □

## References

1. R. C. Alperin, Warren Dicks and J. Porti, *The boundary of the Gieseking tree in hyperbolic three-space*, Topology Appl. **93** (1999), 219–259.
2. S. Betley, J. H. Przytycki and T. Żukowski, *Hyperbolic structures on Dehn filling of some punctured-torus bundles over $S^1$*, Kobe J. Math. **3** (1986), 117–147.
3. Brian H. Bowditch, *The Cannon-Thurston map for punctured-surface groups*, preprint (2002).
4. G. Burde and H. Zieschang, *Knots*, deGruyter Studies in Mathematics Vol. 5, Walter de Gruyter, New York, 1985.

5. J. W. Cannon, *The recognition problem: what is a topological manifold?*, Bull. Amer. Math. Soc. **84** (1978), 832–866.
6. J. W. Cannon, *The Peano curve defined by a punctured-torus bundle*, (in preparation).
7. J. W. Cannon and W. P. Thurston, *Group invariant Peano curves*, preprint (1989).
8. D. Cooper, *Automorphisms of free groups have finitely generated fixed point sets*, J. Algebra **111** (1987), 453–456.
9. Robert J. Daverman, *Decompositions of Manifolds*, Pure and Applied Mathematics Series, V. 124, Academic Press, London, 1986.
10. Warren Dicks and J. Porti, *On the Hausdorff dimension of the Gieseking fractal*, Topology Appl. (to appear).
11. Kenneth Falconer, *Fractal geometry. Mathematical foundations and applications*, John Wiley and Sons, New York, 1990.
12. W. Floyd and A. Hatcher, *Incompressible surfaces in punctured-torus bundles*, Topology Appl. **13** (1982), 263–282.
13. D. Gaboriau, A. Jaeger, G. Levitt and M. Lustig, *An index for counting fixed points of automorphisms of free groups*, Duke Math. J. **93** (1998), 425–452.
14. T. Jørgensen, *On pairs of punctured tori*, preprint.
15. A. Martino, *An index theorem for automorphisms of free products*, J. Group Theory **2** (1999), 199–211.
16. Curtis T. McMullen, *Local connectivity, Kleinian groups and geodesics on the blowup of the torus*, Invent. Math. **146** (2001), 35–91.
17. Yair N. Minsky, *The classification of punctured-torus groups*, Ann. of Math. (2) **149** (1999), 559–626.
18. R. L. Moore, *Concerning a set of postulates for plane analysis situs*, Trans. Amer. Math. Soc. **20** (1919), 169–178.
19. R. L. Moore, *Concerning upper semi-continuous collections of continua*, Trans. Amer. Math. Soc. **27** (1925), 416–428.
20. James R. Munkres, *Topology, Second Edition*, Prentice Hall, Upper Saddle River, N. J., 2000.
21. J. Nielsen, *Die Isomorphismen der allgemeinen, unendlichen Gruppe mit zwei Erzeugenden*, Math. Ann. **78** (1917), 385–397.
22. Wan Lin, *Presentations of the automorphism group of free nilpotent groups*, Comm. Alg. **24** (1996), 3403–3411.
23. Raymond Louis Wilder, *Topology of Manifolds*, American Mathematical Society Colloquium Publications, Volume 32, Providence, R.I., 1963.

James W. Cannon,
Department of Mathematics,
Brigham Young University,
Provo UT 84602, U.S.A.
*E-mail address*: cannon@math.byu.edu

Warren Dicks,
Departament de Matemàtiques,
Universitat Autònoma de Barcelona,
E-08193 Bellaterra (Barcelona), Spain
*E-mail address*: dicks@mat.uab.es