REALIZATION OF THE MAPPING CLASS GROUP BY HOMEOMORPHISMS

VLADIMIR MARKOVIC

Abstract. In this paper, we show that the mapping class group of a closed surface can not be geometrically realized as a group of homeomorphisms of that surface. More precisely, let $Pr : \mathcal{H}omeo(M) \to \mathcal{MC}(M)$ denote the standard projection of the group of homeomorphisms to the mapping class group of a closed surface M of genus $\mathbf{g} > 5$. We show that there is no homomorphism $\mathcal{E} : \mathcal{MC}(M) \to \mathcal{H}omeo(M)$, such that $Pr \circ \mathcal{E}$ is the identity. This answers a question by Thurston (see [10]).

1. Introduction

1.1. **The main results.** Let M be a closed surface of genus $\mathbf{g}$, with $\mathbf{g} \geq 2$. The mapping class group $\mathcal{MC}(M)$ is the group of homotopy classes of orientation preserving homeomorphisms of M (all homeomorphisms in this paper are assumed to be orientation preserving). By $\mathcal{H}omeo(M)$ we denote the group of homeomorphisms of M. If $\widetilde{f} \in \mathcal{H}omeo(M)$ is a homeomorphism, then $[\widetilde{f}] = f \in \mathcal{MC}(M)$ denotes the corresponding homotopy class. The group $\mathcal{MC}(M)$ can be seen as the quotient of the group $\mathcal{H}omeo(M)$ by the subgroup of homeomorphisms of M that are isotopic to the identity (recall that two homeomorphisms from $\mathcal{H}omeo(M)$ are isotopic if an only if they are homotopic to each other). There is the associated projection

$$Pr : \mathcal{H}omeo(M) \to \mathcal{MC}(M).$$

The main result of this paper is the following theorem.

Theorem 1.1. *Let M be a closed surface of genus $\mathbf{g} > 5$. Then, there is no homomorphism $\mathcal{E} : \mathcal{MC}(M) \to \mathcal{H}omeo(M)$, such that the composite homomorphism $Pr \circ \mathcal{E}$ is the identity.*

Let $G < \mathcal{MC}(M)$ be a finite group. The question of whether the corresponding homomorphism $\mathcal{E} : G \to \mathcal{H}omeo(M)$ exists, was known as the Nielsen realization problem. Kerckhoff [9] solved this famous problem by showing the existence of $\mathcal{E}$ in this case (an alternative proof of this theorem is a corollary of [6]). There are other subgroups of $\mathcal{MC}(M)$ for which the corresponding homomorphism exists (one example are Abelian subgroups of $\mathcal{MC}(M)$). But generally, there is no criteria which would help decide whether a given subgroup of $\mathcal{MC}(M)$ can be realized by homeomorphisms. We believe that the methods developed in this paper may help obtain further results in this direction.

The version of Theorem 1.1 was proved by Morita [14], [15], where the space of homeomorphisms of M is replaced by diffeomorphisms of M, and where the genus of M is at least 5. The projection $Pr : \mathcal{H}omeo(M) \to \mathcal{MC}(M)$ induces the

2000 *Mathematics Subject Classification.* Primary 20H10, 37F30.

map between the corresponding cohomology groups. In the diffeomorphic case, it was shown that not all respective cohomology groups of the groups $Diff(M)$ and $\mathcal{MC}(M)$ are the same. This shows that there can not exist a homomorphism like $\mathcal{E}$. However, it follows from the work of Mather and Thurston [18], [12], [14],[15], that there is no cohomological obstructions to the existence of $\mathcal{E}$ in the general homeomorphisms case. It has been conjectured [10] (also see [2], [14], [15]), that such $\mathcal{E}$ does not exists in the general case.

Remark. The mapping class group of the torus can be represented by homeomorphisms (that is, the corresponding homomorphism $\mathcal{E}$ exists). In fact, it can be represented as the group of affine transformations $\mathbf{SL}_2(\mathbf{Z})$. The proof that the corresponding $\mathcal{E}$ does not exist, works for surfaces of genus $\mathbf{g} > 5$. However, most of the analysis in this paper is valid for every surface of genus $\mathbf{g} \geq 2$. The only place we use that $\mathbf{g} > 5$ is at the very end of the proof, at the end of Section 5. It seems likely that one can alter the proof to extend Theorem 1.1 to the surfaces of genus 3, 4 and 5. The case $\mathbf{g} = 2$ seems to require a different type of analysis (although the results from sections 2,3 are still applicable).

Let $\alpha \subset M$ be a simple closed curve. By $t_\alpha \in \mathcal{MC}(M)$ we denote the standard twist about α. By $[\alpha]$ we denote the corresponding homotopy class of curves on M. Clearly, if $\alpha_1 \in [\alpha]$ then $t_{\alpha_1} = t_\alpha$. Now, let α and β be two simple closed curves such that neither of them separates M (this means that $M \setminus \alpha$ and $M \setminus \beta$ are connected sets). Assume that the intersection number between the homotopy classes of α and β is one. Then we have the following standard relation in the mapping class groups (see the Ivanov's book [19] for background on the mapping class groups and more), called the Artin's relation

$$t_\alpha \circ t_\beta \circ t_\alpha = t_\beta \circ t_\alpha \circ t_\beta. \tag{1.1}$$

Set $f = t_\beta \circ t_\alpha$. Then f conjugates t_β to t_α in $\mathcal{MC}(M)$. We have

$$\mathcal{E}(t_\alpha) \circ \mathcal{E}(t_\beta) \circ \mathcal{E}(t_\alpha) = \mathcal{E}(t_\beta) \circ \mathcal{E}(t_\alpha) \circ \mathcal{E}(t_\beta). \tag{1.2}$$

Another useful type of relation in $\mathcal{MC}(M)$ is the following. Let $\widetilde{f}, \widetilde{g} : M \to M$ be two homeomorphisms that have disjoint supports (a closed set is said to be support of $\widetilde{f}$, if $\widetilde{f}$ is the identity outside that set). In particular, $\widetilde{f} \circ \widetilde{g} = \widetilde{g} \circ \widetilde{f}$, so $\widetilde{f}$ and $\widetilde{g}$ commute. We have

$$\mathcal{E}([\widetilde{f}]) \circ \mathcal{E}([\widetilde{g}]) = \mathcal{E}([\widetilde{g}]) \circ \mathcal{E}([\widetilde{f}]).$$

The following is an old theorem (going back to Hurwitz). The standard way to prove it (see [11]) is by using the Ahlfors trick (see [1]).

Proposition 1.1. *Let a homeomorphism $\widetilde{f} : M \to M$ be periodic, that is, $\widetilde{f}^k = id$, for some $k \geq 1$. Then, there exists a complex structure on M, and a homeomorphisms $\widetilde{g} : M \to M$, such that $\widetilde{g}^{-1} \circ \widetilde{f} \circ \widetilde{g}$ is a conformal automorphism with respect to this complex structure. Moreover, if a periodic homeomorphism $\widetilde{f}$ is homotopic to a conformal automorphism $e : M \to M$ (with respect to some complex structure), then we can choose $\widetilde{g}$ to be homotopic to the identity homeomorphism of M, and $\widetilde{g}^{-1} \circ \widetilde{f} \circ \widetilde{g} = e$.*

Let $e \in \mathcal{MC}(M)$ be a periodic element. Assume that for some complex structure on M there exists a conformal automorphism $\widetilde{e}$ of M such that $[\widetilde{e}] = e$. Then $\mathcal{E}(e)$ is conjugated to a conformal automorphism, by a homeomorphism $\widetilde{g} : M \to M$ which is homotopic to the identity. We can conjugate the whole homomorphism $\mathcal{E}$ by $\widetilde{g}$, to obtain a new homomorphism $\mathcal{E}' = \widetilde{g}^{-1} \circ \mathcal{E} \circ \widetilde{g}$, for which we have that $\mathcal{E}'(e)$ is a conformal automorphism of M (when M is endowed with the appropriate complex structure).

In the end, we prove Theorem 1.1 by employing Artin's relations to derive a contradiction to the existence of the homomorphism $\mathcal{E}$. However, we need to do a bit of analysis first. The analysis that precedes the proof of Theorem 1.1 is intended for proving that every twist $\mathcal{E}(t_\alpha)$ is semi-conjugated to the identity mapping outside some closed annulus that is homotopic to α (in fact, this annulus is compactly contained in a bigger open annulus homotopic to the same curve). One important step in establishing this is based on the analysis of Anosov like elements of $\mathcal{MC}(M)$ (see Section 4). These are elements of $\mathcal{MC}(M)$ that are Anosov (or in general, pseudo-Anosov) on some subsurface of M. The methods we deploy have similarity with the theory of global shadowing by Anosov (and in general pseudo-Anosov) homeomorphisms (see the papers by Franks [5], and Handel [8], where some of these classical results were proved, and for further references). However, there are important differences. The main difference is that in our analysis we are focused on analyzing homeomorphisms that commute with these Anosov-like maps. In any case, the methods we give can be used to give perhaps more direct proof of the theorem which states that a homeomorphism that is homotopic to an Anosov map of the torus (or the punctured torus), is in fact semi-conjugated to that map. It can be generalized to pseudo-Anosov maps as well (in this case the corresponding statement is somewhat different, because a homeomorphism that is homotopic to a pseudo-Anosov map is not necessarily semi-conjugated to that map).

In Section 2 and Section 3, we define the concept of an admissible decomposition $\mathbf{S}$ for a group $G < \mathcal{H}omeo(M)$ of homeomorphism of M. The decomposition $\mathbf{S}$ of M is an upper semi-continuous decomposition (we recall this definition below), where every element of $\mathbf{S}$ is setwise preserved by every homeomorphism in G. Also, we require that acyclic components are in certain sense dense among all elements of $\mathbf{S}$. If G is the cyclic group generated by $\widehat{f} \in \mathcal{H}omeo(M)$, we say that $\mathbf{S}$ is an admissible decomposition for $\widehat{f}$. The main result we prove in Section 3 is Theorem 3.1 which says that every such group G has the minimal decomposition. The minimal decomposition is the unique admissible decomposition for G that is contained in any other admissible decomposition for G. The main step in proving this is Proposition 3.1 where we prove that the descendant of any finite set of admissible decompositions for G is also an admissible decomposition for G.

For a generic homeomorphism of M the minimal decomposition will contain only one component, namely M itself. However, when homeomorphisms satisfy certain algebraic relations (as certain homeomorphism from $\mathcal{E}(\mathcal{MC}(M))$ do), then the corresponding minimal decomposition is non-trivial. The purpose of Section 4 is to show that for certain homeomorphism in $\mathcal{E}(\mathcal{MC}(M)) < \mathcal{H}omeo(M)$, there are non-trivial admissible decompositions for these homeomorphisms. The key results in Section 4 are Proposition 4.1 and Theorem 4.1.

The main results we prove in this paper do not require M to have any structure other than being a closed, orientable, topological surface. However, when analyzing closed sets in M, and what is an important assumption for us, connected subsets of M, it is convenient to assume that M is a Riemann surface. We will make this assumption often, but the main reason for this is to be able to use the Riemann mapping theorem (and its corollaries) to "parametrize" certain closed and connected subsets of M by geometric circles. Another reason for introducing Riemann surfaces is to use Proposition 1.1.

There are some interesting results we prove in this paper that are independent of the assumption that such a homomorphism $\mathcal{E} : \mathcal{MC}(M) \to \mathcal{H}omeo(M)$ exists. These results are Proposition 3.1, Theorem 3.1, Proposition 4.1. In fact, all the results we prove before the beginning of Subsection 4.4 are independent of the existence of $\mathcal{E}$. However, at the beginning of Subsection 4.4 we make the assumption that such $\mathcal{E}$ exists, and from there on we work under that assumption.

In recent years, I have been involved in a fruitful collaboration with David Epstein on several papers. One of these papers [4], deals with the quasiconformal version of Theorem 1.1, where instead of a closed surface one considers the unit disc, and where instead of homeomorphisms one restricts to the group of quasiconformal maps. I would like to express my gratitude to David for all the help and encouragement he gave me. I am grateful to the referee for numerous suggestions that have improved this paper. Also, I am thankful to J. Giblin and D. Saric for reading this paper and for making useful remarks and suggestions.

2. Upper semi-continuous decomposition of a surface

Let M be a surface. If M is equipped with a complex structure we say that M is a Riemann surface. Recall the following definition.

Definition 2.1. *Let $\mathbf{S}$ be a collection of closed, connected subsets of M. We say that $\mathbf{S}$ is an upper semi-continuous decomposition of M if the following holds:*

1. *If $S_1, S_2 \in \mathbf{S}$, then $S_1 \cap S_2 = \emptyset$.*
2. *If $S \in \mathbf{S}$, then S does not separate M, that is $M \setminus S$ is a connected set.*
3. *We have*
$$M = \bigcup_{S \in \mathbf{S}} S.$$
4. *If $S_n \in \mathbf{S}$, $n \in \mathbf{N}$, is a sequence that has the Hausdorff limit S_0, then there exists $S \in \mathbf{S}$ such that $S_0 \subset S$.*

It is well known that if $\mathbf{S}$ is an upper semi-continuous decomposition then every $S \in \mathbf{S}$ has an arbitrarily small open neighborhood which contains every element in $\mathbf{S}$ that it intersects (see [3]). We will prove this fact late.

Recall that a set $S \subset M$ is essential if its fundamental group injects into the fundamental group of M. If M is a hyperbolic Riemann surface (can be equipped with the complete hyperbolic metric), the definition of essential subsets is equivalent to the following one. Let $\widetilde{S} \subset \mathbf{D}$ denote all the points in the unit disc that are mapped into S under the standard covering map $\mathbf{D} \to M$. Let $\widetilde{S}'$ be a connected component of $\widetilde{S}$. Then we say that S is essential (or an essential subset of M) if no component of $\mathbf{D} \setminus \widetilde{S}'$ is compactly contained in $\mathbf{D}$. The following lemma is elementary.

Lemma 2.1. *Let M be a surface and let $S \subset M$ be a closed and connected set which does not separate M. Then, either $M \setminus S$ is simply connected, or S is an essential subset of M. Also, let S_1 and S_2 be two closed, connected subsets of M so that neither one of them separates M. Then every component of the intersection $S_1 \cap S_2$ is itself a closed, connected subset of M which does not separate M.*

Definition 2.2. *Let M be a surface and let $S \subset M$ be a closed, connected subset of M which does not separate M. We say that S is acyclic if there is a simply connected open set $U \subset M$ such that $S \subset U$ and $U \setminus S$ is homeomorphic to an annulus.*

Remark. Note that this definition agrees with the standard definition of acyclic closed subsets of manifolds. Assume that M is a hyperbolic Riemann surface. Let $S \subset M$ be a closed, connected set that does not separate M. Then S is acyclic if and only if there is a lift of S to the unit disc $\mathbf{D}$ (the universal cover of M) which is a compact subset of $\mathbf{D}$.

Let M be a closed surface of genus $g \geq 2$. In what follows, by a curve we mean a simple closed curve. We recall a few standard definitions. Let $M_1 \subset M$ be a subsurface. This means that M_1 is an open, connected subset of M, that is M_1 is a surface itself. One can define the notion of ends in M_1 (see [1]). To each end of M_1, there corresponds a nested sequence of curves $\beta_n \subset M_1$, $n \in \mathbf{N}$, (mutually homotopic in M_1) that leave every compact subset of M_1. This sequence defines an end of M_1. Since $M_1 \subset M$, we have that β_n are mutually homotopic in M too. We will say that the homotopy type of a given end of M_1, is the corresponding homotopy class (in M) of the curves β_n (we also say that the end is homotopic to each curve in the homotopy class of β_n). The frontier component that corresponds to a given end is the Hausdorff limit of the sequence β_n. This limit is a closed, connected subset of M, but it does not belong to M_1.

Definition 2.3. *We say that M_1 is a proper subsurface of M if one of the following two cases occurs:*

(1) *The curves β_n are not homotopically trivial in M, for any end of M_1.*
(2) *The surface M_1 is a simply connected surface.*

Remark. Let α and β be two simple closed curves on a closed Riemann surface. In this paper, the intersection number between the homopotopy classes $[\alpha]$ and $[\beta]$ denotes the minimal number of intersections that any two simple closed curves from $[\alpha]$ and $[\beta]$ respectively must have. In particular, the intersection number between $[\alpha]$ and $[\beta]$ is zero if and only if there are simple closed curves $\alpha_1 \in [\alpha]$ and $\beta_1 \in [\beta]$, so that $\alpha_1 \cap \beta_1 = \emptyset$.

Note that M_1 is a proper subsurface of M if and only if the surface M_1 is an essential subset of M. Let $\alpha, \beta \subset M$ be simple closed curves. Assume that the classes $[\alpha]$ and $[\beta]$ are the homotopy types of two different ends of M_1. Then the intersection number between $[\alpha]$ and $[\beta]$ is zero. This includes the possibility $[\alpha] = [\beta]$. Therefore, if M_1 is a proper subsurface of M (recall that we assume that M is closed), then M_1 has at most finitely many ends (the upper bound depends only on the genus $\mathbf{g}$ of M). Therefore, M_1 is the interior of a compact surface with boundary, or as it is sometimes called a bordered surface.

Let $\mathbf{S}$ be an upper semi-continuous decomposition of M. For every point $p \in M$, there exists a unique $S_p \in \mathbf{S}$ that contains p. The set of all points $p \in M$ such that the corresponding S_p is acyclic is denoted by $M_{\mathbf{S}}$.

Proposition 2.1. *Let $\mathbf{S}$ be a proper upper semi-continuous decomposition of a closed surface M of genus at least two. Then every component of $M_{\mathbf{S}}$ is a proper subsurface of M. In particular, every such component represents the interior of a compact subsurface of M with finitely many ends.*

Proof. Let $p_n \in M$ be a sequence of points such that the corresponding sets $S_{p_n} \in \mathbf{S}$ are not acyclic. Let $p \in M$ be an accumulation point for the sequence p_n, and $S_p \in \mathbf{S}$ the corresponding component. Then, by passing to a subsequence we may assume that S_{p_n} converges in the Hausdorff topology. This limit is a closed and connected subset of M, but it is not acyclic (it is clear that this set does not have any lift to $\mathbf{D}$, which is a compact subset of $\mathbf{D}$). Therefore, S_p is not acyclic. This shows that $M_{\mathbf{S}}$ is open in M.

Let M_1 be a component of $M_{\mathbf{S}}$. Choose an end of M_1, and let β_n be the corresponding nested sequence. If β_n is homotopically trivial in M, then β_n bounds a topological disc in M. Fix β_n, and denote this disc by D. If $D \subset M_1$ then M_1 is simply connected. Therefore, we suppose that D is not entirely contained in M_1.

Let $p \in D$. Assume that $S_p \in \mathbf{S}$ is not acyclic. But, such S_p can not be contained in a topological disc D. This means that S_p must intersect $\beta_n \subset M_1$, which is a contradiction. Therefore, every component of $M_{\mathbf{S}}$ is a proper subsurface of M. $\square$

The following theorem is a classical result known as the Moore Theorem (see [3]).

Theorem 2.1. *Let M be a surface and let $\mathbf{S}$ be an upper semi-continuous decomposition of M so that every component in $\mathbf{S}$ is acyclic. Then there is a continuous map $\phi : M \to M$ that is homotopic to the identity map on M and such that for every $q \in N$ we have $\phi^{-1}(q) \in \mathbf{S}$. Moreover, for different points q_1 and q_2, the corresponding pre-images are different.*

Originally Moore had proved this when M is the plane $\mathbf{R}^2$, but the general case easily follows from this by passing to the universal cover of M.

The following two lemmas will be needed later.

Lemma 2.2. *Let P be a compact subset of a surface M, so that every component of P is acyclic. Let U_n, $n \in \mathbf{N}$, be a decreasing family of open sets, $U_{n+1} \subset U_n$, and assume that*

$$\bigcap_{n \in \mathbf{N}} U_n = P.$$

If U_n does not separate M for each $n \in \mathbf{N}$, then for n large enough, every component of U_n is simply connected.

Proof. Since P is a compact set, whose every component is acyclic, we can find a simply connected set $D \subset M$, so that $P \subset D$. Then, from that fact that U_n is decreasing sequence of sets, for n large enough we have that the closure of U_n is contained in D. Therefore, for n large enough, every component U'_n of U_n is a subsurface of D, and it is either simply connected or it separates D. But if U'_n separates D, then since the closure of U'_n is contained in D, we conclude that U'_n would separate M as well, which is impossible. So, U'_n is simply connected.

□

Also, we will need the following lemma.

Lemma 2.3. *Let M be a closed surface of genus $g \geq 2$ and let $F : M \to M$ be a homeomorphism. Suppose that $\mathbf{S}$ be an upper semi-continuous decomposition of a surface M so that every element of $\mathbf{S}$ is acyclic, and so that if $S \in \mathbf{S}$ then $F(S) \in \mathbf{S}$. Furthermore, assume that for each $S \in \mathbf{S}$ there exists a decreasing sequence of simply connected open sets $U(S,n) \subset M$ so that $\bigcap_{n\in\mathbf{N}} U(S,n) = S$, and so that each $U(S,n)$ contains every element of $\mathbf{S}$ it intersects. If for each $U(S,n)$ there exists $k(S,n) \in \mathbf{N}$, so that $F^{k(S,n)}(U(S,n)) = U(S,n)$, then F is homotopic to a periodic homeomorphism of M.*

Proof. We first apply the Moore Theorem to the decomposition $\mathbf{S}$. From the assumptions in the lemma that F and $U(S,n)$ are compatible with $\mathbf{S}$ we can assume from now on that $\mathbf{S}$ is the trivial decomposition of M, that is each element of $\mathbf{S}$ is a point in M.

We endow M with a complex structure and the corresponding hyperbolic metric. Let d_0 be the lower bound on the injectivity radius of M (with respect to the hyperbolic metric). Since each $S \in \mathbf{S}$ is a point, and since $\bigcap_{n\in\mathbf{N}} U(S,n) = S$, where $U(S,n)$ is a decreasing sequence, we have that for n large enough the hyperbolic diameter of every set $U(S,n)$ is smaller than $\frac{d_0}{4}$. Now, for each point $S \in \mathbf{S}$ we discard all the sets $U(S,n)$ whose hyperbolic diameter is greater than $\frac{d_0}{4}$.

The union of all open sets $U(S,n)$, $S \in \mathbf{S}$, is an open cover of M. We choose a finite sub-cover of this open cover. That is, there are sets $S_1,...,S_m \in \mathbf{S}$, and numbers $n_1,...,n_m$, so that the union of the open sets $U(S_1,n_1),...,U(S_m,n_m)$ covers M. For simplicity we set $U(S_i,n_i) = U_i$, $1 \leq i \leq m$. Let $k(i) \in \mathbf{N}$, $1 \leq i \leq m$, so that $F^{k(i)}(U_i) = U_i$. Let N be the product of all the m numbers $k(1),...,k(m)$. Then the homeomorphism $G = F^N$ setwise preserves each U_i. We show that G is homotopic to the identity. This conclusion yields that F is homotopic to a periodic homeomorphism (see [2], [7]).

Let $\widetilde{U_1}$ be a lift of U_1 to the universal cover $\mathbf{D}$ of M. Let $\widetilde{g} : \mathbf{D} \to \mathbf{D}$ be a lift of G that setwise preserves $\widetilde{U_1}$. We have that $\widetilde{U_1}$ is a simply connected, compactly contained subset of $\mathbf{D}$, of the same hyperbolic diameter as U_1. Suppose that U_2 intersects U_1. Let $\widetilde{U_2}$ be a lift of U_2 to $\mathbf{D}$ that intersects $\widetilde{U_1}$. In fact, $\widetilde{U_2}$ is the only such lift because the diameter of both U_1 and U_2 is less than $\frac{d_0}{4}$. This implies that $\widetilde{g}$ setwise preserves $\widetilde{U_2}$ as well. By repeating the same process we see that $\widetilde{g}$ setwise preserves every lift of every U_i. Since the hyperbolic diameter of each U_i is less than $\frac{d_0}{4}$ we conclude that the hyperbolic distance between $\widetilde{g}(x)$ and x is less than $\frac{d_0}{4}$ for every $x \in \mathbf{D}$. But this implies that $\widetilde{g}$ extends to the identity on the boundary of $\mathbf{D}$ which implies that G is homotopic to the identity.

□

2.1. A few lemmas on conformal maps of planar domains.

We recall the following collection of classical results (see [16], [1], [17]).

Proposition 2.2. *Let M be a planar Riemann surface (possibly with infinitely generated fundamental group), and let T be a component of the frontier of M. Assume that there exists a relatively open subset $U \subset \overline{M}$ so that T is the only frontier component of M that U intersects (such frontier components are also called "very strong" frontier components, see [17]). Then there exists a domain $\Omega \subset \mathbf{D}$ and a conformal map $\phi : M \to \Omega$, so that the frontier component of Ω that corresponds to T is the unit circle. Also, the closure $\overline{\phi(U \cap M)}$ (in $\overline{\Omega} \subset \overline{\mathbf{D}}$) contains a relatively open subset I of the unit circle. Moreover, every point in I contains a relatively open neighborhood in $\overline{\Omega}$ which is contained in $\overline{\phi(U \cap M)}$.*

Remark. In fact, on this set I one can prove that the map ϕ^{-1} has radial limits almost everywhere, and these are dense in $U \cap T$ (see [16]).

Another standard result we need is the following (see [16], [4]).

Proposition 2.3. *With the notation and the assumptions from the previous proposition we have the following. Assume that $\overline{M}$ is a subset of another (open) Riemann surface M', and let $F : M' \to M'$ be a homeomorphism, so that $F(M) = M$. Then the homeomorphism $\phi \circ F \circ \phi^{-1} : \Omega \to \Omega$ extends continuously to the set $I \subset \overline{\phi(U \cap M)}$ which is an open subset of the unit circle.*

Remark. In the above proposition one can replace the assumption that F is defined on a larger Riemann surface M' with the weaker assumption that F is a homeomorphism of $\overline{M}$.

Now we prove the following two lemmas that will be used later in this paper.

Lemma 2.4. *Let $\Omega \subset \mathbf{C}$ be an annulus, and let $\phi : \Omega \to N$ be a conformal map (here N is a geometric annulus). Let A and B be two connected and relatively closed subsets of Ω, such that the closure of either one of them connects the two frontier components of Ω. Assume that the closures of $\overline{A}$ and $\overline{B}$ are disjoint subsets of $\overline{\Omega}$. Then, the closures of the sets $\phi(A)$ and $\phi(B)$ are disjoint subsets of $\overline{N}$.*

Proof. To prove this, let U_1 and U_2 denote the two components of $\Omega \setminus (A \cup B)$, such that either one of them connects the two frontier components of Ω. Since A and B are connected, it is clear that there are exactly two such components (but there can be many components of $\Omega \setminus (A \cup B)$ that do not connect the two frontier components of Ω). Let p be a point in the boundary of Ω, that is not in the closure of either A or B, and such that p is in the closure of U_1. Then, there is a relative neighborhood of p in the corresponding frontier component of Ω, that is contained in the closure of U_1. This shows that the intersection between U_1 and either of the frontier components of Ω contains a relatively open set. Similarly for U_2.

It follows that $\phi(A)$ and $\phi(B)$ are connected, and we conclude that there is an interval In_A (In_B) that contains the intersection between the closure of A (the closure of B) and a chosen boundary circle. We assume that In_A and In_B are the minimal intervals with these properties. It is clear that the interiors of In_A and In_B do not meet. To see that the corresponding closed intervals do not meet, we use the fact that the intersection between the closure of $\phi(U_1)$ and either boundary circle contains an open interval. Similarly for $\phi(U_2)$. These open intervals in fact separate the closed intervals In_A and In_B. Similarly for the other boundary circle of N. This proves the lemma.

□

Next, we have:

Lemma 2.5. *Let $S \subset \mathbf{D}$ denote an acyclic subset of the open unit disc $\mathbf{D}$ and assume that S has no interior. Let $F : \mathbf{D} \to \mathbf{D}$ be a homeomorphism such that $F(S) = S$. Let C be an open subset of $\mathbf{D}$ so that there is a sequence $m(n) \in \mathbf{N}$, $n \in \mathbf{N}$, so that the sequence of maps $F^{m(n)}$ converges to the identity uniformly on compact sets in C when $n \to \infty$. In addition, suppose that there is a sequence of points $z_k \in (\mathbf{D} \setminus S)$, $k \in \mathbf{N}$, so that $F(z_k) = z_k$ for each k, and so that the accumulation set of the sequence z_k is a subset of S. Then F fixes every point in $C \cap S$.*

Remark. If we drop the assumption on the existence of the sequence of fixed points x_k which tends to S, then one could expect that the conclusion of the lemma would be that some fixed power of F is the identity on $C \cap S$.

Proof. The set $\mathbf{D} \setminus S$ is a conformal annulus. This means that there is a geometric annulus $N = \{z \in \mathbf{C} : r < |z| < 1\}$, $0 < r < 1$, and a conformal map $\phi : \mathbf{D} \setminus S \to N$. Assume that the the unit circle is being mapped onto the unit circle by ϕ (in this case ϕ continuously extends to a homeomorphism of the unit circle). Since S has no interior the frontier component of $\mathbf{D} \setminus S$ that is not the unit circle is equal to S. Since F is a homeomorphism of $\mathbf{D}$ we have that the map $G = \phi \circ F \circ \phi^{-1}$ is a homeomorphism of the closed annulus $\overline{N}$ (see Proposition 2.3). Let R denote the other boundary circle of N (the one of radius r). Then the accumulation set of the points $\phi(z_k)$ is contained in R. Since these points are all fixed under G we conclude that the restriction $g : R \to R$ of the map G to R has a fixed point $z_0 \in R$, $g(z_0) = z_0$.

Let $Fix(g)$ be the set of fixed points of g. This is a closed subset of R, and the restriction of g on any maximal open interval $I \subset (R \setminus Fix(g))$ is conjugated to a translation map of the real line. Although the map ϕ^{-1} does not extend continuously to R, one can identify each point in R with the corresponding prime-end in S (see [16]). We adopt the following standard convention. For each $x \in R$ we say that $\phi^{-1}(x)$ is the impression of the corresponding prime end (see [16] for various equivalent definitions of the impression of a prime-end). Note that it is possible that impressions of different prime-ends have non-empty intersection. On the other hand, since the closure of $\mathbf{D} \setminus S$ contains S, each $z \in S$ belongs to the impression of at least one prime-end (and possibly to the impressions of many different prime-ends). We say that $\phi(z) \subset R$ is the union of all the points that correspond to all the prime-ends whose impressions contain z.

We have $\phi(C) \subset \overline{N}$. Since C is an open subset of $\mathbf{D}$ we have that $\phi(C) \cap N$ is an open subset of N. Also, we have that the interior of $\phi(C) \cap R$ is dense in $\phi(C) \cap R$. This follows from the fact that C is open and the fact that the points in S that are the radial limits of ϕ^{-1} are everywhere dense in S (see [16]). Also, the corresponding points in R at which the radial limit for ϕ^{-1} exists are dense in R (and in fact have the full Lebesgue measure in R). Set $K = \phi(C) \cap R$. Then the interior $int(K)$ is an open subset of R everywhere dense in K. Moreover, for every point $x \in int(K)$ there exists a relatively open neighborhood of x in $\overline{N}$ that is contained in $\phi(C)$. This implies that since $F^{m(n)} \to id$, $n \to \infty$, uniformly on compact sets in C, we have that $G^{m(n)} \to id$, uniformly on compact sets in $int(K)$.

Suppose that there exists $x \in (int(K) \cap (R \setminus Fix(g)))$. Then there is a maximal open interval $I \subset (R \setminus Fix(g))$ so that $x \in I$. We have that $g^n(x) \to y$, $n \to \infty$, where $y \in Fix(g)$ is the fixed point of g that is the corresponding end of the interval I. But this is not possible since $g^{m(n)}(x) \to x$. We conclude that $int(K) \subset Fix(g)$ and therefore $K \subset Fix(g)$. Since $int(K)$ is open and dense in K, for a dense set of points in K there exists the radial limit of ϕ^{-1} at that point. The map F fixes each such radial limit. On the other hand, these radial limits are dense in $C \cap S$. This proves the lemma. □

3. The minimal decomposition that corresponds to a group of homeomorphism of a closed surface

Let M be a surface of genus $\mathbf{g} \geq 2$ and as usual let $\mathcal{H}omeo(M)$ denote the group of homeomorphisms of M.

Definition 3.1. *Let $\mathbf{S}$ be a upper semi-continuous decomposition of M. Let G be a subgroup of $\mathcal{H}omeo(M)$. We say that $\mathbf{S}$ is admissible for the group G if the following holds.*

1. *Each $\widehat{f} \in G$ preserves setwise every component of $\mathbf{S}$.*
2. *Let $S \in \mathbf{S}$. Then every point, in every frontier component of the surface $M \setminus S$, is a limit of points from $M \setminus S$ that belong to acyclic components of $\mathbf{S}$ (note that not every point of S need to be in a frontier component of the subsurface $M \setminus S$).*

If G is a cyclic group generated by a homeomorphism $\widehat{f} : M \to M$ we say that $\mathbf{S}$ is an admissible decomposition for $\widehat{f}$.

For a generic homeomorphism $\widehat{f} : M \to M$, the only admissible decomposition is the trivial one, which is the one that contains only one set, namely M itself.

Lemma 3.1. *Let $\mathbf{S}$ be an admissible decomposition for a homeomorphism $\widehat{f} : M \to M$, such that $M_{\mathbf{S}} = M$. Then $\widehat{f}$ is homotopic to the identity map.*

Proof. So, we assume that every component of $\mathbf{S}$ is acyclic ($M_{\mathbf{S}} = M$). By lifting $\mathbf{S}$ to the unit disc $\mathbf{D}$, we obtain an admissible decomposition that corresponds to a lift of $\widehat{f}$ to $\mathbf{D}$. This lift of $\mathbf{S}$, which we denote by $\mathbf{S}'$ is in particular an upper semi-continuous decomposition of the (open) unit disc, whose every component is a compact, acyclic subset of $\mathbf{D}$. Let $f' : \mathbf{D} \to \mathbf{D}$ be a lift of $\widehat{f}$ which setwise preserve an element S'_0 of $\mathbf{S}'$. Let $A_0 \subset \mathbf{S}'$ be the union of all elements of $\mathbf{S}'$ that are setwise preserved by f'. It is clear that f' setwise preserves every element of $\mathbf{S}'$ that is in a small enough open neighborhood of a given $S' \in A_0$. Also, every sequence of elements in A_0 converges to an element that is in A_0. Therefore, we have $A_0 = \mathbf{S}'$.

It is clear that there is an upper bound on the hyperbolic diameter of every component of $\mathbf{S}'$ (this follows from the fact that $\mathbf{S}$ is an upper semi-continuous decomposition of a closed surface). Since the homeomorphism f' preserves every element in $\mathbf{S}'$ we conclude that the hyperbolic distance between $f'(x)$ and x is uniformly bounded for every $x \in \mathbf{D}$. But this implies that f' extends continuously to the identity map of the boundary of $\mathbf{D}$. This shows that $\widehat{f}$ is homotopic to the identity. □

We have already mentioned that from the fact that $\mathbf{S}$ is upper semi-continuous it follows that every $S \in \mathbf{S}$ has an arbitrarily small neighborhood that contains every element of $\mathbf{S}$ it intersects.

Lemma 3.2. *Let $\mathbf{S}$ be an admissible decomposition for a homeomorphism $\widehat{f} : M \to M$. Let $S \in \mathbf{S}$ be an acyclic component. Then there exists a sequence of neighborhoods U_n, $n \in \mathbf{N}$, of S so that the following holds:*

1. *Each U_n is a simply connected subset of M.*
2. *Each U_n contains every element of $\mathbf{S}$ it intersects.*
3. *We have $S = \bigcap_{n \in \mathbf{N}} U_n$.*
4. *We have $\widehat{f}(U_n) = U_n$, $n \in \mathbf{N}$.*

Proof. Let $N = M'_{\mathbf{S}}$ be the connected component of $M_{\mathbf{S}}$ that contains S. By the Theorem 2.1 (the Moore Theorem) there exists a continuous map $\phi : N \to N$ such that for every $q \in N$ we have $\phi^{-1}(q) \in \mathbf{S}$ and for different points q_1 and q_2 the corresponding pre-images are different. Let $\phi(S) = q_0$. Let Δ_n be a sequence of small topological discs so that

$$\bigcap_{n \in \mathbf{N}} \Delta_n = q_0.$$

Let $U_n = \phi^{-1}(\Delta_n)$. Then

$$\bigcap_{n \in \mathbf{N}} U_n = S.$$

Since ϕ is continuous we have that every U_n is a connected open set. Clearly every element of U_n contains every element of $\mathbf{S}$ it intersects. This implies that $\widehat{f}(U_n) = U_n$.

It remains to see that U_n is simply connected. The collection of sets from the decomposition $\mathbf{S}$ that are contained in U_n gives an upper semi-continuous decomposition of U_n and all the components in this decomposition are acyclic (note that we have already seen that U_n is an open and connected set, which means that U_n is a surface). By Theorem 2.1 (the Moore theorem) there is the corresponding continuous map $\psi : U_n \to U_n$. The map ψ is homotopic to the identity map and it maps every component of the decomposition onto a point. The different components are mapped onto different points. That is why there exists a unique map $\varphi : \Delta_n \to U_n$ so that $\psi = \varphi \circ \phi$. The map φ is continuous, and bijective (this follows from the fact that $\mathbf{S}$ is an upper semi-continuous decomposition). The same is true for its inverse, so φ is a homeomorphism. In particular, the surface U_n is simply connected (since it is homeomorphic to Δ_n). ☐

Let $\mathbf{S}_k$, $1 \leq k \leq m$, $m \in \mathbf{N}$ be a finite set of admissible decompositions for a given group G. We define a new collection of sets $\mathbf{S}$ as follows. For a given point $p \in M$, let $S_{p,k} \in \mathbf{S}_k$ be the unique element such that $p \in S_{p,k}$. Let S_p be the connected component of the intersection between sets $S_{p,k}$, $1 \leq k \leq m$, that contains the point p. The elements of the collection $\mathbf{S}$ are the sets S_p, $p \in M$. In this case we say that $\mathbf{S}$ is the descendant of the decompositions $\mathbf{S}_k$, $1 \leq k \leq m$.

Proposition 3.1. *Let G be a group of homeomorphisms of a closed surface M and let $\mathbf{S}_k$, $1 \leq k \leq m$, be a finite set of admissible decompositions for G. Let $\mathbf{S}$ be* *the descendant of the decompositions* $\mathbf{S}_k$*,* $1 \leq k \leq m$*. Then* $\mathbf{S}$ *is an admissible decomposition for* G*.*

Proof. Note the following. Let $1 < m' < m$, $m' \in \mathbf{N}$. Denote by $\mathbf{S}'$ the descendant of the decompositions $\mathbf{S}_k$, $1 \leq k \leq m'$ and by $\mathbf{S}''$ the descendant of the decompositions $\mathbf{S}_k$, $m' < k \leq m$. Then it follows from the definition that the descendant of the decompositions $\mathbf{S}'$ and $\mathbf{S}''$ is the decomposition $\mathbf{S}$ which is the descendant of all of $\mathbf{S}_k$, $1 \leq k \leq m$. In other words, the operation of taking the descendant is transitive.

The proof of the proposition is given by induction. Suppose that the statement is true for $m-1$. For simplicity denote the descendant of the decompositions $\mathbf{S}_k$, $1 \leq k \leq m-1$ by $\mathbf{S}_1$ and the decomposition $\mathbf{S}_m$ by $\mathbf{S}_2$. We have that both $\mathbf{S}_1$ and $\mathbf{S}_2$ are admissible decompositions for G. Let $\mathbf{S}$ be the descendant of $\mathbf{S}_1$ and $\mathbf{S}_2$. It remains to prove that $\mathbf{S}$ is an admissible decomposition for the group G.

It is straightforward that $\mathbf{S}$ is a proper upper semi-continuous decomposition of M. What is not obvious is to show that $\mathbf{S}$ satisfies the two conditions in Definition 3.1. To verify (1) of Definition 3.1 we must show that every $\widehat{f} \in G$ preserves setwise every component of $S_1 \cap S_2$, $S_1 \in \mathbf{S}_1$, $S_2 \in \mathbf{S}_2$. Fix $\widehat{f} \in G$. It follows from the definition that if $S \in \mathbf{S}$ then $\widehat{f}(S) \in \mathbf{S}$. We first claim that $\widehat{f}$ setwise preserves every component of the intersection $S_1 \cap S_2$, for any given $S_1 \in \mathbf{S}_1$ and for any acyclic $S_2 \in \mathbf{S}_2$. We prove this claim below.

Actually to prove this claim it is enough to prove that $\widehat{f}$ setwise preserves every component of the intersection $S_1 \cap S_2$, providing that both S_1 and S_2 are acyclic. We can also assume that S_1 is not a subset of S_2 and that S_2 is not a subset of S_1 because then the claim is trivial. We already have that S_2 is acyclic but we now show that one can assume that both S_1 and S_2 are acyclic.

Let S be a component of $S_1 \cap S_2$. We know that S_2 is not a subset of S_1, so S_2 has non-empty intersection with the surface $M \setminus S_1$. Since S_2 is connected (in particular, S_2 "connects" the set S and some point in the surface $M \setminus S_1$), we conclude that there exists $p \in S$, so that p is a frontier point of the surface $M \setminus S_1$. Since $\mathbf{S}_1$ and $\mathbf{S}_2$ are admissible decompositions, there exists a sequence of points $p_n \in M$, $n \in \mathbf{N}$, so that $p_n \to p$, $n \to \infty$, and so that for each n the component $S_{p_n,1} \in \mathbf{S}_1$, $p_n \in S_{p_n,1}$ is acyclic. Let $S_{p_n,2} \in \mathbf{S}_2$ be the corresponding components that contain p_n. Since the set of acyclic components of $\mathbf{S}_2$ is open we know that for n large enough each $S_{p_n,2}$ is acyclic as well. If we prove that $\widehat{f}$ setwise preserves every component of the intersection $S_{p_n,1} \cap S_{p_n,2}$, then $\widehat{f}$ will preserve S as well. Therefore, it is enough to prove our claim if both S_1 and S_2 are acyclic.

The decomposition $\mathbf{S}$ is upper semi-continuous. As usual, let $M_{\mathbf{S}} \subset M$ be the set of acyclic components of $\mathbf{S}$. If either one of $S_1 \in \mathbf{S}_1$ or $S_2 \in \mathbf{S}_2$ is acyclic then every component of the corresponding intersection must be acyclic. This implies that $M_{\mathbf{S}_1}, M_{\mathbf{S}_2} \subset M_{\mathbf{S}}$. Recall that if $S \in \mathbf{S}$ then $\widehat{f}(S) \in \mathbf{S}$. Let M_1 be a component of $M_{\mathbf{S}}$. After applying the Moore theorem on M_1, and accordingly semi-conjugating the restrictions of $\mathbf{S}_1, \mathbf{S}_2$ and $\widehat{f}$ on M_1, we may assume that for two acyclic components $S_1 \in \mathbf{S}_1$ and $S_2 \in \mathbf{S}_2$, $S_1, S_2 \subset M_1$, we have that every components of $S_1 \cap S_2$ is a point.

Lemma 3.3. *With the notation as above we have the following. Let* $C \subset M_1$ *denote the set of points that are components of the intersection* $S_1 \cap S_2$*, where* $S_1 \in \mathbf{S}_1$ *and*

$S_2 \in \mathbf{S}_2$ are arbitrary acyclic components that are contained in M_1. Then C is an open subset of M_1. Moreover, there exists a sequence $m(k) \in \mathbf{N}$, $k \in \mathbf{N}$, so that the sequence of maps $\widehat{f}^{m(k)}$ converges to the identity map uniformly on compact sets in C.

Proof. Let $S_1 \in \mathbf{S}_1$, $S_2 \in \mathbf{S}_2$ be acyclic so that $S_1, S_2 \subset M_1$ and $C(S_1,S_2) = S_1 \cap S_2 \neq \emptyset$. Each $C(S_1,S_2)$ is totally disconnected set, and C is the union of all such $C(S_1,S_2)$. Since the sets of acyclic components of $\mathbf{S}_1$ and $\mathbf{S}_2$ are open it follows that C is open as well. Now fix such $S_1 \in \mathbf{S}_1$ and $S_2 \in \mathbf{S}_2$. Let $U_n(i)$, $n \in \mathbf{N}$, denote the decreasing sequence of simply connected sets from Lemma 3.2 so that

$$\bigcap_{n\in\mathbf{N}} U_n(i) = S_i.$$

We have $U_n(i) \subset M_1$. The sequence of sets $V_n = U_n(1) \cap U_n(2)$ is a decreasing sequence of open sets, and we have

$$\bigcap_{n\in\mathbf{N}} V_n = C(S_1,S_2).$$

Now we show that for a fixed $n \in \mathbf{N}$, and a given component K_n of V_n, there is $k = k(K_n) \in \mathbf{N}$ so that $\widehat{f}^k(K_n) = K_n$.

The set K_n is open and simply connected (every component of the intersection between two simply connected sets is simply connected). Let $q \in K_n \subset V_n = U_n(1) \cap U_n(2)$. It follows from Lemma 3.2 (and in particular the fact that $U_n(i)$ contains every component of $\mathbf{S}_i$ it intersects), that the closure of the orbit $\widehat{f}^k(q)$, $k \in \mathbf{N}$, is contained in V_n. On the other hand, every point in the frontier component of K_n is contained in either the frontier component of $U_n(1)$ or the frontier component of $U_n(2)$. This means that every accumulation point of the sequence $\widehat{f}^k(\partial K_n)$ is also contained in the union of the frontiers of $U_n(1)$ and $U_n(2)$. This implies that $\widehat{f}^k(p)$ is a uniform distance apart from $\widehat{f}^k(\partial K_n)$, and this lower bound on the distance is independent from k. Since the sets $\widehat{f}^k(K_n)$ and $\widehat{f}^l(K_n)$, $k \neq l$, are either the same or disjoint, we conclude that there could be at most finitely many different sets $\widehat{f}^k(K_n)$. So, for some $k \in \mathbf{N}$, we have $\widehat{f}^k(K_n) = K_n$.

For any $p \in C$, denote the component of V_n that contains p by $V_n(p)$. This is a decreasing sequence of open sets whose intersection is p. Moreover, the Hausdorff limit of this sequence is equal to p (since it is a decreasing sequence), so its diameter tends to zero. Let $C_0 \subset C$ be a compact subset in C. Fix an $n \in \mathbf{N}$. Then the sets $V_n(p)$, $p \in C_0$, give an open cover of C_0. Take a finite sub-cover. Also, note that we can assume that the sets $V_n(p)$, $p \in C_0$, have uniformly small diameter (this is easily achieved by renumbering the sequence $V_n(p)$). Clearly there is a number $m(n) \in \mathbf{N}$, so that $\widehat{f}^{m(n)}$ is uniformly close to the identity map on C_0. By letting $n \to \infty$ we prove the lemma. □

In order to finish the proof of our claim that $\widehat{f}$ setwise preserves every component of $S_1 \cap S_2$, it is enough to assume that S_2 has no interior. There are at most countably many components in $\mathbf{S}_2$ that have non-empty interior. If $p \in S_2$, but p is not in the interior of S_2, there is a sequence of points $p_n \to p$ so that the corresponding components $S_{p_n,2} \in \mathbf{S}_2$ have no interior. If we prove that $\widehat{f}$ fixes every point of $C \cap S_{p_n,2}$, then $\widehat{f}$ would fix every point of $(C \cap S_2) \setminus int(S_2)$ (here

$int(S_2)$ denotes the interior of S_2 and we use the fact that C is open). On the other hand, let $p \in (C \cap int(S_2))$. Let $S_{p,1} \in \mathbf{S}_1$ be the corresponding component that contains p. Since $p \in C$ we have that every component of the intersection $S_{p,1} \cap S_2$ is a point. Let U denote a small simply connected neighborhood of p that is contained in S_2. We have that the component of the intersection $S_{p,1} \cap U$ that contains p is a point. This shows that $S_{p,1} \cap U = p$. Since $S_{p,1}$ is connected we conclude that $S_{p,1} = p$. This implies that $\widehat{f}(p) = p$, and it shows that $\widehat{f}$ fixes every point of $C \cap S_2$.

We now assume that S_2 has no interior. Let z_n be a sequence of points that accumulates in S_2, and let $S_{z_n,2}$ be the corresponding acyclic components in $\mathbf{S}_2$. Since $\widehat{f}(S_{z_n,2}) = S_{z_n,2}$, by the Browder fixed point theorem we conclude that there exists a fixed point for $\widehat{f}$ that is contained in $S_{z_n,2}$. Then these fixed points accumulate on S_2. We now can apply Lemma 2.5 and conclude that $\widehat{f}$ fixes every point in the set $C \cap S_2$. Therefore, $\widehat{f}$ fixes every points in C. This means that $\widehat{f}$ setwise preserves every component of the intersection $S_1 \cap S_2$ providing that S_2 is acyclic.

Let S be a component of $S_1 \cap S_2$, where $S_1 \in \mathbf{S}_1$ and $S_2 \in \mathbf{S}_2$ are arbitrary sets. We show that $\widehat{f}$ preserves setwise every component of $S_1 \cap S_2$. Let S be such component. If S_1 is not contained in S_2, we have that S must have a non-empty intersection with the frontier component of the subsurface $M \setminus S_2$. Let p be a point in this intersection, and let $p_n \in (M \setminus S_2)$ be a sequence of points that converge to p, and such that the corresponding components $S_{p_n,2} \in \mathbf{S}_2$ that contain points p_n are acyclic. Let $S_{p_n,1} \in \mathbf{S}_1$ contain the point p_n. We have proved above that $\widehat{f}$ setwise preserves every component of $S_{p_n,1} \cap S_{p_n,2}$. In particular, $\widehat{f}$ setwise preserves the component of $S_{p_n,1} \cap S_{p_n,2}$ that contains p_n. This yields the conclusion that $\widehat{f}$ setwise preserves S.

The same argument as in the previous paragraph shows that the decomposition $\mathbf{S}$ satisfies the second condition from Definition 3.1. We have proved that $\mathbf{S}$ is an admissible decomposition for every group element $\widehat{f}$ which implies that it is admissible for the group G as well. This proves the proposition. □

For a given group G of homeomorphisms denote by $\mathbf{S}_\alpha$, $\alpha \in \mathbf{A}$, the collection of all admissible decompositions indexed by the set $\mathbf{A}$. The collection $\mathbf{S}_\alpha$, $\alpha \in \mathbf{A}$, has a natural ordering. We say that $\mathbf{S}_\alpha$ and $\mathbf{S}_\beta$ are in the relation $\mathbf{S}_\alpha \leq \mathbf{S}_\beta$ if every element of $\mathbf{S}_\alpha$ is contained in an element of $\mathbf{S}_\beta$. In this case we will also say that $\mathbf{S}_\alpha$ is contained in $\mathbf{S}_\beta$. The relation $\leq$ is clearly transitive. If $\mathbf{S}_\alpha \leq \mathbf{S}_\beta$ and $\mathbf{S}_\beta \leq \mathbf{S}_\alpha$ then $\mathbf{S}_\alpha = \mathbf{S}_\beta$. This means that $\leq$ defines an ordering on the collection $\mathbf{S}_\alpha$, $\alpha \in \mathbf{A}$.

Definition 3.2. *An admissible decomposition* $\mathbf{S}$ *for a group* G *will be called the minimal decomposition for* G *if* $\mathbf{S}$ *is contained in every admissible decomposition for* G*.*

The following is the main result of this section.

Theorem 3.1. *Every group* G *has a unique minimal decomposition.*

Proof. Since $\mathbf{S}_\alpha$, $\alpha \in \mathbf{A}$ is ordered by the Zorn's lemma there exists a maximal totally ordered subset (or a chain) of this collection. For $p \in M$ let S_p be the intersection of all the corresponding components (that contain p) from all the sets from this chain. We define the decomposition $\mathbf{S}$ so its elements are the sets S_p, $p \in M$. Clearly, $\mathbf{S}$ is upper semi-continuous. Since the family $\mathbf{S}_\alpha$ is totally ordered, and since $\widehat{f}$ setwise preserve each S_p, where S_p is in some $\mathbf{S}_\alpha$, it follows that $\widehat{f}$ setwise preserves each element of $\mathbf{S}$. Similarly one can check that the second condition of Definition 3.1 is satisfied, so $\mathbf{S}$ is an admissible decomposition for G.

The hard part of is to prove that $\mathbf{S}$ is the minimal decomposition. But, this is where Proposition 3.1 comes into play. Let $\mathbf{S}_1$ be an arbitrary admissible decomposition for G and let $\mathbf{S}_2$ be the descendant decomposition of $\mathbf{S}$ and $\mathbf{S}_1$. We have that $\mathbf{S}_2$ is contained in $\mathbf{S}$. The fact that the chain we have chosen above is maximal implies that $\mathbf{S}_2 = \mathbf{S}$. In turn this shows that $\mathbf{S} \leq \mathbf{S}_1$. Therefore, $\mathbf{S}$ is contained in every admissible decomposition for G so $\mathbf{S}$ is the minimal decomposition. □

4. The torus analysis

4.1. **Hyperbolic linear transformations.** Let $\mathbf{T}^2$ be the 2-dimensional torus, thought of as $\mathbf{R}^2/\mathbf{Z}^2$, with the euclidean metric. Denote by $\pi : \mathbf{R}^2 \to \mathbf{R}^2/\mathbf{Z}^2 = \mathbf{T}^2$ the universal covering map. Let $A : \mathbf{T}^2 \to \mathbf{T}^2$ be a homeomorphism induced by an element $A_L \in \mathbf{SL}_2(\mathbf{Z})$, which is Anosov or, in a different terminology, a hyperbolic linear transformation (A_L is a lift of A to the universal cover $\mathbf{R}^2$ of $\mathbf{T}^2$). If (x, y) is the parameter in the plane $\mathbf{R}^2$ then we can choose A_L to be $A_L(x, y) = (2x + y, x + y)$. Note that for this particular choice of A the only fixed point it has is the point on the torus $\mathbf{T}^2$ that is represented by the origin $(0, 0) \in \mathbf{R}^2$. For $n > 1$, the map $A^n : \mathbf{T}^2 \to \mathbf{T}^2$ has more than one, but finitely many, fixed points.

Such A_L has two positive real eigenvalues, one greater than 1 and one less than 1. These eigenvalues are irrational algebraic numbers. The map A preserves two transverse foliations of $\mathbf{T}^2$. These foliations are by geodesics, parallel to the eigenvectors of A_L. $\mathbf{T}^2$ has a basepoint $0_{\mathbf{T}^2}$ (represented by the origin $(0, 0) \in \mathbf{R}^2$ in the universal cover $\mathbf{R}^2$) and this basepoint is preserved by A.

We can 'blow up' the basepoint of $\mathbf{T}^2$ to a circle, as follows. Let D be a round disc centered at the basepoint, of radius $\epsilon > 0$. We choose ϵ small enough so that D does not intersect itself in $\mathbf{T}^2$. We think of $\mathbf{T}^2$ as the union of the closed round disk D and its complement $\mathbf{T}^2 \setminus D$. Let $TB\mathbf{T}^2$ denote the tangent bundle of $\mathbf{T}^2$, that is $TB\mathbf{T}^2$ is the set of pairs (x, v), where $x \in \mathbf{T}^2$, and v is in the tangent space to $\mathbf{T}^2$ at x. We define the map $i' : \mathbf{T}^2 \setminus \{0_{\mathbf{T}^2}\} \to TB\mathbf{T}^2$ as follows. We set

$$i'(x) = (x, \frac{(\epsilon - ||x||)x}{\epsilon ||x||}),$$

for $x \in D$ and $i'(x) = (x, 0)$ for x that is not in D. The map i' is a homeomorphism of $\mathbf{T}^2 \setminus \{0_{\mathbf{T}^2}\}$ onto its image. Moreover, the map i' is smooth everywhere except on the circle that is the boundary of the disc D. We now define a new map $i : \mathbf{T}^2 \setminus \{0_{\mathbf{T}^2}\} \to TB\mathbf{T}^2$ so that $i'(x) = i(x)$ for x not in D and for x that belongs to the disc centered at the basepoint, of radius $\frac{\epsilon}{2}$. On the rest we define the map

i so that it is a diffeomorphism (and smooth in particular) of $\mathbf{T}^2 \setminus \{0_{\mathbf{T}^2}\})$ onto its image.

Let $T = i(\mathbf{T}^2 \setminus \{0_{\mathbf{T}^2}\})$. Then T is diffeomorphic to $\mathbf{T}^2 \setminus \{0_{\mathbf{T}^2}\}$. The closure $\overline{T}$ of T in the tangent bundle of $\mathbf{T}^2$ adjoins the circle of unit tangent vectors to $\mathbf{T}^2$ at $0_{\mathbf{T}^2}$. The map A (same as any other element of $\mathbf{SL}_2(\mathbf{Z})$) is in an obvious way conjugated to a map $A' : \overline{T} \to \overline{T}$.

Let M be the surface of genus $\mathbf{g} \geq 2$. We build up M as follows. We start with T, a closed annulus N and a compact surface R' of genus $\mathbf{g}-1$ with one end. We glue N to T by identifying one frontier component of N to ∂T, and we glue N to R' by identifying the other frontier component with $\partial R'$. This gives us $M = R' \cup N \cup T$. Define

$$\mathcal{P} : M = R' \cup N \cup T \to \mathbf{T}^2$$

as follows. $\mathcal{P}$ sends $R' \cup N$ to the base point $0_{\mathbf{T}^2} \in \mathbf{T}^2$. On T, we define $\mathcal{P}$ to be the diffeomorphism $i^{-1} : T \to \mathbf{T}^2 \setminus \{0_{\mathbf{T}^2}\}$ defined above. Then $\mathcal{P}$ is smooth.

Let A_M be a diffeomorphism of M defined as follows. On T, A_M agrees with the diffeomorphism $A' : \overline{T} \to \overline{T}$ defined above. The effect of A' on the boundary circle ∂T is that there are four fixed points corresponding to the two eigenvectors, and in each complementary interval on the circle, there is a diffeomorphism moving points away from the contracting eigenvector and towards the expanding eigenvector. We extend the action of A' to the annulus N by constructing an isotopy of each of the four diffeomorphisms to the identity. This extension defines A_M on N. Finally, we define A_M to be the identity on R'. In this construction all the operations and all mappings are C^∞. We have a commutative diagram

$$\mathcal{P} \circ A_M = A \circ \mathcal{P} : M = R' \cup N \cup T \to \mathbf{T}^2. \tag{4.1}$$

Let $\pi_M : \widetilde{M} \to M$ be the normal cover induced by $\mathcal{P}$. $\widetilde{M}$ is the corresponding covering surface with group of covering transformations equal to the free Abelian group on two generators (see Figure 1). We define the lift $\widetilde{\mathcal{P}} : \widetilde{M} \to \mathbf{R}^2$ so that the relation

$$\pi \circ \widetilde{\mathcal{P}} = \mathcal{P} \circ \pi_M : \widetilde{M} \to \mathbf{T}^2 \tag{4.2}$$

holds. The inverse image of any integer point in $\mathbf{R}^2$ under $\widetilde{\mathcal{P}}$ is a copy of $R = R' \cup N$. We denote the inverse image of such a point $(m,n) \in \mathbf{Z} \times \mathbf{Z}$ by $R_{m,n}$. A covering transformation of $\widetilde{M}$ induces a diffeomorphism $R_{m,n} \to R_{p,q}$. Since the identification of $R_{m,n}$ with R is given by the covering map π_M, the covering transformation corresponds to the identity map on R. We choose a basepoint for $\widetilde{M}$ in $R_{0,0}$. We lift $A_M : M \to M$ to $\widetilde{A}_M : \widetilde{M} \to \widetilde{M}$, so that it fixes $R_{0,0}$ pointwise. Let m and n be integers. We define integers m_1 and n_1 by $A_L(m,n) = (m_1, n_1)$. Then the restriction of the map $\widetilde{A}_M$ on $R_{m,n}$ agrees with the corresponding covering transformation and $\widetilde{A}_M$ sends $R_{m,n}$ to R_{m_1,n_1}.

Note that $\widetilde{M}$ minus the union of all $R_{m,n}$, $m,n \in \mathbf{Z}$, is a planar surface. We will need the following lemma.

Lemma 4.1. *Let α be a simple closed curve in M such that for some $n \in \mathbf{N}$ the n-th power of A_M preserves the homotopy class of α, that is we have $[A_M{}^n(\alpha)] = [\alpha]$. Then α is homotopic to a simple closed curve that is contained in R.*

Proof. We may assume that α has the least intersections with the boundary circle of the surface R among all simple closed curves in its homotopy class. If this intersection number is 0 then α is either contained in R or it is contained in T. If α was contained in T, and α was not homotopic to the boundary circle of the surface R, then $\mathcal{P}(\alpha)$ would be a simple closed curve on the torus $\mathbf{T}^2$ and such that $[A^n(\mathcal{P}(\alpha))] = [\mathcal{P}(\alpha)]$ (if α is homotopic to the boundary circle of R then it is homotopic to a simple closed curve that is contained in R). But this is impossible since A is an Anosov map of the torus.

If α intersects the boundary circle of R we proceed in a similar way. Note that $\mathcal{P}(\alpha)$ is a union of finitely many simple closed curves on the torus $\mathbf{T}^2$. If either one of these curves was homotopically trivial on $\mathbf{T}^2$ we would have that α does not have the smallest number of intersections with the boundary circle of R in its homotopy class. Therefore we have that A^n does not preserve the homotopy class of any of the simple closed curves on $\mathbf{T}^2$ whose union is equal to $\mathcal{P}(\alpha)$. This contradicts the assumption $[{A_M}^n(\alpha)] = [\alpha]$. $\square$

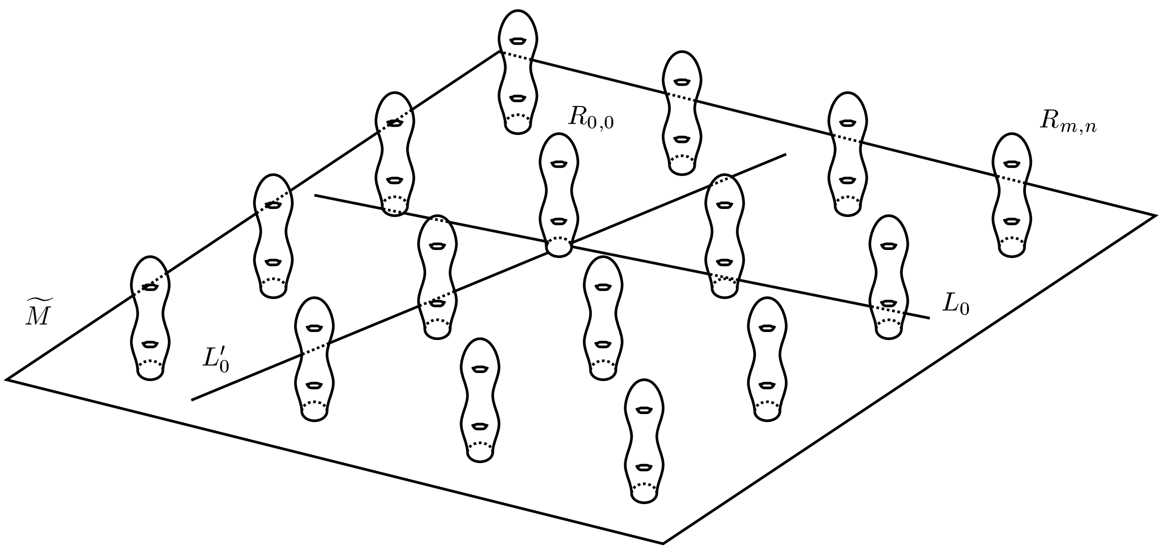

FIGURE 1. This figure illustrates the surface $\widetilde{M}$, when the corresponding closed surface M is of genus $\mathbf{g} = 3$. We choose the base point for $\widetilde{M}$ in the set $R_{0,0}$. The sets L_0 and L'_0 are defined below.

Let $A_{\mathcal{E}} : M \to M$ be any homeomorphism that is isotopic to A_M. We lift the isotopy from A_M to $A_{\mathcal{E}}$ to obtain an isotopy between $\widetilde{A_M}$ and $\widetilde{A_{\mathcal{E}}}$. Here $\widetilde{A_M}, \widetilde{A_{\mathcal{E}}} : \widetilde{M} \to \widetilde{M}$. We choose a hyperbolic metric on M, and this induces a hyperbolic metric on $\widetilde{M}$. Denote the corresponding distance on M by $\mathbf{d}$ and by $\tilde{\mathbf{d}}$ the corresponding distance on $\widetilde{M}$. There is an associated complex structure on M (and consequently on $\widetilde{M}$) that corresponds to $\mathbf{d}$. That is, we consider M and $\widetilde{M}$ as Riemann surfaces.

A standard argument using the compactness of M shows that there is a number $\epsilon_1 > 0$ such that $\mathbf{d}(\widetilde{A_M}, \widetilde{A_{\mathcal{E}}}) < \epsilon_1$. Given $m, n \in \mathbf{Z}^2$, we denote by $\tau(m,n) : \widetilde{M} \to \widetilde{M}$ the corresponding covering transformation which maps $R_{0,0}$ to $R_{m,n}$, that is $\tau(m,n)(R_{0,0}) = R_{m,n}$.

Let $\mathcal{F}$ be the set of homeomorphisms $\widetilde{h}$ of $\widetilde{M}$ that are lifts of homeomorphisms of M, such that, for $m, n \in \mathbf{Z}^2$,

$$\widetilde{h} \circ \tau(m,n) \circ \widetilde{h}^{-1} = \tau(A_L(m,n)).$$

Lemma 4.2. *Given any continuous path of homeomorphisms of $\widetilde{M}$ that are lifts of homeomorphisms of M, if the path meets $\mathcal{F}$, then it is completely contained in $\mathcal{F}$. In particular, $\widetilde{A_M}$ and $\widetilde{A_{\mathcal{E}}}$ are both elements of $\mathcal{F}$.*

Proof. Let the given path of homeomorphisms be $\{\widetilde{f}_t\}$. Then

$$\pi_M \circ \widetilde{f}_t \circ \tau(m,n) \circ \widetilde{f}_t^{-1} = \pi_M.$$

Therefore $\widetilde{f}_t \circ \tau(m,n) \circ \widetilde{f}_t^{-1}$ is a covering transformation, varying continuously with t. But the group of covering transformations is discrete, so $\widetilde{f}_t \circ \tau(m,n) \circ \widetilde{f}_t^{-1}$ is independent of t. Therefore the path $\{\widetilde{f}_t\} \subset \mathcal{F}$. We have $A_L \circ \tau(m,n) \circ {A_L}^{-1} = \tau(A_L(m,n))$ on $\mathbf{R}^2$. Lifting the diagram (4.1) above to the appropriate coverings shows that $\widetilde{A_M} \in \mathcal{F}$. Since A_M is isotopic to $A_{\mathcal{E}}$, there is a path connecting A_M to $A_{\mathcal{E}}$ and its lift connecting $\widetilde{A_M}$ to $\widetilde{A_{\mathcal{E}}}$, and so $\widetilde{A_{\mathcal{E}}} \in \mathcal{F}$. □

We also have the following lemma.

Lemma 4.3. *There exists $\epsilon > 0$ such that for each point $x \in \widetilde{M}$,*

$$dist(\widetilde{\mathcal{P}}(\widetilde{A_{\mathcal{E}}}(x)), A_L(\widetilde{\mathcal{P}}(x))) = dist(\widetilde{\mathcal{P}}(\widetilde{A_{\mathcal{E}}}(x)), \widetilde{\mathcal{P}}(\widetilde{A_M}(x))) < \epsilon.$$

Here dist stands for the standard Euclidean distance in the plane $\mathbf{R}^2$.

Proof. This follows since $\widetilde{\mathcal{P}} : \widetilde{M} \to \mathbf{R}^2$ is uniformly continuous and the hyperbolic metric of $\widetilde{M}$ is a path metric (so that one long distance in $\widetilde{M}$ can be split up into many small distances). □

4.2. **Decomposition of the surface $\widetilde{M}$.** Let λ_1 and λ_2 be the two eigenvalues of A_L, with $\lambda_1 > 1 > \lambda_2 > 0$. Every line in $\mathbf{R}^2$ parallel to the λ_2-eigenvector meets at most one point of $\mathbf{Z}^2$. Therefore the inverse image under $\widetilde{\mathcal{P}}$ of such a line is either a copy of $\mathbf{R}$ embedded in $\widetilde{M}$ or the union of two half-lines with a copy of $\overline{R}$, where $R = N \cup R'$ is a surface (defined above) of genus $\mathbf{g} - 1$ with one end. For each $a \in \mathbf{R}$, let $\ell_a \subset \mathbf{R}^2$ be the line parallel to the λ_2-eigenvector and at signed distance a from the origin. To fix the sign, we fix the eigenvectors e_i corresponding to λ_i, for $i = 1, 2$. Points on ℓ_a have coordinates of the form $x_1 e_1 + x_2 e_2$, with x_1 fixed and x_2 varying. We declare that a has the same sign as x_1. Since the vectors e_1 and e_2 are not orthogonal to each other (the two eigenvectors of an Anosov map can not be mutually orthogonal) we have that $|a| < |x_1|$.

Let $L_a = \widetilde{\mathcal{P}}^{-1}(\ell_a) \subset \widetilde{M}$ (see Figure 1). The subsets $\{L_a\}_{a \in \mathbf{R}}$ partition $\widetilde{M}$. Each subset L_a is closed and separates $\widetilde{M}$ into exactly two components. If $b < c$, let $B_{b,c} \subset \widetilde{M}$ be the inverse image of the closed subspace of $\mathbf{R}^2$ lying between ℓ_b and ℓ_c. So $B_{b,c} \subset \widetilde{M}$ can also be defined as the closed subspace lying between and including L_b and L_c.

Lemma 4.4. *Let $k > \epsilon/(\lambda_1 - 1)$, where $\epsilon > 0$ is the constant of Lemma 4.3 We fix $a \in \mathbf{R}$. For each $n \geq 0$, let $P_{a,n} = B_{b(n),c(n)}$, where $b(n) = \lambda_1^n a - k$ and $c(n) = \lambda_1^n a + k$. Then $\widetilde{A_\mathcal{E}}(P_{a,n}) \supset P_{a,n+1}$.*

Proof. By Lemma 4.3 we have that $\widetilde{A_\mathcal{E}}(L_{b(n)}) \subset B_{r,s}$, where

$$r = \lambda_1 b(n) - \epsilon = b(n+1) - (\lambda_1 - 1)k - \epsilon$$

and

$$s = \lambda_1 b(n) + \epsilon = b(n+1) - (\lambda_1 - 1)k + \epsilon.$$

We have a similar situation for $\widetilde{A_\mathcal{E}}(L_{c(n)})$. Therefore $\widetilde{A_\mathcal{E}}(P_{a,n}) \supset B_{x,y}$, where

$$x = b(n+1) - (\lambda_1 - 1)k + \epsilon < b(n+1)$$

and

$$y = c(n+1) + (\lambda_1 - 1)k - \epsilon > c(n+1).$$

This gives the desired result. ☐

Definition 4.1. *Let k and $P_{a,n}$ be as in the statement of the previous lemma. We define*

$$Q_a = Q_a(k) = \bigcap_{n \geq 0} \widetilde{A_\mathcal{E}}^{-n}(P_{a,n})$$

which is the intersection of the decreasing sequence of closed subsets of $\widetilde{M}$.

Remark. Note that in the above definition we can take k to be any number that is large enough to define $Q_a(k)$. If we take $k_1 < k_2$ then it is an interesting question to determine whether $Q_a(k_1) = Q_a(k_2)$. From now on, since we have fixed $A_\mathcal{E}$, we fix one k large enough and take $Q_a = Q_a(k)$.

We will say that ℓ_a is to the left of ℓ_b if $a < b$. Orienting the lines ℓ_a in the direction of the eigenvector e_2 of A_L, we can talk of the subspace of $\mathbf{R}^2$ to the left or the right of ℓ_a. The complement of L_a in $\widetilde{M}$ also consists of two components, one to the left and one to the right.

Lemma 4.5. *The set Q_a is a closed subset of $\widetilde{M}$, separating $\widetilde{M}$ into exactly two components, one on each side of Q_a.*

Proof. Note that $Q_a \subset P_{a,0}$. Let p and q be points of $\widetilde{M}$ very far to the left and right respectively of L_a. We assume by contradiction that there is a path in the complement of Q_a joining p to q. For each n, the path intersects $\widetilde{A_\mathcal{E}}^{-n}(P_{a,n})$ in a compact set. This is a decreasing sequence of compact subsets which therefore has a non-empty intersection. So the path meets Q_a which is a contradiction. This shows that the complement of Q_a contains at least two components, one extending arbitrarily far to the left of Q_a and the other arbitrarily far to the right.

We next assume by contradiction that there are three points p, q and r in different components of the complement of Q_a. For n sufficiently large, these three points are in the complement of $\widetilde{A_\mathcal{E}}^{-n}(P_{a,n})$. Since $P_{a,n}$ has exactly two complementary components, and $\widetilde{A_\mathcal{E}}$ is a homeomorphism, at least two of the three points must lie in the same component of the complement of $\widetilde{A_\mathcal{E}}^{-n}(P_{a,n})$. Therefore they are in the same component of the complement of Q_a, a contradiction. This proves that there are exactly two components of $\widetilde{M} \setminus Q_a$, one extending infinitely far to the left and the other infinitely far to the right. $\square$

Lemma 4.6. *If $a < b$, then $Q_a \cap Q_b = \emptyset$ and Q_b is to the right of Q_a.*

Proof. For n large, $P_{a,n} \cap P_{b,n} = \emptyset$ and $P_{b,n}$ is to the right of $P_{a,n}$. The result follows. $\square$

Remark. The set Q_a does not have to be connected in $\widetilde{M}$. We can compactify $\widetilde{M}$ by adding the point at infinity. This new space $\widetilde{M_1}$ is a compact metric space but it is not a surface. Moreover, the homeomorphisms $\widetilde{A_M}$ and $\widetilde{A_{\mathcal{E}}}$ extend to homeomorphisms of $\widetilde{M_1}$ and both maps fix the added point. If to each set $P_{a,n}$ we add this new point then the set Q_a (plus the added point) is obtained as the intersection of a decreasing family of connected subsets of a complete and compact metric space. Therefore, the extended set Q_a is a connected subset of $\widetilde{M_1}$. This implies that the set Q_a does not have any components that are compact subsets of the surface $\widetilde{M}$.

Definition 4.2. *Given $a \in \mathbf{R}$, we define T_a to be the set of $x \in \widetilde{M}$, such that, if $b > a$, then x is to the left of Q_b and, if $b < a$, then x is to the right of Q_b. By the above lemma, $Q_a \subset T_a$.*

Remark. One interesting question is whether $Q_a = T_a$. This is related to the previously raised question of whether the definition of Q_a depends on the parameter k we choose (see the remark after Definition 4.1).

Lemma 4.7. *For each $a \in \mathbf{R}$, the following holds*

(1) *$\widetilde{M}$ is the disjoint union of the closed sets T_a.*
(2) *T_a is a closed subset of $\widetilde{M}$, whose complement has exactly two components, each contained in the corresponding component of the complement of Q_a.*
(3) *For $a < b$ the open set $\widetilde{M} \setminus (T_a \cup T_b)$ has three unbounded components.*

Proof. Suppose that $a < b$. Choose c with $a < c < b$. Then T_a and T_b are on opposite sides of Q_c. Therefore T_a and T_b are disjoint.

Let $x \in \widetilde{M}$. We want to show that for some $a \in \mathbf{R}$ we have $x \in T_a$. Using the fact that $Q_b \subset P_{b,0}$, we easily find numbers $c > b$ such that x lies to the left of Q_c and to the right of Q_b. We define

$$I = \{c : x \text{ lies to the left of } Q_c\}, J = \{b : x \text{ lies to the right of } Q_b\}.$$

Then I and J are disjoint semi-infinite intervals. One sees that $\overline{I} \cap \overline{J} = \{a\}$ for some a (otherwise the sets I and J could be enlarged). It follows that $x \in T_a$, as required. We next show that T_a is closed. Let $\{x_n\}$ be a sequence of points in T_a, converging to some point $x_\infty \in \widetilde{M}$. We have just shown that $x_\infty \in T_b$ for some b. If $b \neq a$, we easily get a contradiction from Lemma 4.6. This proves (1).

Let $\Omega_a^L = \bigcup_{b<a} T_b$ and $\Omega_a^R = \bigcup_{b>a} T_b$. It follows from (1) and Lemma 4.5 that Ω_a^L and Ω_a^R are connected open sets. This proves (2). Set $\Omega_a^L = \bigcup_{c<a} T_c$, $\Omega_{a,b} = \bigcup_{a<c<b} T_c$ and $\Omega_b^R = \bigcup_{c>b} T_c$. Similarly we have (3). $\square$

We also have the following lemma.

Lemma 4.8. *Let $a \in \mathbf{R}$.*

(1) *$T_a \subset P_{a,0}$.*
(2) *Let a_n, $n \in \mathbf{N}$, be a sequence of real numbers such that $a_n < a$ and $a_n \to a$, $n \to \infty$. Then the Hausdorff limit of the sets T_{a_n} exists and is contained in T_a.*
(3) *Let a_n and b_n be two sequences of real numbers such that $a_n < a < b_n$ for each $n \in \mathbf{N}$, and $a_n, b_n \to a$, $n \to \infty$. Let $\Omega_n = \Omega_{a_n,b_n}$ be the component of $\widetilde{M} - (T_{a_n} \cup T_{b_n})$ that is in between T_{a_n} and T_{b_n}. Then*

$$T_a = \bigcap_{n \in \mathbf{N}} \Omega_n.$$

Proof. If $b < a$ then T_a is to the right of $Q_b \subset P_{b,0}$. This means that T_a is to the right of every L_{b-k}, whenever $b < a$ and where k is the constant from Lemma 4.4. Similarly T_a is to the left of every L_{c+k}, whenever $c > a$. This implies that $T_a \subset P_{a,0} = B_{a-k,a+k}$. This proves (1).

Let T'_a be a Hausdorff limit of a subsequence of the sequence T_{a_n}. Then for n large enough, every T_{a_n} is either in this subsequence or it is in between two elements of this subsequence. This implies that T'_a is the limit of the whole sequence T_{a_n}. If $x \in T'_a$, then x either belongs to T_a or to Ω^L_a. If $x \in \Omega^L_a$ then there exists $b < a$ such that $x \in T_b$. But we could find n large enough so that T_{a_n} is to the right of T_b which is a contradiction. This proves (2). The proof of (3) follows the same way. $\square$

The following lemma follows readily from the definition of the sets Q_a and T_a, $a \in \mathbf{R}$.

Lemma 4.9. *Let $a \in \mathbf{R}$ and $m, n \in \mathbf{Z}$. Define $b \in \mathbf{R}$ by $\tau(m,n)(L_a) = L_b$. Then $\tau(m,n)(Q_a) = Q_b$ and $\tau(m,n)(T_a) = T_b$. Also, if we define $c \in \mathbf{R}$ by $A_L(\ell_a) = \ell_c$ (or equivalently by $\widetilde{A_M}(L_a) = L_c$), then $\widetilde{A_{\mathcal{E}}}(Q_a) = Q_c$ and $\widetilde{A_{\mathcal{E}}}(T_a) = T_c$.*

Remark. It follows from the Remark after Lemma 4.6 and the definition of T_a, that the set T_a does not have any components that are compact subsets of $\widetilde{M}$.

The map ${A_L}^{-1}$ is an Anosov diffeomorphism with the eigenvalues $\frac{1}{\lambda_1} < 1 < \frac{1}{\lambda_2}$. For each $a \in \mathbf{R}$, let $\ell'_a \subset \mathbf{R}^2$ be the line parallel to the λ_1-eigenvector and at signed distance a from the origin. In the same way as above, we define $L'_a = \widetilde{\mathcal{P}}^{-1}(\ell'_a) \subset \widetilde{M}$. The subsets $\{L'_a\}_{a \in \mathbf{R}}$ have the similar properties as the subsets $\{L_a\}_{a \in \mathbf{R}}$. If $b < c$, let $B'_{b,c} \subset \widetilde{M}$ be the inverse image of the closed subspace of $\mathbf{R}^2$ lying between ℓ'_b and ℓ'_c. Repeating the same process as above we define

$$Q'_a = \bigcap_{n \geq 0} \widetilde{A_{\mathcal{E}}}^n (P'_{a,n})$$

where $P'_{a,n}$ has the obvious meaning. Now, we define T'_a in the appropriate way and it clearly satisfies the statements similar to lemmas 4.4, 4.5, 4.6, 4.7, 4.8, 4.9. The sets T_a and T'_a are called the separators.

For $x, y \in \mathbf{R}$, set

$$\widetilde{S}_{x,y} = T_x \cap T'_y.$$

Since the separators T_x and T'_y are contained in $P_{x,0}$ and $P'_{y,0}$ respectively, the set $\widetilde{S}_{x,y}$ is a compact subset of $\widetilde{M}$. From Lemma 4.9 we have that for every x, y, no two points of the set $\widetilde{S}_{x,y}$ are identified under the action of the covering group that acts on $\widetilde{M}$ (excluding the action of the identity element of the group).

We define the set $S_{x,y} = \pi_M(\widetilde{S}_{x,y}) \subset M$. If r, s are such that for some element τ of the covering group we have $\tau(T_x) = T_r$ and $\tau(T'_y) = T'_s$, then $\tau(\widetilde{S}_{x,y}) = \widetilde{S}_{r,s}$, and we have that $S_{x,y} = S_{r,s}$. Also, we have that the restriction of the covering map π_M on each $\widetilde{S}_{x,y}$ is a homeomorphism. Note that Lemma 4.9 implies that for each x, y we have $\widetilde{A_{\mathcal{E}}}(\widetilde{S}_{x,y}) = \widetilde{S}_{r,s}$ for suitable r, s.

4.3. **Some important properties of the decomposition S.** We will say that $\widetilde{S}_{x,y}$ is equivalent to $\widetilde{S}_{r,s}$ if $\tau(\widetilde{S}_{x,y}) = \widetilde{S}_{r,s}$ for some covering transformation (this is the same as saying that $S_{x,y} = S_{r,s}$).

Lemma 4.10. *Let $\widetilde{S}$ be a component of $\widetilde{S}_{x,y}$ for any $x, y \in \mathbf{R}$. Then $\widetilde{S}$ does not separate $\widetilde{M}$.*

Proof. $\widetilde{S}$ is a compact subset of $\widetilde{M}$ and it is contained in both T_x and T'_y. If $\widetilde{S}$ is to separate $\widetilde{M}$, then one of the open sets in its complement would be compactly contained in $\widetilde{M}$. This is not possible, since by Lemma 4.7, T_x (and T'_y as well) separates $\widetilde{M}$ into exactly two components, and both of these are unbounded in $\widetilde{M}$. So, $\widetilde{S}$ does not separate $\widetilde{M}$. □

Lemma 4.11. *Let $x, y \in \mathbf{R}$ be such so that the sets $\widetilde{A_{\mathcal{E}}}^m(\widetilde{S}_{x,y})$ and $\widetilde{A_{\mathcal{E}}}^n(\widetilde{S}_{x,y})$ are not equivalent for any two integers $m \neq n$ and $1 \leq m, n \leq 3g-2$, where g is the genus of the surface M. Let V denote the union of all such sets $\widetilde{S}_{x,y}$. Then every component of the set $\widetilde{S}_{x,y}$ is acyclic. Also, the set V is open, connected and unbounded subset of $\widetilde{M}$.*

Proof. Let $\widetilde{S}$ be a component of $\widetilde{S}_{x,y}$ where $x, y \in \mathbf{R}$ are as in the statement of the lemma. Let $U \subset \widetilde{M}$ be an open connected set that contains $\widetilde{S}$. Since $\widetilde{S}$ does not separate $\widetilde{M}$ we can choose U, so that U does not separate $\widetilde{M}$ too. Note that the covering map is a homeomorphism on the union

$$\bigcup_{1 \leq i \leq 3g-2} \widetilde{A_{\mathcal{E}}}^i(\widetilde{S}).$$

So we can choose U small enough so that the covering map π_M is a homeomorphism on the union

$$\bigcup_{1 \leq i \leq 3g-2} \widetilde{A_{\mathcal{E}}}^i(U).$$

Note that since U does not separate $\widetilde{M}$ neither does the above union.

Suppose that $\alpha \subset U$ is a simple closed curve that is not homotopically trivial in $\widetilde{M}$. Set $\alpha_i = \widetilde{A_{\mathcal{E}}}^i(\alpha)$, $1 \leq i \leq 3g-2$. Note that α_i is disjoint from $\tau(\alpha_j)$, for any covering transformation τ and any $i \neq j$. Suppose that α_i is homotopic to $\tau(\alpha_j)$, for some $i \neq j$, and for some covering transformation τ. Then these two curves bound an annulus $Ann_{i,j}$ in $\widetilde{M}$ because these two curves are disjoint (here we use the fact that two disjoint simple closed curves on any surface, that are homotopic and not homotopically trivial, always bound an annulus on that surface).

Since the sets $\widetilde{A_{\mathcal{E}}}^i(U)$ and $\tau(\widetilde{A_{\mathcal{E}}}^j(U))$ are disjoint, and since the union $\widetilde{A_{\mathcal{E}}}^i(U) \cup \tau(\widetilde{A_{\mathcal{E}}}^j(U))$ does not separate $\widetilde{M}$, we conclude that the surface $\widetilde{M} \setminus (\widetilde{A_{\mathcal{E}}}^i(U) \cup \tau(\widetilde{A_{\mathcal{E}}}^j(U)))$ is contained in the annulus $Ann_{i,j}$. This is impossible because the surface $\widetilde{M} \setminus (\widetilde{A_{\mathcal{E}}}^i(U) \cup \tau(\widetilde{A_{\mathcal{E}}}^j(U)))$ has infinite genus. This gives a contradiction and proves that α_i is not homotopic to $\tau(\alpha_j)$, for any $i \neq j$, and for any τ.

Let $\beta_i = \pi_M(\alpha_i)$. Since the covering map from $\widetilde{M}$ to M is injective on U, it follows that β_i, $1 \leq i \leq 3g-2$, are mutually non-homotopic and homotopically non-trivial simple closed curves on M. But there could be at most $3g-3$ such curves which gives a contradiction to the assumption that α is homotopically non-trivial curve in $\widetilde{M}$. Because U is connected we conclude that U is simply connected and this shows that $\widetilde{S}$ is acyclic.

We now study the set V. Assume that $x, y \in \mathbf{R}$ are such so that $\widetilde{S}_{x,y}$ is not contained in V. This means that the sets $\widetilde{A_{\mathcal{E}}}^m(\widetilde{S}_{x,y})$ and $\widetilde{A_{\mathcal{E}}}^n(\widetilde{S}_{x,y})$ are equivalent for some $1 \leq m < n \leq 3g-2$. Denote by $z_{x,y}$ the point on the torus $\mathbf{T}^2$ that is obtained as the projection of the point with coordinates $(xe_1 + ye_2) \in \mathbf{R}^2$, that is $(xe_1 + ye_2) = z_{x,y} \in \mathbf{T}^2$ (recall that e_1 and e_2 are the eigenvectors that correspond to the Anosov map $A : \mathbf{T}^2 \to \mathbf{T}^2$). It follows from the construction of the sets $\widetilde{S}_{x,y}$ that if $\widetilde{A_{\mathcal{E}}}^m(\widetilde{S}_{x,y})$ and $\widetilde{A_{\mathcal{E}}}^n(\widetilde{S}_{x,y})$ are equivalent for some $1 \leq m < n \leq 3g-2$, then $A^m(z_{x,y}) = A^n(z_{x,y})$. It is well known (see [2]) that there are at most finitely many such points on $\mathbf{T}^2$. This implies that there are only countably many sets $\widetilde{S}_{x,y}$ that are not in V. Moreover, in every compact subset of $\widetilde{M}$ there at most finitely many such sets. Since none of them separates $\widetilde{M}$, then the countable union does not separate $\widetilde{M}$ either. Also, since every accumulation point of $\widetilde{M} \setminus V$ that is in $\widetilde{M}$ also belongs to $\widetilde{M} \setminus V$, we conclude that V is open. Clearly, V is unbounded.

□

Let $\widetilde{\mathbf{S}}$ be the collection of all components of the sets $\widetilde{S}_{x,y}$. Set $\mathbf{S} = \pi_M(\widetilde{\mathbf{S}})$.

Proposition 4.1. $\mathbf{S}$ *is a proper upper semi continuous decomposition of* M*. Moreover, there exists a simple closed curve* γ*, that is homotopic to* N*, such that if* $p \in M$ *belongs to the component of* $M \setminus \gamma$ *that is homotopic to* T*, then the component of* $\mathbf{S}$ *that contains* p *is acyclic.*

Remark. Although the curve γ is homotopic to N it still might have non-empty intersection with T. This is why the component of $M \setminus \gamma$ that is homotopic to T does not have to contain T.

Proof. It follows from Lemma 4.8, and the fact that the Hausdorff limit of connected sets is a connected set (in this particular setting, because the corresponding space is compact and connected), that $\widetilde{\mathbf{S}}$ is an upper semi continuous decomposition of $\widetilde{M}$. This also shows that $\mathbf{S}$ is an upper semi-continuous decompositions of M.

Let $M_{\mathbf{S}}$ be the set of acyclic components of $\mathbf{S}$ (see Proposition 2.1). This is an open set so that every component of $M_{\mathbf{S}}$ is a proper subsurface of M. Let $\widetilde{M_{\widetilde{\mathbf{S}}}}$ be the set of acyclic components of $\widetilde{\mathbf{S}}$. Then the open set $\widetilde{M_{\widetilde{\mathbf{S}}}}$ is the union of all the lifts of $M_{\mathbf{S}}$. It is clear that $A_{\mathcal{E}}$ does not setwise preserve each component of $\mathbf{S}$, but if $S \in \mathbf{S}$ then $A_{\mathcal{E}}(S) \in \mathbf{S}$ (and also $\widetilde{S} \in \widetilde{\mathbf{S}}$ implies $\widetilde{A_{\mathcal{E}}}(S) \in \widetilde{\mathbf{S}}$). We have that $A_{\mathcal{E}}(M_{\mathbf{S}}) = M_{\mathbf{S}}$.

Let $\widetilde{M_1}$ be the component of $\widetilde{M_{\widetilde{\mathbf{S}}}}$ that contains the set V from the previous lemma. Set $\pi_M(\widetilde{M_1}) = M_1$. We have that M_1 is a component of the set $M_{\mathbf{S}}$, and its lift to $\widetilde{M}$ is unbounded. This implies that M_1 is not simply connected. Therefore, M_1 is a proper subsurface of M with finitely many ends. Assume it has k ends, $k \in \mathbf{N}$. Each end of M_1 corresponds to the homotopy class of a simple closed curve $\alpha_i \subset M_1$, $1 \leq i \leq k$. Also, each α_i is homotopically non-trivial in M.

For every $j \in \mathbf{N}$, we have that ${A_{\mathcal{E}}}^j(M_1)$ is either equal to M_1 or it has no intersection with M_1. Since M_1 is not simply connected, and it is a proper subsurface, it follows that for some $n \in \mathbf{N}$ we have ${A_{\mathcal{E}}}^n(M_1) = M_1$. Since ${A_{\mathcal{E}}}^n$ permutes the ends of M_1, by taking even larger n we may assume that $[{A_{\mathcal{E}}}^n(\alpha_i)] = [\alpha_i]$. By Lemma 4.1 we have that each α_i is homotopic to a curve that is contained in $R \subset M$.

Let $M_2 \subset M$ be a proper subsurface of M that is homotopic to M_1 and so that its frontier components are simple closed curves (that are necessary in the corresponding homotopy classes of the curves α_i). We can also assume that all the boundary curves of M_2 are contained in R. Since the lift of M_1 to $\widetilde{M}$ is unbounded we have that the lift of M_2 to $\widetilde{M}$ is unbounded too. This implies that M_2 contains the one holed torus T. Therefore, there is a simple closed curve $\gamma \subset M$ which is homotopic to the annulus N, so that M_2 contains the torus minus the disc which is the corresponding component of $M \setminus \gamma$. $\square$

Remark. The previous proposition shows that the set of acyclic components of $\mathbf{S}$ is substantial. However, we do not address here the interesting questions like: What are the non-acyclic components of $\mathbf{S}$? How many non-acyclic components $\mathbf{S}$ has? Does $\mathbf{S}$ has at least one non-acyclic component?

Recall that $A_{\mathcal{E}} : M \to M$ denotes a homeomorphism homotopic to A_M. Let $\widehat{f} : M \to M$ be a homeomorphism that commutes with $A_{\mathcal{E}}$, and such that $\widehat{f}$ is homotopic to a homomorphism $\widehat{f_1} : M \to M$ that is the identity outside the set R' (note that this implies that $\widehat{f_1}$ commutes with A_M). Let $\widetilde{F}$ denote the lift of $\widehat{f}$ to $\widetilde{M}$ that commutes with $\widetilde{A_{\mathcal{E}}}$ (such a lift is unique). This particular lift can be constructed by lifting any isotopy between $\widehat{f}$ and $\widehat{f_1}$ from M to $\widetilde{M}$. One end of the lifted isotopy (the one that is the lift of $\widehat{f_1}$) is the identity outside every $R_{m,n}$, $m, n \in \mathbf{Z}$. The other end of the lifted isotopy is $\widetilde{F}$. This also shows that there is a positive number $c(\widetilde{F})$ such that

$$\widetilde{\mathbf{d}}(x, \widetilde{F}(x)) < c(\widetilde{F}),$$

for every $x \in \widetilde{M}$ because the similar is true for the corresponding lift of $\widehat{f_1}$.

Lemma 4.12. *For each $a \in \mathbf{R}$, $\widetilde{F}$ preserves setwise each of the sets T_a and T'_a. Let $a < b < c$. Then $\widetilde{F}$ preserves setwise each component of the complement of the set $T_a \cup T_b \cup T_c$ (similarly for the set $T'_a \cup T'_b \cup T'_c$).*

Proof. Suppose, by contradiction, that for some $x \in T_{a_1}$, $\widetilde{F}(x) \in T_{b_1}$, with $a_1 \neq b_1$. Suppose $a_1 < b_1$. We can find $a_1 < a < b < b_1$, such that x is to the left of Q_a, and such that $\widetilde{F}(x)$ is to the right of Q_b. Now ${\widetilde{A_{\mathcal{E}}}}^n(Q_a) \subset P_{a,n}$ and ${\widetilde{A_{\mathcal{E}}}}^n(Q_b) \subset P_{b,n}$. Therefore, the distance between ${\widetilde{A_{\mathcal{E}}}}^n(Q_a)$ and ${\widetilde{A_{\mathcal{E}}}}^n(Q_b)$ increases exponentially with n. Since $\widetilde{F}$ commutes with $\widetilde{A_{\mathcal{E}}}$ we have

$$({\widetilde{A_{\mathcal{E}}}}^n \circ \widetilde{F})(x) = (\widetilde{F} \circ {\widetilde{A_{\mathcal{E}}}}^n)(x).$$

The left-hand side is to the right of $\widetilde{A_{\mathcal{E}}}^n(Q_b)$, and the right-hand side is within $c(\widetilde{F})$ hyperbolic distance of the space that is to the left of $\widetilde{A_{\mathcal{E}}}^n(Q_a)$. For large n this is impossible, and the result is proved.

It is clear that $\widetilde{F}$ maps every component of the complement of $T_a \cup T_b \cup T_c$ to another such component. For each such component choose $d \in \mathbf{R}$, such that this component contains T_d. Then the other components do not intersect with T_d. But $\widetilde{F}$ preserves T_d setwise. This proves the second part of the lemma. □

The previous lemma implies that $\widetilde{F}$ preserves setwise each set $\widetilde{S}_{x,y}$, $x, y \in \mathbf{R}$.

4.4. **The homomorphism $\mathcal{E}$.** From this point until the end of the paper we assume that there exists a homomorphism

$$\mathcal{E} : \mathcal{MC}(M) \to \mathcal{H}omeo(M),$$

such that the composite homomorphism $Pr \circ \mathcal{E}$ is the identity.

Let $A_{\mathcal{E}} : M \to M$ be the homeomorphism given by $\mathcal{E}([A_M]) = A_{\mathcal{E}}$. We have that $A_{\mathcal{E}}$ is homotopic to A_M and as before let $\mathbf{S}$ be the decomposition of M that corresponds to $A_{\mathcal{E}}$. In Proposition 4.1 we have seen that the set $M_{\mathbf{S}}$ of acyclic components of $\mathbf{S}$ contains a torus minus a disc where the frontier component of this torus is a simple closed curve homotopic to the annulus N. Under the newly introduced assumption that the homeomorphism $\mathcal{E}$ exits we can show the following.

Lemma 4.13. *Let $M_{\mathbf{S}}$ be the set of acyclic components of the decomposition $\mathbf{S}$. There exists one and only one component of $M_{\mathbf{S}}$ that is not simply connected. This component is a proper subsurface of M and it has only one end and this end is homotopic to the annulus N. In particular, the surface V (from the statement of Lemma 4.11) is a planar subsurface of $\widetilde{M}$.*

Proof. We assume at first that $M_{\mathbf{S}}$ is not equal to M. Then every component of $M_{\mathbf{S}}$ has boundary.

Recall that $R' \subset M = T \cup N \cup R'$ is a surface of genus $g-1$ with one end ($g \geq 2$ is the genus of M). We also used the notation $R = N \cup R'$. Let $\widehat{g}$ be a pseudo-Anosov homeomorphism of a closed surface of genus $g-1$ onto itself (or just Anosov if $g = 2$) which has a fixed point. By blowing up the fixed point (in the same manner as we did at the beginning of this section in connection with the Anosov map $A : \mathbf{T}^2 \to \mathbf{T}^2$) one obtains a pseudo-Anosov (or just Anosov) map $\widehat{f} : R \to R$ with the following properties. The map $\widehat{f}$ is the identity on the boundary circle of R. If we collapse the boundary circle of R onto a point to obtain a closed surface of genus $g-1$, then the map $\widehat{f}$ is well defined on this surface and it is equal to $\widehat{g}$. By extending $\widehat{f}$ to be the identity on T we obtain the map $\widehat{f} : M \to M$. Let $F : M \to M$ be given by $F = \mathcal{E}([\widehat{f}])$.

In particular, the map F commutes with $A_{\mathcal{E}}$. Let $\widetilde{F} : \widetilde{M} \to \widetilde{M}$ be the corresponding lift. From Lemma 4.12 we have that $\widetilde{F}$ setwise preserves every $S_{x,y}$. This implies that F maps every element of $\mathbf{S}$ onto another element of $\mathbf{S}$. This implies that $F(M_{\mathbf{S}}) = M_{\mathbf{S}}$. Let M_1 be a component of $M_{\mathbf{S}}$ that contains a torus minus a disc which is homotopic to $T \subset M$. We know by Proposition 2.1 that M_1 is a proper subsurface of M with finitely many ends. Fix an end of M_1. Let α be a simple closed curve so that this end is homotopic to α. From Lemma 4.1 it follows that one can choose α to be a subset of R'. Since M_1 has finitely many ends, there exists $n \in \mathbf{N}$ so that $[F^n(\alpha)] = [\alpha]$. If α is not homotopic to N then after collapsing the boundary circle of R to obtain a closed surface of genus $g-1$, the curve α becomes a simple closed curve that is homotopically non-trivial. Also, we get that the pseudo-Anosov (or just Anosov) map $(\widehat{g})^n$ fixes a homotopy class of α which is not trivial. This is of course impossible so the conclusion is that α is homotopic to N.

Let M_2 be an arbitrary component of $M_{\mathbf{S}}$ that is not simply connected. Then the same argument as above shows that the only end of M_2 is homotopic to N. This yields that M_2 is a surface of genus $g-1$ with one end, and this end is homotopic to N. Assume that M_2 is not the same as M_1. Then, there exists an annulus $N' \subset M$ which is homotopic to N and so that every component of $\mathbf{S}$ that is not acyclic is contained in N'. Let $S \in \mathbf{S}$ be non-acyclic (such a component exists since we assumed that $M_{\mathbf{S}} \neq M$). Then $S \subset N'$. If S does not separate the two ends of N', then we have that $N' \setminus S$ is conformally equivalent to an annulus minus a point or to an annulus minus a disc. Since S does not separate M this implies that S is acyclic which is impossible. We conclude that S does separate the two ends of N'. But then S separates M as well which is a contradiction. This shows that $M_2 = M_1$.

It remains to show that $M_{\mathbf{S}}$ is not equal to M. The proof is by contradiction. Assume that $M_{\mathbf{S}} = M$. Let $x, y \in \mathbf{R}$. Let a_n and b_n (c_n and d_n) be two sequences of real numbers such that $a_n < x < b_n$ ($c_n < y < d_n$) for each $n \in \mathbf{N}$, and $a_n, b_n \to x$ ($c_n, d_n \to y$), $n \to \infty$. Let Ω_n (Ω'_n) be the component of $\widetilde{M} \setminus (T_{a_n} \cup T_{b_n})$ ($\widetilde{M} \setminus (T'_{c_n} \cup T'_{d_n})$) that is in between T_{a_n} and T_{b_n} (T'_{c_n} and T'_{d_n}). We have that Ω_n is an open set. From Lemma 4.8 we have

$$\widetilde{S}_{x,y} = \bigcap_{n \in \mathbf{N}} (\Omega_n \cap \Omega'_n).$$

Since the restriction of the covering map from $\widetilde{M}$ to M on $\widetilde{S}_{x,y}$ is injective, we have that for n large enough the corresponding covering map is injective on $\Omega_n \cap \Omega'_n$. By the same argument that shows that no component of $\widetilde{S}_{x,y}$ separates $\widetilde{M}$ one shows that no component of $\Omega_n \cap \Omega'_n$ separates $\widetilde{M}$. Then, from Lemma 2.2 (since no component of $\Omega_n \cap \Omega'_n$ separates $\widetilde{M}$) we have that for n large enough every component of $\Omega_n \cap \Omega'_n$ is simply connected. We choose n to be large enough.

Let U be a component of $\Omega_n \cap \Omega'_n$. We show that there exists $k \in \mathbf{N}$ so that $\widetilde{F}^k(U) = U$. On the contrary, assume that the sets $\widetilde{F}^i(U)$, $i \in \mathbf{N}$, are disjoint. The boundary of the set U is a subset of $T_{a_n} \cup T_{b_n} \cup T'_{c_n} \cup T'_{d_n}$. Let $p \in U$, and let r, s be such that $p \in T_r \cap T'_s$. Here clearly $a_n \neq r \neq b_n$ and $c_n \neq s \neq d_n$. So, we have a sequence of disjoint, open sets, that are all contained in some compact subset of the surface $\widetilde{M}$. Clearly, every accumulation point $q \in \widetilde{M}$ of the sequence $\widetilde{F}^i(p)$, $i \in \mathbf{N}$, must also be an accumulation point of the sequence of sets $\widetilde{F}^i(T_{a_n} \cup T_{b_n} \cup T'_{c_n} \cup T'_{d_n})$, $i \in \mathbf{N}$, because $\widetilde{F}^k(T_{a_n} \cup T_{b_n} \cup T'_{c_n} \cup T'_{d_n})$ contains the boundary of the subsurface $\widetilde{F}^k(U)$. Then q must be in $T_{a_n} \cup T_{b_n} \cup T'_{c_n} \cup T'_{d_n}$ since this set is closed and preserved setwise by $\widetilde{F}$. But for the same reasons q is also in $T_r \cap T'_s$. This is a contradiction since $a_n \neq r \neq b_n$ and $c_n \neq s \neq d_n$, and we conclude that there exists $k \in \mathbf{N}$ so that $\widetilde{F}^k(U) = U$.

By projecting the sets $\Omega_n \cap \Omega'_n$ to M we have that for each element of $S \in \mathbf{S}$ there is a decreasing sequence of simply connected sets U_n so that $S = \bigcap_{n\in\mathbf{N}} U_n$. Also, for each U_n there is $k_n \in \mathbf{N}$ so that $F^{k_n}(U_n) = U_n$. From Lemma 2.3 we conclude that F is homotopic to a periodic map which is a contradiction. So, we have proved that $M \neq M_{\mathbf{S}}$.

Finally, it follows from what we have already proved that any lift of any component of $M_{\mathbf{S}}$ to $\widetilde{M}$ is a planar surface. Since V is a subsurface of $\widetilde{M}$, and the projection of V to M is contained in $M_{\mathbf{S}}$, we have that V is a planar surface.

□

Lemma 4.14. *Let $\widehat{f} : M \to M$ be a homeomorphism that is the identity on $T \subset M$. Set $F = \mathcal{E}([\widehat{f}])$. Let $\widetilde{S}$ be a component of $\widetilde{S}_{x,y}$, for some $x, y \in \mathbf{R}$. Then $\widetilde{F}$ preserves setwise the set $\widetilde{S}$.*

Proof. We first prove the lemma under the assumption that $x, y \in \mathbf{R}$ are such so that the sets $\widetilde{A_{\mathcal{E}}}^m(\widetilde{S}_{x,y})$ and $\widetilde{A_{\mathcal{E}}}^n(\widetilde{S}_{x,y})$ are not equivalent when $m \neq n$ and $1 \leq m, n \leq 3g-2$, where g is the genus of the surface M. Recall from Lemma 4.11 that we denote the union of all such $\widetilde{S}_{x,y}$ by V. In this case we know by Lemma 4.11 that every component of the set $\widetilde{S}_{x,y}$ is acyclic.

Let a_n and b_n (c_n and d_n) be two sequences of real numbers such that $a_n < x < b_n$ ($c_n < y < d_n$) for each $n \in \mathbf{N}$, and $a_n, b_n \to x$ ($c_n, d_n \to y$), $n \to \infty$. Let Ω_n (Ω'_n) be the component of $\widetilde{M} \setminus (T_{a_n} \cup T_{b_n})$ ($\widetilde{M} \setminus (T'_{c_n} \cup T'_{d_n})$) that is in between T_{a_n} and T_{b_n} (T'_{c_n} and T'_{d_n}). From Lemma 4.8 we have

$$\widetilde{S}_{x,y} = \bigcap_{n\in\mathbf{N}} (\Omega_n \cap \Omega'_n).$$

In order to show that $\widetilde{F}$ preserves $\widetilde{S}$ setwise, it is enough to show that $\widetilde{F}$ preserves each component of the set $(\Omega_n \cap \Omega'_n)$, for large n. We will prove it for n large enough.

Since every component of the set $\widetilde{S}_{x,y}$ is acyclic by Lemma 4.11 we can choose $n \in \mathbf{N}$ large enough so that every component of $\Omega_n \cap \Omega'_n$ is an acyclic subset of $\widetilde{M}$, and therefore simply connected. Let U be any component of $(\Omega_n \cap \Omega'_n)$. We first show that either $\widetilde{F}(U) = U$, or all the sets in the sequence $\widetilde{F}^k(U)$, $k \in \mathbf{N}$, are mutually disjoint.

Set $D_n = \Omega_n \cap V$ and $D'_n = \Omega'_n \cap V$. We have

$$\widetilde{S}_{x,y} = \bigcap_{n\in\mathbf{N}} (D_n \cap D'_n).$$

Let $\Omega^R_{a_n}$ be the subsurface of $\widetilde{M}$ that is to the right of T_{a_n}, and set $H = \Omega^R_{a_n} \cap V$ (also, let $\Omega^L_{b_n}$ we denote the subsurface of $\widetilde{M}$ that is to the left of T_{b_n}). Then H is a planar surface and $\widetilde{F}$ maps H onto itself (because $\widetilde{F}$ preserves every $\widetilde{S}_{x,y}$ setwise). In particular, $\widetilde{F}$ maps the frontier component of H that is contained in T_{a_n} onto itself. Note that the frontier component of H that is contained in T_{a_n} is the only unbounded frontier component of H.

Note that H contains $(D_n \cap D'_n)$. Since the set V is a planar domain, by Proposition 2.2 and Proposition 2.3 there is a conformal map $\phi : H \to H_1$ with the following properties. The conformal map ϕ maps H onto a Riemann surface H_1 that is contained in the upper half plane. Also, the frontier component of H, that is contained in T_{a_n}, under the map ϕ corresponds to the real line $\mathbf{R}$ which is the boundary of the half plane. Let $\widetilde{F}_1 = \phi \circ \widetilde{F} \circ \phi^{-1}$. Since ϕ is conformal, by Proposition 2.3 we conclude that $\widetilde{F}_1$ extends continuously to a homeomorphism of the real line $\mathbf{R}$.

Next, we show that the intersection between the closure of U and the frontier component of either $\Omega^R_{a_n}$ or $\Omega^L_{b_n}$ is non-empty and contains a relatively open subset of the corresponding frontier component. To see this let $\overline{\Omega_n}$ be the closure of Ω_n in $\widetilde{M}$. This set is contained in the union $\Omega_n \cup T_{a_n} \cup T_{b_n}$. Every component of the intersection between Ω'_n and $\overline{\Omega_n}$ is a relatively open subset of $\overline{\Omega_n}$ (because Ω'_n is open), and the intersection between every such component with at least one the frontier components contains a relatively open subset of that component (otherwise, one would have that T_{a_n} and T_{b_n} have a point in common). On the other hand, each component of the set $\Omega'_n \cap \overline{\Omega_n}$ contains the appropriate component of $(\Omega_n \cap \Omega'_n)$ as a dense subset.

Now we can assume that the corresponding frontier component is the frontier component of $\Omega^R_{a_n}$ and that the intersection between the closure of U and the unbounded frontier component of H contains a relatively open set. Let $O = \phi(U)$. It follows that $\partial O \cap \mathbf{R}$ is non-empty. Since the closure of U contains a relatively open subset of the frontier components of H that is contained in T_{a_n}, we have that $\partial O \cap \mathbf{R}$ is a compact subset of $\mathbf{R}$ which contains at least one interval. Let $U_1 = \widetilde{F}(U)$, and set $\phi(U_1) = O_1 = \widetilde{F}_1(O)$. Assume that U_1 is not the same as U. Then they are disjoint open subsets of H. This implies that O and O_1 are disjoint (note that this does not necessarily imply that $\partial O \cap \mathbf{R}$ and $\partial O_1 \cap \mathbf{R}$ are disjoint). However, if I is an open interval, $I \subset (\partial O \cap \mathbf{R})$, then clearly I has empty intersection with $\partial O_1 \cap \mathbf{R}$ since for every point in I there exists a relative neighborhood (in $\overline{H_1}$) of this point that is contained in O.

Assume now that for some $k \geq 2$ we have that $\widetilde{F}^k(U) = U$ (this implies $\widetilde{F}^k_1(O) = O$ and $\widetilde{F}^k_1(O_1) = O_1$). Then the intervals I, $\widetilde{F}_1(I)$, $\widetilde{F}^k_1(I)$, $\widetilde{F}^{k+1}_1(I)$, are mutually disjoint and in this order on $\mathbf{R}$. Let $\alpha \subset O$ be a curve that connects I and $\widetilde{F}^k_1(I)$, and let $\beta \subset V_1$ be a curve that connects $\widetilde{F}_1(I)$ and $\widetilde{F}^{k+1}_1(I)$. Since H_1 is contained in the upper half plane we have that α and β must intersect. This is a contradiction since O and O_1 are disjoint. This shows that either $\widetilde{F}(U) = U$, or any two sets in the sequence $\widetilde{F}^k(U)$, $k \in \mathbf{N}$, are disjoint.

We consider the case when any two sets in the sequence $\widetilde{F}^k(U)$, $k \in \mathbf{N}$, are disjoint (note that $n > n_0$ is fixed). The boundary of the set U is a subset of $T_{a_n} \cup T_{b_n} \cup T'_{c_n} \cup T'_{d_n}$. Let $p \in U$, and let r, s be such that $p \in T_r \cap T'_s$. Here clearly $a_n \neq r \neq b_n$ and $c_n \neq s \neq d_n$. We have that each $\widetilde{F}^k(U)$, $k \in \mathbf{N}$, is a planar surface (since it is a subset of a planar surface). So, we have a sequence of disjoint surfaces, that are all contained in some compact subset of the surface $\widetilde{M}$. Clearly, every accumulation point $q \in \widetilde{M}$ of the sequence $\widetilde{F}^k(p)$, $k \in \mathbf{N}$, must also be an accumulation point of the sequence of sets $\widetilde{F}^k(T_{a_n} \cup T_{b_n} \cup T'_{c_n} \cup T'_{d_n})$, $k \in \mathbf{N}$, (because $\widetilde{F}^k(T_{a_n} \cup T_{b_n} \cup T'_{c_n} \cup T'_{d_n})$ is the boundary of the subsurface $\widetilde{F}^k(U)$). Then

q must be in $T_{a_n} \cup T_{b_n} \cup T'_{c_n} \cup T'_{d_n}$ since this set is closed and preserved setwise by $\widetilde{F}$. But for the same reasons q is also in $T_r \cap T'_s$. This is a contradiction.

Let $a, b, r, s \in \mathbf{R}$, $a < b$ and $c < d$. Let Ω (Ω')be the component of $\widetilde{M} \setminus (T_a \cup T_b)$ ($\widetilde{M} \setminus (T'_c \cup T'_d)$) that is in between T_a and T_b (T'_c and T'_d). Let U be a component of the intersection $\Omega \cap \Omega'$. We have that U contains at least one component from uncountably many different sets $\widetilde{S}_{x,y}$. On the other hand, we proved above that $\widetilde{F}$ setwise preserves every component of all but countably many sets $\widetilde{S}_{x,y}$. Since $\widetilde{F}(U)$ is either equal to U or disjoint with U, we conclude that $\widetilde{F}$ setwise preserves U. This shows that $\widetilde{F}$ setwise preserves every component of every set $\widetilde{S}_{x,y}$.

□

Remark. It also follows from the argument in the previous paragraph that if S is a component of $\mathbf{S}$ then every point in the frontier component of the subsurface $M \setminus S$ is a limit point of acyclic components of $\mathbf{S}$.

The summary of the results we have proved in this section is as follows.

Theorem 4.1. *Let M be a surface of genus $\mathbf{g} \geq 2$, and $\alpha \subset M$ a simple closed curve, such that α separates M into a torus minus a disc and a compact surface of genus $\mathbf{g} - 1$ with one end. Assume that $f \in \mathcal{MC}(M)$ can be realized by a homeomorphism that is the identity inside the corresponding subsurface of M that is homeomorphic to a torus minus a disc. Then there exists an admissible decomposition of M for $\mathcal{E}(f)$ with the following property. There exists a simple closed curve β, homotopic to α, such that if $p \in M$ belongs to the torus minus a disc (which is one of the two components obtained after removing β from M), then the component of the decomposition that contains p is acyclic.*

Proof. As above, let $\mathcal{E}([A_M]) = A_{\mathcal{E}}$, and let $\mathbf{S}$ be the corresponding decomposition of M. We have already observed that $\mathbf{S}$ is an upper semi-continuous decomposition of M. The density of acyclic components was shown in the proof of Lemma 4.14 (see the previous remark). Since $\mathcal{E}([f]) = F$ commutes with $A_{\mathcal{E}}$ (and f can be realized to be the identity outside the torus minus the disc) we can apply the above proposition to conclude that $\mathcal{E}([f])$ setwise preserves each component of $\mathbf{S}$. This shows that $\mathbf{S}$ is an admissible decomposition for $\mathcal{E}([f])$.

□

5. The structure of homeomorphisms from $\mathcal{E}(\mathcal{MC}(M))$

5.1. **The characteristic annulus of a twist t_α.** Let M be a closed surface of genus $g \geq 2$, and $\alpha \subset M$ a simple, closed, non-separating curve. Let $t_\alpha \in \mathcal{MC}(M)$ be the standard twist which can be realized by a homeomorphism of M that is supported in an annulus homotopic to α. Denote by $\mathbf{S}_\alpha$ the minimal decomposition of M associated to $\mathcal{E}(t_\alpha)$ (see Definition 3.2). By $M_{\mathbf{S}_\alpha} = M_\alpha \subset M$ we denote the open set of points that are contained in acyclic components of $\mathbf{S}_\alpha$.

Let $f \in \mathcal{MC}(M)$, and let $\beta \subset M$ be a simple closed curve such that $[\beta] = f^*[\alpha]$. We have $f \circ t_\alpha = t_\beta \circ f$. Let $\mathbf{S}_\beta$ be the minimal decomposition that corresponds to $\mathcal{E}(t_\beta)$. Then, $\mathcal{E}(f)(\mathbf{S}_\alpha)$ is an admissible decomposition that corresponds to $\mathcal{E}(t_\beta)$. Since both $\mathbf{S}_\alpha$ and $\mathbf{S}_\beta$ are the minimal decompositions, we conclude that $\mathcal{E}(f)(\mathbf{S}_\alpha) = \mathbf{S}_\beta$.

Lemma 5.1. *M_α contains one and only one component A_α which is non-planar. Moreover, A_α is a proper subsurface of M. It has exactly two ends, and both have the homotopy type of the curve α.*

Remark. Our proof does not exclude the possibility that the two frontier components of A_α can meet (or even coincide) as closed subsets of M. By simple analysis one can show that this is not possible, and in fact the interior of the set $M \setminus A_\alpha$ contains an annulus homotopic to α. But we do not need this fact.

Proof. Let $\gamma \subset M$ be a closed curve which separates M into a torus minus a disc, and a compact surface of genus $\mathbf{g}-1$ with one end. Assume that α belongs to the second surface and $\alpha \cap \gamma = \emptyset$. By Theorem 4.1 there is curve $\gamma_1 \subset M$, homotopic to γ, such that the torus minus a disc that is obtained by removing γ_1 is contained in M_α (we can also choose γ_1 so that $\gamma_1 \subset M_\alpha$). Let A_α be the component of M_α that contains this torus (and γ_1). Note that this implies that A_α is not simply connected.

From Proposition 2.1 we know that A_α is a proper subsurface of M. Assume that $M_\alpha \neq M$. This implies that the set $M \setminus A_\alpha$ is not homotopically trivial, because if it was that would imply that $M \setminus A_\alpha$ is a subset of M_α, which would mean that $M_\alpha = M$. So, A_α has at least one end. The corresponding frontier component we denote by K. We have that K is a closed subset of M, and K belongs to the closure of A_α, but no point from K belongs to A_α. Therefore, no point from K belongs to M_α (here we use that M_α is an open set and that A_α is a connected component of it). Let $\beta_n \subset A_\alpha$, $n \in \mathbf{N}$, be a nested sequence of simple closed curves that determine the end K. These curves are all in the same homotopy class. β_n converges to K in the Hausdorff metric. Then β_n, $n \in \mathbf{N}$, is not homotopically trivial because both A_α, and $M \setminus A_\alpha$ are not homotopically trivial (this also follows from the fact that A_α is a proper subsurface which is not simply connected, and therefore by definition, A_α can not have any ends so that the corresponding curves β_n are homotopically trivial). Note that the intersection number between $[\beta_n]$ and $[\gamma]$ must be zero. If it was not zero, we would have that γ_1 intersects every β_n, and therefore it would have to intersect K. This is impossible, since no point from K is in M_α.

Before we prove that every β_n is homotopic to α we need the following simple proposition.

Proposition 5.1. *Let η_i, $i = 1, 2, 3$, be simple closed curves in M, such that the intersection numbers between $[\eta_2]$ and $[\eta_1]$, and, $[\eta_2]$ and $[\eta_3]$, are both not zero. On the other hand, assume that the intersection number between $[\eta_1]$ and $[\eta_3]$ is zero. Then, the intersection number between $t^*_{\eta_2}[\eta_1]$ and $[\eta_3]$ is not zero.*

Proof. Let l_1, l_2, l_3, denote the lifts (to the universal cover $\mathbf{D}$) of the simple closed geodesics that correspond to $[\eta_1]$, $[\eta_2]$, $[\eta_3]$, respectively. We choose these lifts such that l_1 and l_3 intersect l_2, and such that the hyperbolic distance between the points $l_1 \cap l_2$ and $l_3 \cap l_2$, is less than the length of the closed geodesic which lifts to the geodesic in the class $[\eta_2]$. Since l_1 and l_3 are disjoint, it readily follows that the geodesic in $\mathbf{D}$, with the same endpoints as any curve in the class $[t^*_{l_2} l_1]$, has to intersect l_3. $\square$

By contradiction, assume that β_n is not homotopic to α. Therefore, we can find a simple closed curve $\delta \subset M$, such that the intersection number between $[\delta]$ and

$[\beta_n]$ is not zero, and such that $\delta \cap \alpha = \emptyset$. Moreover, since $\gamma \cap \alpha = \emptyset$, we can assume that the intersection number between $[\delta]$ and $[\gamma]$ is not zero either. Let $f = t_\delta$. Then, that the intersection number between the homotopy classes $f^*[\gamma]$ and $[\beta_n]$ is not zero.

We conclude that $\mathcal{E}(f)(\gamma_1)$ intersects each β_n, and then it has to intersect K. On the other hand, f commutes with t_α, and so $\mathcal{E}(f)(\mathbf{S}_\alpha) = \mathbf{S}_\alpha$. This yields a contradiction, since no point from K is in the set M_α. So, β_n (and the corresponding end of A_α) must be homotopic to α.

The same conclusion holds for any other end of A_α. It now follows that A_α satisfies the conclusion of the lemma, or we have $A_\alpha = M_\alpha = M$. If $M_\alpha = M$, then by Lemma 3.1 $\mathcal{E}(t_\alpha)$ would be homotopic to the identity map, which is a contradiction.

Choose two disjoint curves β_n and β_n', where each of them corresponds to one of the two ends of A_α. Then both β_n and β_n' are homotopic to α, and therefore every other component of M_α must be in the annulus bounded by these two curves. Therefore, every such component must be planar. $\square$

For a non-separating simple closed curve $\alpha \subset M$, let

$$B_\alpha = M \setminus A_\alpha. \tag{5.1}$$

Clearly B_α is a closed, connected subset of M (note that we do not claim that $B_\alpha \cap M_\alpha = \emptyset$). Also, $\mathcal{E}(t_\alpha)(B_\alpha) = B_\alpha$. Since A_α is connected we have that B_α does not separate M.

Lemma 5.2. *Let α and α_1 be two disjoint, non-separating, simple closed curves on M, such that every component of $M \setminus (\alpha \cup \alpha_1)$ is of genus at least one. Then B_α and B_{α_1} are disjoint. Also, $\mathcal{E}(t_\alpha)$ setwise preserves B_{α_1}.*

Proof. Since homeomorphisms $\mathcal{E}(t_\alpha)$ and $\mathcal{E}(t_{\alpha_1})$ commute (note $t^*_{\alpha_1}\alpha = \alpha$), we have that $\mathcal{E}(t_{\alpha_1})(\mathbf{S}_\alpha) = \mathbf{S}_\alpha$. That $\mathcal{E}(t_\alpha)$ setwise preserves B_{α_1} follows from the fact that $\mathbf{S}_{\alpha_1}$ is the minimal decomposition for $\mathcal{E}(t_{\alpha_1})$, and because we have that $\mathcal{E}(t_\alpha)$ commutes with $\mathcal{E}(t_{\alpha_1})$. From Lemma 5.1 we have that $\mathcal{E}(t_\alpha)$ setwise preserves A_{α_1} (because A_{α_1} is the unique component with the given properties), and it therefore has to setwise preserve the complement of A_{α_1}.

Denote by $G = G(\alpha, \alpha_1)$ the subgroup of $\mathcal{H}omeo(M)$ generated by $\mathcal{E}(t_\alpha)$ and $\mathcal{E}(t_{\alpha_1})$. This group is a two-generator free Abelian group. Let $\mathbf{S}_G$ denote the minimal decomposition of M for the group G. By $M_G \subset M$ we denote the open set of points that are contained in acyclic components of $\mathbf{S}_G$. Since M_α and M_{α_1} are the minimal decomposition we have $M_G \subset (M_\alpha \cap M_{\alpha_1})$. Combining this with the fact that every component of $M \setminus (\alpha \cup \alpha_1)$ is of genus at least one, and by repeating the argument from the proof of Lemma 5.1, we get the following. To every component of the set $M \setminus (\alpha \cup \alpha_1)$, there corresponds a homotopically non-trivial component of the set M_G. Every component of M_G is a proper subsurface of M. Moreover, every end, of every homotopically non-trivial component of M_G, has the homotopy type either of α, or of α_1. This shows that the complement of the set M_G (this complement is a closed set) is contained in the union of two disjoint (open) annuli around α and α_1 respectively. This implies that B_α and B_{α_1} are disjoint. $\square$

Let $\beta_i \subset A_\alpha$, $i = 1, 2$, be simple closed curves that are homotopic to the two ends of A_α respectively. Each curve is homotopic to the corresponding frontier component of A_α. Therefore, β_i, $i = 1, 2$, are both homotopic to α. The open set $A_\alpha \setminus (\beta_1 \cup \beta_2)$ has three components. We denote the first one by $A_\alpha(\beta_1, \beta_2)$. This component has the same homotopy type as A_α. The other two components are annuli, each of them bounded by the corresponding curve β_i and by the corresponding frontier component of A_α. Let

$$\overline{A_\alpha(\beta_1, \beta_2)} = A_\alpha(\beta_1, \beta_2) \cup \beta_1 \cup \beta_2$$

denote the corresponding closure. Denote by $C_\alpha(\beta_1, \beta_2)$ the union of all elements of the decomposition $\mathbf{S}_\alpha$ that do intersect $\overline{A_\alpha(\beta_1, \beta_2)}$. Since $\overline{A_\alpha(\beta_1, \beta_2)}$ is a connected, compact subset of A_α, we have that $C_\alpha(\beta_1, \beta_2)$ is a compact, connected subset of A_α. Also,

$$\mathcal{E}(t_\alpha)(C_\alpha(\beta_1, \beta_2)) = C_\alpha(\beta_1, \beta_2),$$

since $C_\alpha(\beta_1, \beta_2)$ is a union of sets that are all setwise preserved by $\mathcal{E}(t_\alpha)$. In particular, there is a neighborhood of the set B_α which does not intersect $C_\alpha(\beta_1, \beta_2)$.

Let $D_\alpha(\beta_1, \beta_2)$ be the component of the set $M \setminus C_\alpha(\beta_1, \beta_2)$ that contains B_α. So, $D_\alpha(\beta_1, \beta_2)$ is an open, connected set which does not separate M (because $C_\alpha(\beta_1, \beta_2)$ is connected). Also, $D_\alpha(\beta_1, \beta_2)$ is contained in $M \setminus A_\alpha(\beta_1, \beta_2)$. Therefore $D_\alpha(\beta_1, \beta_2)$ is an (open) annulus homotopic to α. Note that

$$\mathcal{E}(t_\alpha)(D_\alpha(\beta_1, \beta_2)) = D_\alpha(\beta_1, \beta_2).$$

Remark. If β_1 and β_2 approach the corresponding frontier components of A_α (recall that these frontier components are contained in B_α) then the set $D_\alpha(\beta_1, \beta_2)$ converges in the Hausdorff metric to the set B_α. In addition, one can show that $D_\alpha(\beta_1, \beta_2)$ is the only component of the set $M \setminus C_\alpha(\beta_1, \beta_2)$. We do not need this last fact though.

The annulus $D_\alpha(\beta_1, \beta_2)$ has two frontier components we denote by K_1 and K_2. Note that $K_1, K_2 \subset C_\alpha(\beta_1, \beta_2)$. We have that K_1 and K_2 belong respectively to the closures of the two annuli that are components of $A_\alpha \setminus (\beta_1 \cup \beta_2)$. Therefore $K_1 \cap K_2 = \emptyset$. Moreover, each K_i, $i = 1, 2$, is compactly contained in A_α.

Definition 5.1. *Let $\alpha \subset M$ be a simple, closed, non-separating curve. For given β_1 and β_2 (introduced above) we say that $D_\alpha(\beta_1, \beta_2)$ is the characteristic annulus that corresponds to the curves α, β_1 and β_2.*

Let $f \in \mathcal{MC}(M)$ and let $\alpha_1 \subset M$ be a simple closed curve such that $[\alpha_1] = f^*[\alpha]$. We have $f \circ t_\alpha = t_{\alpha_1} \circ f$. Then $\mathcal{E}(f)(D_\alpha(\beta_1, \beta_2)) = D_{\alpha_1}(\mathcal{E}(f)(\beta_1), \mathcal{E}(f)(\beta_2))$. The set $D_{\alpha_1}(\mathcal{E}(f)(\beta_1), \mathcal{E}(f)(\beta_2))$ is a characteristic annulus that corresponds to the curves α_1, $\mathcal{E}(f)(\beta_1)$, $\mathcal{E}(f)(\beta_2)$.

Let $\psi : D_\alpha(\beta_1, \beta_2) \to N$ be a conformal map. Here $N = \{1 < |z| < r\}$ is the geometric annulus, where $r > 1$ is the appropriate number that depends on the conformal modulus of $D_\alpha(\beta_1, \beta_2)$. Set $E_\alpha = \psi \circ \mathcal{E}(t_\alpha) \circ \psi^{-1}$. Since $\mathcal{E}(t_\alpha)$ is a homeomorphism of M we conclude that E_α extends to a homeomorphism of the closed annulus $\overline{N}$. Moreover, E_α preserves each boundary circle of N.

Let F_α denote the set of all fixed points of E_α that belong to the boundary circles of the annulus N. By R_α we denote the standard twist homeomorphism (which is in fact a diffeomorphism) on N.

Lemma 5.3. *The set F_α contains at least one point from either of the two boundary circles of N. Moreover, the maps R_α and E_α are homotopic on $\overline{N}$ modulo the set $F_\alpha \cap \partial N$.*

Proof. Since R_α is the identity on the boundary circles of N we have that

$$\widetilde{R}_\alpha = \psi^{-1} \circ R_\alpha \circ \psi$$

is a homeomorphism of the closure of the set $D_\alpha(\beta_1, \beta_2)$. We define $\widetilde{R}_\alpha$ to be the identity on the rest of the surface M. So, $\widetilde{R}_\alpha$ is a homeomorphism that is in the homotopy class of the twist t_α, and it is therefore isotopic to $\mathcal{E}(t_\alpha)$.

Recall that K_1 and K_2 are the frontier components of $D_\alpha(\beta_1, \beta_2)$. Let $p, q \in K_1$ are chosen so that they belong to different components of the decomposition $\mathbf{S}_\alpha$. Let $p_n, q_n \in D_\alpha(\beta_1, \beta_2) \cap A_\alpha$, $n \in \mathbf{N}$, be two sequences of points that converge to the points p and q respectively. Denote by S_n and S'_n the corresponding elements of $\mathbf{S}_\alpha$ that contain the points p_n and q_n respectively. It follows from the construction that S_n and S'_n are compactly contained in $D_\alpha(\beta_1, \beta_2)$ (or otherwise each of them would be entirely in the set $C_\alpha(\beta_1, \beta_2)$). Let $S \in \mathbf{S}_\alpha$ be the set that contains the point $p \in K_1$, and $S' \in \mathbf{S}_\alpha$ the set that contains q. Note $S \cap S' = \emptyset$. By the same argument we have that $S \cap D_\alpha(\beta_1, \beta_2) = \emptyset$. The same is true for S'. We have that the accumulation set of the sequence S_n is contained in S. The accumulation set of S'_n is contained in S'.

Let $z_n = \psi(p_n) \in N$, and $w_n = \psi(q_n)$. Set $Z_n = \psi(S_n) \subset N$ and $W_n = \psi(S'_n) \subset N$. Note that $E_\alpha(Z_n) = Z_n$, and $E_\alpha(W_n) = W_n$. After passing to a subsequence if necessary, we may assume that Z_n has Hausdorff limit Z. Similarly, W_n has Hausdorff limit W. Clearly, Z and W are contained in the boundary circle of N that corresponds to K_1. Also, $E_\alpha(Z) = Z$ and $E_\alpha(W) = W$.

Since Z_n are connected sets we have that Z is connected and (as the Hausdorff limit) it is closed as well. The same applies to W. Now we show that $Z \cap W$ does not contain any open subinterval of the corresponding boundary circle. Assume that an open interval I is contained in $Z \cap W$. For every point $x \in I$, let $r_x \subset N$ be the geometric ray that ends at x (the geometric ray is the geodesic ray with respect to the hyperbolic metric on N). Then for almost every $x \in I$ we have that ψ^{-1} has the radial limit (see [16]). Let $x_0 \in I$ be such a point. Since Z_n and W_n are connected, we have that $\psi^{-1}(x_0)$ is contained in both S and S'. But this is impossible.

So, at least one of the sets Z and W is either a point or a closed interval which is not the full boundary circle. If Z is a point, then E_α has a fixed point in the corresponding boundary circle. If Z is an interval, then E_α fixes each boundary point of Z. This again shows that F_α is non-empty. We repeat the same argument for the other boundary circle (the one that corresponds to K_2).

Let β be a simple arc that connects the two boundary circles of N. Since we proved that F_α has points in both boundary circles, to prove this lemma it remains to compare the amount of twisting of the curves $E_\alpha(\beta)$ and $R_\alpha(\beta)$. We verify that these amounts are the same as follows. Let $\mathbf{S}'_\alpha$ denote the following decomposition of M. Every point in the (open) set $D_\alpha(\beta_1, \beta_2)$ is a component of $\mathbf{S}'_\alpha$. For every point of M that is outside $D_\alpha(\beta_1, \beta_2)$ we choose the corresponding component of $\mathbf{S}_\alpha$. It follows from the construction above that $\mathbf{S}'_\alpha$ is an upper semi continuous decomposition of M whose all components are acyclic. Moreover, both maps $\widetilde{R}_\alpha$

and $\mathcal{E}(t_\alpha)$ respect this decomposition. By applying the Moore Theorem, we obtain the projection map $\sigma : M \to M$ which semi-conjugates the maps $\widetilde{R}_\alpha$ and $\mathcal{E}(t_\alpha)$ to homeomorphisms R'_α and E'_α respectively (these are homeomorphisms of M). Here σ is homotopic to the identity map. Set $D'_\alpha(\beta_1,\beta_2) = \sigma(D_\alpha(\beta_1,\beta_2))$. We have that σ is a homeomorphism on $D_\alpha(\beta_1,\beta_2)$ (but it is not necessarily a homeomorphism of the closure of $D_\alpha(\beta_1,\beta_2)$). Also, the maps R'_α and E'_α are both identity outside $D'_\alpha(\beta_1,\beta_2)$. Since they are homotopic on M they must be homotopic on $D'_\alpha(\beta_1,\beta_2)$, where the corresponding homotopy preserves pointwise the frontier components of $D'_\alpha(\beta_1,\beta_2)$. By returning to $D_\alpha(\beta_1,\beta_2)$ by σ^{-1} and by applying ψ we conclude that R_α and E_α are homotopic on N modulo the set $F_\alpha \cap \partial N$. □

5.2. **Proof of Theorem 1.1.** We now assume that the genus of M is at least 3. The construction below does not make sense for $\mathbf{g} = 2$. We continue to assume that there exists the homomorphism $\mathcal{E} : \mathcal{MC}(M) \to \mathcal{H}omeo(M)$. We will derive a contradiction to this assumption at the end of this subsection.

Represent surface M as in the Figure 2. Let α, α_1,...,$\alpha_{\mathbf{g}-1}$ be simple closed curves as in the Figure 2. The curves α_i and α_j, $i \neq j$, are mutually disjoint, non-separating curves, and every component of $M \setminus (\alpha_i \cup \alpha_j)$ has genus at least one. This implies that α_i and α_j satisfy the assumptions of Lemma 5.2. The main idea of the proof below is that for a simple closed curve γ in M the corresponding homeomorphism $\mathcal{E}(t_\gamma)$ behaves similarly to an actual twist homeomorphism that corresponds to γ (any twist homeomorphism is the identity outside an annulus homotopic to γ). If $\mathcal{E}(t_\gamma)$ was an actual twist homeomorphism for every simple closed curve γ, we would have that for two such curves γ_1 and γ_2, where the intersection number between $[\gamma_1]$ and $[\gamma_2]$ is zero, the corresponding annuli that support the homeomorphisms $\mathcal{E}(t_{\gamma_1})$ and $\mathcal{E}(t_{\gamma_2})$ are disjoint from each other. One can than show that this contradicts the Artin's relations in the mapping class group.

There exists an obvious finite order homeomorphism $\widehat{e} : M \to M$, of order $\mathbf{g}-1$, where $\widehat{e}$ setwise preserves α, and maps α_i to α_{i+1}, for $1 \leq i \leq \mathbf{g}-1$, where $\alpha_{\mathbf{g}}$ is identified with α_1. Set $[\widehat{e}] = e \in \mathcal{MC}(M)$. Clearly one can choose a complex structure on M such that the homeomorphism $\widehat{e}$ is a conformal automorphism of M. We assume that M is endowed with that complex structure. Also, by Proposition 1.1 we may assume that $\mathcal{E}(e)$ is that conformal automorphism. Note that $\mathcal{E}(t_\alpha)$ commutes with $\mathcal{E}(e)$. By Artin's relations (1.1) and (1.2), we have

$$(\mathcal{E}(t_{\alpha_i}) \circ \mathcal{E}(t_\alpha))(B_{\alpha_i}) = B_\alpha$$

and

$$(\mathcal{E}(t_\alpha) \circ \mathcal{E}(t_{\alpha_i}))(B_\alpha) = B_{\alpha_i}.$$

The sets B_α, B_{α_1},.., $B_{\alpha_{\mathbf{g}-1}}$, have the same meaning as above. Another important observation is that $\mathcal{E}(e)(B_\alpha) = B_\alpha$, and $\mathcal{E}(e)(B_{\alpha_i}) = B_{\alpha_{i+1}}$. By Lemma 5.2 we have that

$$B_{\alpha_i} \cap B_{\alpha_j} = \emptyset. \tag{5.2}$$

One can choose a characteristic annulus $D_\alpha(\beta_1,\beta_2)$ for α (and with respect to the appropriate curves β_1 and β_2), such that $D_\alpha(\beta_1,\beta_2)$ is invariant as a set under the map $\mathcal{E}(e)$ (this is easily achieved by choosing curves β_1 and β_2 to be invariant

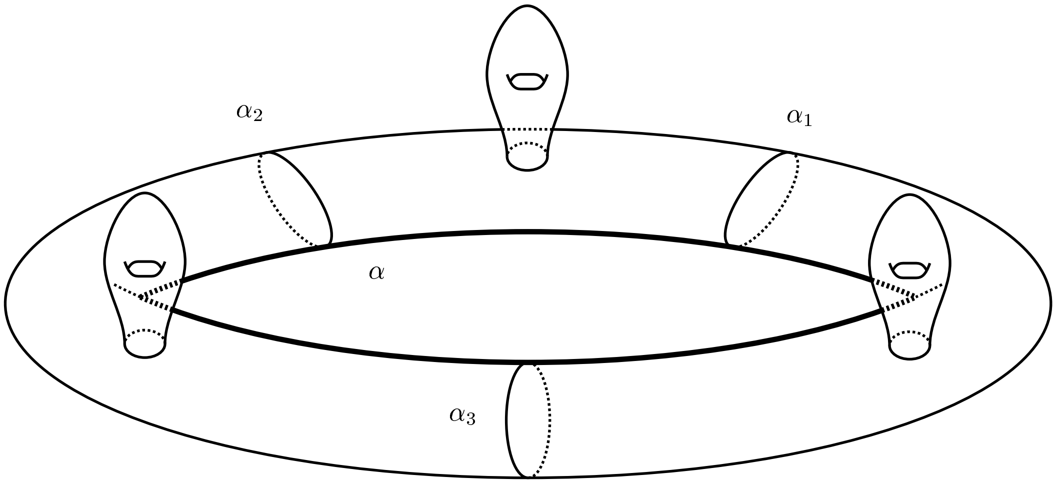

FIGURE 2. In this figure the surface M has genus $\mathbf{g} = 4$. We represent M as the base torus with $\mathbf{g}-1$ handles attached to it. The conformal automorphism (rotation) of M, setwise preserves the curve α, and it maps α_i to α_{i+1}. The order of this rotation is $\mathbf{g}-1$.

under $\mathcal{E}(e)$). Also, we can choose such $D_\alpha(\beta_1,\beta_2)$ to be arbitrarily close to B_α in the Hausdorff metric. We specify below the exact choice of $D_\alpha(\beta_1,\beta_2)$.

Let

$$\beta_{1,1} = (\mathcal{E}(t_{\alpha_1}) \circ \mathcal{E}(t_\alpha))^{-1}(\beta_1)$$

and

$$\beta_{2,1} = (\mathcal{E}(t_{\alpha_1}) \circ \mathcal{E}(t_\alpha))^{-1}(\beta_2).$$

Then $D_{\alpha_1}(\beta_{1,1},\beta_{2,1})$ is a characteristic annulus for α_1 and

$$(\mathcal{E}(t_{\alpha_1}) \circ \mathcal{E}(t_\alpha))(D_{\alpha_1}(\beta_{1,1},\beta_{2,1})) = D_\alpha(\beta_1,\beta_2).$$

We can choose $D_\alpha(\beta_1,\beta_2,)$ close enough to B_α, so that $D_{\alpha_1}(\beta_{1,1},\beta_{2,1})$ is close enough to B_{α_1}, and so that the closures of the sets $\mathcal{E}(e)^i(D_{\alpha_1}(\beta_{1,1},\beta_{2,1}))$, $i = 1,..,\mathbf{g}-1$, are mutually disjoint. Moreover, by Artin's relations we have that $(\mathcal{E}(t_\alpha) \circ \mathcal{E}(t_{\alpha_1}))(D_\alpha(\beta_1,\beta_2))$ is a characteristic annulus for α_1, but this one is not necessarily the same as $D_{\alpha_1}(\beta_{1,1},\beta_{2,1})$. However, by choosing $D_\alpha(\beta_1,\beta_2)$ close enough to B_α, we can arrange that the closures of the sets

$$\mathcal{E}(e)^i((\mathcal{E}(t_\alpha) \circ \mathcal{E}(t_{\alpha_1}))(D_\alpha(\beta_1,\beta_2))), i = 1,...,\mathbf{g}-1,$$

are mutually disjoint. This is how we choose the annulus $D_\alpha(\beta_1,\beta_2)$.

Remark. The sets $\mathcal{E}(e)^i(D_{\alpha_1}(\beta_{1,1},\beta_{2,1}))$, and $\mathcal{E}(e)^i((\mathcal{E}(t_\alpha) \circ \mathcal{E}(t_{\alpha_1}))(D_\alpha(\beta_1,\beta_2)))$, $2 \le i \le \mathbf{g}-1$, are the characteristic annuli for the curves α_i, $2 \le i \le \mathbf{g}-1$. However, $\mathcal{E}(t_{\alpha_1})$ does not have to setwise preserve these sets. But since these sets are very close to the corresponding sets B_{α_i}, $2 \le i \le \mathbf{g}-1$, we have that $\mathcal{E}(t_{\alpha_1})$ nearly preserves these sets. We will not make a use of this remark, but it helps understand the picture.

From now on we use the following notation

$$D_\alpha(\beta_1,\beta_2) = D_0,$$

$$\mathcal{E}(e)^i(D_{\alpha_1}(\beta_{1,1},\beta_{2,1})) = D_{i+1},$$

and

$$\mathcal{E}(e)^i((\mathcal{E}(t_\alpha)\circ\mathcal{E}(t_{\alpha_1}))(D_\alpha(\beta_1,\beta_2))) = D'_{i+1}.$$

We define

$$X_i = D_0 \cap D_i, X'_i = D_0 \cap D'_i.$$

Let N be the geometric annulus (defined above) that is the conformal image of D_0. Let $p \in \{|z| = r\}$ and $q \in \{|z| = 1\}$ be two fixed points of E_α (here E_α is the conjugate of the restriction of $\mathcal{E}(t_\alpha)$ by the conformal map $\psi : D_0 \to N$). Let T denote the standard rotation of N of order $\mathbf{g}-1$. We can obtain T as the conjugate of the restriction of $\mathcal{E}(e)$ by the conjugating map ψ. Set $p = p_1$ and $p_{i+1} = T^i(p_1)$. Similarly for q. Let I_i be the subinterval of the circle $\{|z| = r\}$ that lies in between the points p_i and p_{i+1}. Similarly we define the subintervals J_i of $\{|z| = 1\}$. There are $\mathbf{g}-1$ intervals I_i, and the same number of J_i.

Set $Y_i = \psi(X_i)$ and $Y'_i = \psi(X'_i)$. Note that $Y_{i+1} = T^i(Y_1)$ and $Y'_{i+1} = T^i(Y'_1)$. We have that Y_i and Y_j are mutually disjoint subsets of N, $i \neq j$. Similarly for Y'_i and Y'_j.

Let ℓ_1 be a connected component of the set $N \cap \psi(\partial D_1)$. The set ∂D_1 is the union of the two frontier components of D_1. The two corresponding ends have the homotopy type of the curve α_1 which transversally intersects α. Therefore, we can choose ℓ_1, so that ℓ_1 connects the two boundary circles of N. Since ℓ_1 is a connected subset of N (this implies that $\psi^{-1}(\ell_1)$ is connected subset of D_0) we have that the closures of the sets ℓ_1 and $T^i(\ell_1)$, $i \neq 1$, are disjoint subsets of $\overline{N}$. This follows from Lemma 2.4 and the fact that the closures of the sets $\psi^{-1}(\ell_1)$ and $\mathcal{E}(e)(\psi^{-1}(\ell_1))$ are disjoint connected subsets of D_0.

Set $L_1 = E_\alpha(\ell_1)$. We have that L_1 is a connected subset of N.

Lemma 5.4. *The connected set L_1 intersects at least $\mathbf{g}-3$ sets $\psi(B_{\alpha_i})$ (altogether there are $\mathbf{g}-1$ sets $\psi(B_{\alpha_i})$, $i = 1,..,\mathbf{g}-1$).*

Proof. The set $\psi(B_{\alpha_i})$ is contained in Y_i, $i = 1,...,\mathbf{g}-1$. Let B'_{α_1} be a connected component of $\psi(B_{\alpha_1})$, that connects the two boundary circles of N. Such a component exists because the boundary of the set A_{α_1} is contained in B_{α_1}, and it has the homotopy type of α_1 (which transversally intersects α). Set $B'_{\alpha_{(i+1)}} = T^i(B'_{\alpha_1})$, $i = 1,...,\mathbf{g}-1$. The component B'_{α_i} is disjoint from ℓ_1, as is the entire set $\psi(B_{\alpha_i})$ (note that $l_1 \cap B'_{\alpha_1} = \emptyset$). Therefore, we have two connected, disjoint subsets of N, that both connect the two boundary circles. This implies that $R_\alpha(\ell_1)$ intersects each B'_{α_i}, $i = 1,..,\mathbf{g}-1$, inside N.

Let V be a component of the set $N \setminus \ell_1$. The intersection between the closure of V and a boundary circle of N is a connected set. This directly follows from the fact that ℓ_1 is a connected subset of N. There exists one and only one component of the set $N \setminus \ell_1$ whose boundary has non-empty intersection with both boundary circles of N. That there is at most one such component comes from the fact that ℓ_1 is a connected subset of N. On the other hand, $\psi(B_{\alpha_2})$ is contained in a component of the set $N \setminus \ell_1$. This component must connect the boundary circles of N since

$\psi(B_{\alpha_i})$ does it. This proves that there is at least one such component. This particular component of the set $N \setminus \ell_1$ we denote by V_0.

Let P_1 be the subset of $\{|z| = r\}$ which contains all the points on $\{|z| = r\}$ which are either in the closure of ℓ_1, or are not in the closure of the set V_0. Since ℓ_1 is connected, and since V_0 connects the two boundary circles of N, we have that P_1 is a closed, connected subset of the circle $\{|z| = r\}$. P_1 can be either a closed interval or a point. For the other circle $\{|z| = 1\}$, the set Q_1 is defined in the same way.

Since each B'_{α_i} is a connected subset of N, that connects the two boundary circles, we can define the sets $Z_i \subset \{|z| = r\}$, and $W_i \subset \{|z| = 1\}$, that correspond to B'_{α_i} (the sets Z_i and W_i are the analogues of the sets P_i and Q_i respectively). We have $Z_{i+1} = T^i(Z_1)$ and $W_{i+1} = T^i(W_1)$. Set $P_{i+1} = T^i(P_1)$ and $Q_{i+1} = T^i(Q_1)$. P_i is disjoint from P_j, $i \neq j$, because the closures $\overline{\ell_i}$ and $\overline{\ell_j}$ are disjoint. Similarly for Q_i. Also, P_i is disjoint from any Z_j, because the closure of ℓ_i is disjoint from the closure of any B'_{α_i} (similarly as above, we use that B'_{α_i} and l_1 are connected, and that the closures of $\psi^{-1}(B'_{\alpha_i})$ and $\psi^{-1}(l_1)$ are disjoint subsets of $\overline{D_0}$) . Similarly for Q_i and W_j.

We saw that $R_\alpha(\ell_1)$ intersects each B'_{α_i}. By Lemma 5.3 we have that E_α is homotopic to R_α on N modulo at least the sets $p_1, .., p_{\mathbf{g}-1}$ and $q_1, .., q_{\mathbf{g}-1}$. The points p_i and q_i are the fixed points for E_α, and were introduced above (before the statement of Lemma 5.4). P_1 (or Q_1) can contain at most one of the points p_i (or q_i) that are fixed for E_α. If both P_1 and Q_1 contain such points then L_1 intersects every B'_{α_i} (because $R_\alpha(\ell_1)$ does so). We assume that neither P_1 or Q_1 contain any of the fixed points (the case when only one of them does not contain the fixed point is treated similarly). Also, for a fixed $i \in \{1, ..., \mathbf{g} - 1\}$, if both Z_i and W_i contain the corresponding fixed points, we again have that L_1 intersects B'_{α_i}. Therefore, we also assume that neither Z_i or W_i contain the fixed points (the case when only one of them contains the fixed point is treated in a similar way).

After re-numerating the points p_i and q_i we may assume that P_1 is contained in the open interval bounded by the points p_1 and p_2. Q_1 is contained in the open interval bounded by the points q_1 and q_2. If B'_{α_i} is such that the corresponding Z_i and W_i are not subsets of the corresponding intervals (p_1, p_2) and (q_1, q_2), then L_1 intersects B'_{α_i}. This shows that L_1 intersects at least $\mathbf{g} - 3$ sets B'_{α_i}, since there are at most two sets B'_{α_i} such that either one of the corresponding sets Z_i or W_i is in one of the corresponding intervals (p_1, p_2) and (q_1, q_2).

□

To finish the argument it is convenient to introduce a sequence of simple closed curves $\gamma'_n \subset D_1$, $n \in \mathbf{N}$, that converge to the frontier component of D_1 which contains $\psi^{-1}(\ell_1)$. By $\gamma_n \subset N$ we denote the appropriately chosen arcs such that $\psi^{-1}(\gamma_n) \subset \gamma'_n$, γ_n connects the two sides of N, and γ_n converges to ℓ_1.

Remark. If ℓ_1 was an arbitrary connected subset of N, which connects the two boundary sides of N, then one would not be able to choose such a sequence of arcs γ_n. In this general setting, the sequence γ_n would converge only to a subset of ℓ_1 (but this subset would still connect the two sides of N). But in our case, the set $\psi^{-1}(\ell_1)$ is a part of the frontier component of the annulus D_1, and therefore one can choose a sequence γ'_n that converges to the corresponding frontier component.

Let $L_1(n) = E_\alpha(\gamma_n)$. Then $L_1(n) \subset N$, converges to L_1.

Since $(\mathcal{E}(t_{\alpha_1}) \circ \mathcal{E}(t_\alpha))(D_1) = D_0$, and since $\psi^{-1}(\ell_1)$ is a part of a frontier component of D_1, we have that

$$(\mathcal{E}(t_{\alpha_1}) \circ \mathcal{E}(t_\alpha))(\psi^{-1}(\ell_1)) \subset \partial D_0.$$

This implies the following relation

$$\mathcal{E}(t_{\alpha_1})(\psi^{-1}(L_1)) = K \subset \partial D_0. \tag{5.3}$$

In fact, we have that the sequence $\mathcal{E}(t_{\alpha_1})(\psi^{-1}(L_1(n))) \subset D_0$ approaches K in the Hausdorff metric.

Set $H_n = \psi\left(\mathcal{E}(t_{\alpha_1})(\psi^{-1}(L_1(n)))\right)$. After passing to a subsequence if necessary, let H be the limit of H_n. We have that H is a connected subset of one of the boundary circles of N (say $H \subset \{|z| = r\}$). H is connected as the limit of connected sets in the compact and complete space.

Since (by Lemma 5.2) $\mathcal{E}(t_{\alpha_1})$ setwise preserves the sets B_{α_i}, $1 \le i \le \mathbf{g}-1$, and since L_1 (and therefore $L_1(n)$ for large n) intersects at least $\mathbf{g}-3$ sets $\psi(B_{\alpha_i})$, $1 \le i \le \mathbf{g}-1$, we have that the set H contains at least one point from at least $\mathbf{g}-3$ sets Z_i, $1 \le i \le \mathbf{g}-1$ (the sets Z_i were defined in the proof of Lemma 5.4). The sets Z_i are connected subsets of the corresponding boundary circle of N. Since $Z_{i+1} = T(Z_i)$, and since H is connected, we conclude that H contains at least $\mathbf{g}-5$ sets Z_i, $1 \le i \le \mathbf{g}-1$. In conclusion, we have that H contains at least $\mathbf{g}-5$ sets Z_i, $1 \le i \le \mathbf{g}-1$, and also H contains at least one point from at least $\mathbf{g}-3$ sets Z_i, $1 \le i \le \mathbf{g}-1$.

Now, we use the fact that $(\mathcal{E}(t_\alpha) \circ \mathcal{E}(t_{\alpha_1}))(D_0) = D'_1$. Since $\psi^{-1}(L_1) \subset D_0$ we have that $(\mathcal{E}(t_\alpha) \circ \mathcal{E}(t_{\alpha_1}))(\psi^{-1}(L_1)) \subset D'_1$. Combining this with (5.3) we conclude that $\mathcal{E}(t_\alpha)(K) \subset D'_1$. Since K is a subset of ∂D_0, and since $\mathcal{E}(t_\alpha)$ setwise preserves ∂D_0, we have $\mathcal{E}(t_\alpha)(K) \subset (D'_1 \cap \partial D_0)$.

On the other hand, going back to N again, we have that $E_\alpha(H) \subset \{|z| = r\}$ is contained in the closure of the set $\psi(D'_1)$ (the map ψ is not defined on the entire set D'_1, and $\psi(D'_1)$ denotes the image of the corresponding subset of D'_1 where ψ is defined). Let $\widehat{Z}_i = E_\alpha(Z_i)$, $1 \le i \le \mathbf{g}-1$. Since the homeomorphism E_α commutes with the rotation T we have that $T(\widehat{Z}_i) = \widehat{Z}_{i+1}$. Also, the set $E_\alpha(H)$ contains at least $\mathbf{g}-5$ sets $\widehat{Z}_i$, $1 \le i \le \mathbf{g}-1$, and intersects at least $\mathbf{g}-3$ sets $\widehat{Z}_i$, $1 \le i \le \mathbf{g}-1$. Altogether, this implies that the length of the closed, connected set $E_\alpha(H) \subset \{|z| = r\}$ is at least $2\pi r \frac{\mathbf{g}-5}{\mathbf{g}-1}$.

However, the closures of the sets $\mathcal{E}(e)^i(D'_1)$, $i = 1, ..., \mathbf{g}-1$, are disjoint. This implies that the intersection between the closures of the sets $\psi(D'_i)$ and $\psi(D'_j)$, $i \neq j$, can not contain an open interval (note that if we had that $\psi(D'_i)$ and $\psi(D'_j)$ were connected, that would yield $\overline{\psi(D'_i)} \cap \overline{\psi(D'_j)} = \emptyset$). Since H is a subset of the closure of $\psi(D'_1)$, we have that the interiors of the sets $T^i(E_\alpha(H))$, $1 \le i \le \mathbf{g}-1$, are mutually disjoint. Since the length of $E_\alpha(H)$ is at least $2\pi \frac{\mathbf{g}-5}{\mathbf{g}-1}$ we obtain a contradiction if

$$\frac{\mathbf{g}-5}{\mathbf{g}-1} > \frac{1}{\mathbf{g}-1},$$

which is equivalent to $\mathbf{g} > 6$.

If $\mathbf{g} = 6$, the above situation is sustainable. But then every point in the corresponding boundary circle is contained in some of the sets $T^i(E_\alpha(H))$, $1 \le i \le \mathbf{g}-1$. Consequently, every point in the corresponding boundary circle is contained in the closure of some $\psi(D_i')$. Since the closures of the sets $\mathcal{E}(e)^i(D_1')$, $i = 1, ..., \mathbf{g}-1$, are disjoint, we can choose a component of $D_0 \setminus \cup_{1<i<\mathbf{g}-1}\mathcal{E}(e)^i(D_1')$, such that the intersection of the closure of that component and the corresponding frontier component of D_0 contains a relatively open subset of that frontier component of D_0, and which is not in the closure of the union $\cup_{1<i<\mathbf{g}-1}\mathcal{E}(e)^i(D_1')$. This implies that there is an open interval in the corresponding boundary circle of N, and no point from this interval is contained in the closure of some $\psi(D_i')$. Again, a contradiction. This completes the proof of Theorem 1.1.

References

[1] L.Ahlfors, L. Sario, *Riemann surfaces.* Princeton University Press, Princeton, N.J. (1960)

[2] A. Casson, S. Bleiler, *Automorphisms of surfaces after Nielsen and Thurston.* London Mathematical Society Student Texts, 9. Cambridge University Press, Cambridge, (1988)

[3] R.J. Daverman, *Decomposition of manifolds.* volume 124 of Pure and Applied Mathematics, (1986)

[4] D. Epstein, V. Markovic, *Extending homeomorphisms of the circle to quasiconformal homeomorphisms of the disc.* to appear in Topology.

[5] J. Franks, *Anosov diffeomorphisms.* Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968) pp. 61–93, (1970)

[6] D. Gabai, *Convergence groups are Fuchsian groups*, Bull. Amer. Math. Soc. (N.S.) 25 (1991), no. 2, 395-402.

[7] F. Gehring, G. Martin, *Discrete convergence groups.* Complex analysis, 158–167, Lecture Notes in Math., 1275, Springer, Berlin, (1987)

[8] M. Handel, *Global shadowing of pseudo-Anosov homeomorphisms.* Ergodic Theory Dynam. Systems 5 , no. 3, 373–377. (1985)

[9] S. Kerckhoff, *The Nielsen realization problem.* Ann. of Math. (2) 117, no. 2, 235–265. (1983)

[10] R. Kirby, *Problems in low-dimensional topology.* Geometric topology (Athens, GA, 1993), 35–473, Amer. Math. Soc., Providence, RI, (1997)

[11] G. Martin, P. Tukia, *Convergence and Möbius groups.* Holomorphic functions and moduli, Vol. II, 113–140. (1986)

[12] J. Mather, *Integrability in codimension* 1. Comment. Math. Helv. 48, 195–233. (1973)

[13] R. Moore, *Concerning upper semi-continuous collections of continua.* Trans. Amer. Math. Soc., 27, 416-428 (1925)

[14] S. Morita, *Characteristic classes of surface bundles.* Invent. Math. 90, no. 3, 551–577 (1987)

[15] S. Morita, *Geometry of characteristic classes.* Translations of Mathematical Monographs, 199. Iwanami Series in Modern Mathematics. American Mathematical Society, (2001)

[16] Ch. Pommerenke, *Boundary behavior of conformal maps.* Fundamental Principles of Mathematical Sciences, 299. Springer-Verlag, Berlin, (1992)

[17] L. Sario, K. Oikawa, *Capacity functions.* Springer-Verlag, (1969)

[18] W. Thurston, *Foliations and groups of diffeomorphisms.* Bull. Amer. Math. Soc. 80, 304–307. (1974)

[19] N. Ivanov, *Mapping class groups.* Preprint

University of Warwick, Institute of Mathematics, Coventry, CV4 7AL, UK
E-mail address: `markovic@maths.warwick.ac.uk`