ON HARNACK'S THEOREM AND EXTENSIONS: A GEOMETRIC PROOF AND APPLICATIONS

ANTONIO F. COSTA AND HUGO PARLIER

Abstract. Harnack's theorem states that the fixed points of an orientation reversing involution of a compact orientable surface of genus g are a set of k disjoint simple closed geodesic where $0 \leq k \leq g+1$. The first goal of this article is to give a purely geometric, complete and self-contained proof of this fact. In the case where the fixed curves of the involution do not separate the surface, we prove an extension of this theorem, by exhibiting the existence of auxiliary invariant curves with interesting properties. Although this type of extension is well known (see for instance [7] and [9]), our method also extends the theorem in the case where the surface has boundary. As a byproduct, we obtain a geometric method on how to obtain these auxiliary curves. As a consequence of these constructions, we obtain results concerning presentations of Non-Euclidean crystallographic groups and a new proof of a result on the set of points corresponding to real algebraic curves in the compactification of the Moduli space of complex curves of genus g, $\overline{\mathcal{M}_g}$. More concretely, we establish that given two real curves there is a path in $\overline{\mathcal{M}_g}$ which passes through at most two singular curves, a result of M. Seppala [11].

1. Introduction

From the moment Felix Klein observed the equivalence between real algebraic curves and Riemann surfaces with an anticonformal involution, the study of orientation reversing involutions on surfaces became an important problem. The first step in this study is the topological classification of such automorphisms, i. e., two anticonformal involutions σ_1 and σ_2 of a Riemann surface are topologically equivalent if there is a homeomorphism h of S such that $\sigma_1 = h \circ \sigma_2 \circ h^{-1}$. This equivalence was solved by A. Harnack [5], and G. Weichold [12] at the end of the nineteenth century. The theorems of A. Harnack and G. Weichold can be summarized as follows: let S be a Riemann surface of genus g and σ be an anticonformal involution. The set $\mathrm{Fix}(\sigma)$ of fixed points of σ consists in a finite set of disjoint simple closed curves in S, called ovals. The topological equivalence class of σ is determined by the number k of ovals of σ and the connectedness or not of $S - \mathrm{Fix}(\sigma)$. Furthermore, if $S - \mathrm{Fix}(\sigma)$ is connected then $0 \leq k \leq g$ and if $S - \mathrm{Fix}(\sigma)$ is not connected then $1 \leq k \leq g+1$ and $k + g \equiv 1 \mod 2$.

Date: August 25, 2008.

2000 *Mathematics Subject Classification.* Primary 30F10, 32G15; Secondary 14H50, 30F20.

Key words and phrases. Orientation reversing involutions, simple closed geodesics, hyperbolic Riemann surfaces.

The first author was supported in part by BFM 2002-04801, the second author was supported in part by SNFS grant number PBEL2-106180.

The theorems of Harnack and Weichold, which we shall refer to as *Harnack's Theorem*, play a very important role in the theory of real algebraic curves and there are modern proofs of these theorems in the works of S. M. Natanzon and others, see for example [7] and [9].

This article has been written to reach three specific goals. Our first goal is to give an elementary proof of Harnack's Theorem using techniques from the study of geodesics on hyperbolic surfaces. The proof we present is such that with not much extra effort, we obtain an extended version of this theorem (our second goal). Other extensions of this theorem exist see for instance [7] and [9] (pages 63-67). The proof of S. M. Natanzon has some similarities with ours but we deal with several essential points in a very different way. Our extension includes treating the case of *all* hyperbolic orientable Riemann surfaces, as opposed to just *closed* surfaces as is generally considered. Specifically, our main theorem ([7] and [9]) is the following:

Theorem 1.1. *Let S be a hyperbolic Riemann surface with an orientation reversing involution σ. Then $\mathrm{Fix}(\sigma)$ is a set (possibly empty) of disjoint simple complete geodesics.*

(1) *If $\mathrm{Fix}(\sigma)$ is separating, then denote by g the genus of S and n the number of curves in $\mathrm{Fix}(\sigma)$. We have $S \setminus \mathrm{Fix}(\sigma) = S_1 \cup S_2$ where S_1 and S_2 are connected surfaces with boundary such that $\sigma(S_1) = S_2$, n and g are of different parity and $0 < n \leq g+1$.*

(2) *If $\mathrm{Fix}(\sigma)$ is non-separating, denote by g the genus of $S \setminus \mathrm{Fix}(\sigma)$.*
   (a) *If g is even, for all odd k between 1 and $g+1$ there exist sets $\{\Gamma_k\}$ consisting of k disjoint individually invariant simple closed geodesics such that $\mathrm{Fix}(\sigma) \cup \Gamma_k$ separates S into two isometric surfaces S_1 and S_2 and $\sigma(S_1) = S_2$.*
   (b) *If g is odd, for all even k between 2 and $g+1$ there exist sets $\{\Gamma_k\}$ consisting of k disjoint individually invariant simple closed geodesics such that $\mathrm{Fix}(\sigma) \cup \Gamma_k$ separates S into two isometric surfaces S_1 and S_2 and $\sigma(S_1) = S_2$.*

Our final goal is to give two applications of our theorem which we outline below.

In 1966, H. C. Wilkie [13] produced a presentation for *Non-Euclidean Crystallographic groups* (NEC groups) that became a useful tool and has become known as the *canonical* presentation. A first application of our theorem is the obtention of new algebraic presentations for NEC groups that are similar to the canonical presentation but provide additional geometric information about such groups.

The second application is a property of the set of points corresponding to surfaces with anticonformal involutions in the moduli of Riemann surfaces. This last application has a nice interpretation in terms of real algebraic curves: let C_1, C_2 be two (smooth) real algebraic curves. Each curve can be represented as the zeros of a set of real homogeneous polynomials. F. Klein [6] conjectured that one can continuously modify the real coefficients of such polynomials, preserving its real character, to arrive at a set of polynomials defining C_2, although along the way it is possible that the set of polynomials define, for some values, singular curves. This conjecture was proved by M. Seppala in [11]. Using our extension of Harnack's Theorem, we show something stronger than Klein's original conjecture, namely the existence of a path in the set of real algebraic curves from C_1 to C_2 which passes through at most two singular curves. Note that this result is contained in M. Seppala's original proof (Theorem 7.5 of [11]).

The remainder of this article is organized as follows. The next section is dedicated to definitions, notations and a limited amount of prerequisites. The geometric proof of the completed version of Harnack's theorem follows, and the final two sections deal with the applications of this theorem discussed above.

2. Preliminaries

Our main object of study is a Riemann surface S (i.e. an orientable surface with a conformal structure) of finite type (with finitely generated fundamental group), meaning with both finite topological genus g and finite number of boundary components n. If the Euler characteristic of S, denoted $\chi(S)$ and equal to $2-2g-n$, is negative, then by uniformization, S admits a hyperbolic metric which gives S the *same* conformal structure. We shall assume that all our surfaces are endowed with a hyperbolic metric, and a *surface* will always mean a hyperbolic orientable surface. More precisely, a Riemann surface S with negative Euler characteristic is conformally equivalent to $\mathbb{H}/G$ where $\mathbb{H}$ is the hyperbolic plane and G is a *discrete* subgroup of the isometries of $\mathbb{H}$. For practical purposes, in the case where $\mathbb{H}/G$ is not *closed*, it is more convenient to think of S as a compact surface with boundary where the boundary is a (finite) collection of simple closed geodesics and cusps. This subset of S is called the *Nielsen core* of the surface, and is obtained by cutting S along simple closed geodesics to remove infinite funnels. If S is of genus g and has n boundary components, we say the surface is of signature (g,n). The area of the Nielsen core is then $2\pi\chi(S)$ by Gauss-Bonnet, but for hyperbolic surfaces, this can be proved by decomposing the surface into *pairs of pants*, meaning surfaces of signature $(0,3)$, and showing that the area of each pair of pants is 2π by elementary hyperbolic geometry. A geodesic, simple or not, is called *complete* if it can be parametrized over $\mathbb{R}$ at unit speed. For instance, a simple closed geodesic is complete, but a geodesic segment is not. Note that a non-closed simple complete geodesic is naturally isometric to $\mathbb{R}$. A simple closed curve is called *non-trivial* if is not freely homotopic to a boundary curve or it does not bound a disk. For a surface S, the *systole* of S is the (or a) shortest non-trivial simple closed geodesic. Note then that by our definition of non-trivial, a systole is an *interior* simple closed curve, meaning not homotopic to boundary. We call a set Γ of disjoint simple complete geodesics *separating* if the surface obtained by cutting along all the curves in

Γ is disconnected. The main preliminary results we shall use are the two following propositions.

Proposition 2.1. *Let S be a hyperbolic surface. Let α, β be disjoint simple closed geodesics on S. Let c be a simple path from α to β. Then in the free homotopy class of c with endpoints gliding on α and β, there exists a unique shortest curve, denoted $\mathbb{G}(c)$, which meets α and β perpendicularly. Furthermore, if $\tilde{c}$ is also a simple path from α to β such that $c \cap \tilde{c} = \emptyset$, then either $\mathbb{G}(c) = \mathbb{G}(\tilde{c})$ or $\mathbb{G}(c) \cap \mathbb{G}(\tilde{c}) = \emptyset$.*

Proposition 2.2. *Let S be a hyperbolic surface and let c be a homotopically non-trivial simple closed curve on S. Then c is freely homotopic to a unique simple closed geodesic, denoted $\mathbb{G}(c)$. The curve $\mathbb{G}(c)$ is either contained in ∂S or $\mathbb{G}(c) \cap \partial S = \emptyset$. If c is a non-smooth boundary component, then $\mathbb{G}(c)$ and c bound an embedded annulus.*

The proofs to these propositions can be found in [2], pp. 19-23, but the essence of the proofs is the fact that the universal cover for hyperbolic surfaces is $\mathbb{H}$, and in $\mathbb{H}$ for any two disjoint geodesics with distinct endpoints, there is a unique distance minimizing geodesic between them.

We shall also use the following lemma concerning systoles of surfaces.

Lemma 2.3. *Let η_1 and η_2 be distinct systoles of a hyperbolic orientable surface S. Then $\mathrm{int}(\eta_1, \eta_2) \leq 1$ or if $\mathrm{int}(\eta_1, \eta_2) = 2$, then S necessarily has two distinct boundary curves.*

This is not a difficult fact to prove, and is well discussed in [10]. The idea is, in the case where the two systoles intersect more than once, to find two non-trivial loop in the trace of $\eta_1 \cup \eta_2$ which is strictly shorter than the systole length. Because we have supposed that η_1 and η_2 are systoles, this implies that these shorter loops are in fact homotopic to boundary.

3. A geometric proof of the extended Harnack theorem

For a given point p of a surface, the set of equidistant points from p will be called a *circle* if the set bounds a disk. Notice that the radius of a circle centered in p is necessarily less than the injectivity radius of the surface in p.

Lemma 3.1. *If f is an isometry of S such that a circle C of S is fixed pointwise, then $f = \mathrm{id}_S$.*

Proof. First note that the open disk D bounded by C is globally invariant by f, otherwise $f(D)$ is a another disk with boundary C and the surface is a sphere. It follows that the center of C is preserved by f, and thus all radii of C are fixed pointwise. It follows that $f = \mathrm{id}\,|_D$ which implies that $f = \mathrm{id}$. □

Lemma 3.2. *Let f be an orientation reversing isometry with fixed points. Then f is an involution.*

Proof. For $p \in \mathrm{Fix}(f)$ consider the circle C of radius $r < \mathrm{inj}_S(p)$ and center p, i.e., $C = \{x \in S \mid d(x,p) = r\}$. Clearly $f(C_{p,r}) = C_{p,r}$ and for a given orientation of $C_{p,r}$, this orientation is reversed by f. It follows that there are exactly two antipodal points, say p' and p'', of C that are fixed by f. The two segments of C separated by p' and p'', say c_1 and c_2, are interchanged by f. It follows that $f^2 \mid_C = \mathrm{id}$ and the result follows by the previous lemma. □

Using the same idea as the previous lemma, let us characterize the fixed point set of an orientation reversing involution.

Lemma 3.3. *Let σ be an orientation reversing involution of a surface S. The set of fixed points of σ is a union of disjoint simple complete geodesics of S.*

Proof. As in the proof of the previous lemma, for $p \in \mathrm{Fix}(f)$ consider the circle $C_{p,r}$ of radius $r < \mathrm{inj}_S(p)$ and center p, and the two points $\{p', p''\} = C \cap \mathrm{Fix}(\sigma)$. Denote by $c_{p,p'}$ the shortest geodesic path between p and p', resp. $c_{p,p''}$ the shortest geodesic path between p and p''. We now show that $\sigma(c_{p,p'}) = c_{p,p'}$. Suppose this is not the case. Then there is another path $c'_{p,p'}$, non-homotopic to $c_{p,p'}$, such that $\ell(c_{p,p'}) = \ell(c_{p,p'}) < \mathrm{inj}_S(p)$. As $c_{p,p'}$ and $c'_{p,p'}$ are both shortest paths between p and p', it follows that $c_{p,p'} \cap c'_{p,p'} = \{p, p'\}$. The closed curve $c_{p,p'} \cup c'_{p,p'}$ is thus simple and non-homotopically trivial, and of length $< 2\mathrm{inj}_S(p)$, a contradiction. Thus $\sigma(c_{p,p'}) = c_{p,p'}$. Because the end-points of $c_{p,p'}$ are fixed points of f, it follows that $f(q) = q$ for all $q \in c_{p,p'}$. Similarly, $f(q) = q$ for all $q \in c_{p,p''}$. Notice that the path $c := c_{p,p'} \cup c_{p,p''}$ is smooth in p, as p' and p'' are diametrically opposite. Now for any $r' \leq r$, the circle $C_{p,r'}$ of center p and of radius r', also contains exactly two fixed points of σ, and by what precedes, these are exactly the intersection points between $C_{p,r'}$, and c. It follows that the path c is exactly the fixed point set of f contained in the closed disk $\bar{D}_{p,r} := \{x \in S \mid d(x,p) \leq r\}$.

Now we have $c \subset \mathrm{Fix}(\sigma)$. Taking the endpoints of c, the process can then be repeated until the fixed geodesic is complete. By what precedes, this geodesic is simple. So all $p \in \mathrm{Fix}(\sigma)$ lie on a simple complete geodesic. Note that these fixed geodesics cannot intersect by the same reasoning. This proves the lemma. □

Lemma 3.4. *Let γ be a simple complete geodesic such that $\mathrm{Fix}(\sigma) \cap \gamma = \emptyset$ and $\sigma(\gamma) = \gamma$. Then γ is a simple closed geodesic. Furthermore, the image $\sigma(p)$ for all $p \in \gamma$ is the point on γ diametrically opposite from p.*

Proof. Suppose that γ is not a simple closed geodesic. Then γ is isometric to $\mathbb{R}$, and the only fixed point free involution acting on $\mathbb{R}$ is the identity, which of course contains fixed points. So γ is a simple closed geodesic. Now σ is an involution acting on γ, and the only fixed point free isometric involution acting on a circle is the rotation of angle π. □

In order to completely characterize the fixed point set of an orientation reversing involution, the following proposition is necessary.

Proposition 3.5. *Let σ be an orientation reversing involution of a surface S. Let $\Gamma := \{\gamma_1, \dots, \gamma_n\}$ be a set of disjoint complete simple geodesics such that $\sigma(\gamma_k) = \gamma_k$*

$\forall k \in \{1,\ldots,n\}$, $\mathrm{Fix}(\sigma) \subset \Gamma$ and such that $S \setminus \Gamma$ is not connected. Then the connected components of $S \setminus \Gamma$ consists of two surfaces with boundary S_1 and S_2 such that $\sigma(S_1) = S_2$.

Proof. As $\mathrm{Fix}(\sigma) \subset \Gamma$, each γ_k is either a connected component of the fixed point set, or a globally invariant simple closed geodesic as in lemma 3.4.

Consider the set of connected subsurfaces $S_1, \ldots, S_m$ obtained by cutting S along Γ. As σ is an isometry, it acts as a homeomorphism on each S_k. Thus $\sigma(S_k)$ is a surface homeomorphic to S_k for each $k \in \{1,\ldots,m\}$. Now as the elements of Γ are invariant, it follows that for each $k \in \{1,\ldots,m\}$, there exists a k' such that $\sigma(S_k) = S_{k'}$. Furthermore, because the boundary curves of S_k are stable and because σ is an involution, this implies that the boundary curves of S_k and $S_{k'}$ are the same. As S is connected, we have $S \setminus \Gamma = S_k \cup S_{k'}$, thus $m = 2$ and this concludes the proof. □

As a consequence of the above proposition, we can obtain some results on the cardinality of Γ.

Corollary 3.6. *Suppose S is closed of genus g. Let $\Gamma := \{\gamma_1, \ldots, \gamma_n\}$ be the set of geodesics as in proposition 3.5. Then $n \leq g+1$ and if g is even, n is odd, and if g is odd, n is even.*

Proof. To see that $n \leq g+1$, it suffices to remark that $g+1$ simple topologically distinct and disjoint simple loops necessarily separate S. The rest of the corollary essentially follows by area arguments. The previous proposition tells us that Γ separates S into two isometric surfaces S_1 and S_2, so they are of equal area. Their signature is (g', n), so their area is equal to $\pi(2g' - 2 + n)$. We have then $2\pi(2g' - 2 + n) = \pi(2g - 2)$ which implies that n and g are of different parity. □

One could similarly find an equivalent of the above corollary in the case where S is not a closed surface, say of signature (g', k). The above argument is essentially topological if one replaces the area argument by reasoning on the Euler characteristic. To do this, consider the topological surface $\tilde{S}$ obtained by compactifying any boundary S may have. (By this we mean gluing an extra point at a cusp or a closed disk on any boundary curve.) If one denotes g the genus of $\tilde{S}$, one obtains that $n \leq g+1$ and that n and g also have different parity.

Theorem 3.7 (Harnack's Theorem)**.** *If S is closed surface of genus g, and σ is an orientation reversing involution of S, then the set of fixed points of σ is a set Γ of k disjoint simple closed geodesics such that $0 \leq k \leq g+1$. Furthermore, if Γ is separating, then $S \setminus \Gamma = S_1 \cup S_2$ where S_1 and S_2 are connected surfaces with boundary such that $\sigma(S_1) = \sigma(S_2)$.*

Proof. This is a direct consequence of lemma 3.3 and proposition 3.5 once one remarks that $g+1$ disjoint simple closed geodesics necessarily separate S. □

We would now like to extend this result (in an appropriate manner) in two ways. First of all, we would like to extend this to surfaces with boundary, or more precisely, to surfaces whose Nielsen core has boundary. Secondly, we would like to further characterize surfaces with orientation reversing involutions whose fixed point set is *not* separating (for example, for involutions *without* fixed points). The following generalization of Harnack's theorem appears in [7].

Theorem 3.8. *Let S be a hyperbolic Riemann surface with an orientation reversing involution σ. Then $\Gamma := \mathrm{Fix}(\sigma)$ is a set of disjoint simple complete geodesics.*

(1) *If Γ is separating, then $S \setminus \Gamma = S_1 \cup S_2$ where S_1 and S_2 are connected surfaces with boundary such that $\sigma(S_1) = S_2$.*

(2) *If Γ is non-separating, then there exists a set $\tilde{\Gamma}$, consisting of disjoint simple closed geodesics $\{\tilde{\gamma}_k\}_{k=1}^m$ such that $\Gamma' := \Gamma \cup \tilde{\Gamma}$ is a set of disjoint simple complete geodesics all individually globally invariant by σ, $S \setminus \Gamma' = S_1 \cup S_2$ where S_1 and S_2 are connected surfaces with boundary such that $\sigma(S_1) = S_2$.*

To prove this theorem, we shall prove a more general and specific version (see Lemma 1.2, page 65 in [9]).

Theorem 3.9. *Let S be a hyperbolic Riemann surface with an orientation reversing involution σ. Then $\mathrm{Fix}(\sigma)$ is a set (possibly empty) of disjoint simple complete geodesics.*

(1) *If $\mathrm{Fix}(\sigma)$ is separating, then denote by g the genus of S and n the number of curves in $\mathrm{Fix}(\sigma)$. We have $S \setminus \mathrm{Fix}(\sigma) = S_1 \cup S_2$ where S_1 and S_2 are connected surfaces with boundary such that $\sigma(S_1) = S_2$, n and g are of different parity and $0 < n \leq g + 1$.*

(2) *If $\mathrm{Fix}(\sigma)$ is non-separating, denote by g the genus of $S \setminus \mathrm{Fix}(\sigma)$.*
   (a) *If g is even, for all odd k between 1 and $g + 1$ there exist sets $\{\Gamma_k\}$ consisting of k disjoint individually invariant simple closed geodesics such that $\mathrm{Fix}(\sigma) \cup \Gamma_k$ separates S into two isometric surfaces S_1 and S_2 and $\sigma(S_1) = S_2$.*
   (b) *If g is odd, for all even k between 2 and $g + 1$ there exist sets $\{\Gamma_k\}$ consisting of k disjoint individually invariant simple closed geodesics such that $\mathrm{Fix}(\sigma) \cup \Gamma_k$ separates S into two isometric surfaces S_1 and S_2 and $\sigma(S_1) = S_2$.*

Remark 3.10. *In the case where $\mathrm{Fix}(\sigma)$ is not separating, any auxiliary set of individually invariant curves which completes $\mathrm{Fix}(\sigma)$ into a separating set has the cardinality of the theorem by proposition 3.5.*

Proof. All our claims for when $\mathrm{Fix}(\sigma)$ is separating are direct consequences of lemma 3.3 and proposition 3.5.

We proceed to the case when σ is non-separating. The core of proof in this case is the repetition of a trick which we hope will become apparent in what follows.

The first step consists in adding to the fixed point set of σ a set $\tilde{\Gamma}$ of disjoint simple closed geodesics such that each geodesic is globally invariant by σ and such that $\mathrm{Fix}(\sigma)\cup\tilde{\Gamma}$ is separating (or in other terms, $\tilde{\Gamma}$ is separating for the surface $S\setminus\mathrm{Fix}(\sigma)$).

We shall construct the set $\tilde{\Gamma}$. The construction begins with a simple complete geodesic γ such that $\sigma(\gamma)=\gamma$. If $\mathrm{Fix}(\sigma)$ is non-empty, then it suffices to take one of the connected components of $\mathrm{Fix}(\sigma)$. If $\mathrm{Fix}(\sigma)$ is empty, then consider one of the systoles of S, say η. Now either $\sigma(\eta)=\eta$, in which case we set $\gamma:=\eta$, or by lemma 2.3 we have $\mathrm{int}(\eta,\sigma(\eta)\leq 2$. Now if $\sigma(\eta)$ and η intersect once, then their intersection point would be a fixed point of σ, a contradiction. If $\sigma(\eta)$ and η intersect twice then by lemma 2.3, the surface necessarily has at least two boundary curves. We first prove that there is a globally invariant simple closed geodesic γ on such a surface with a fixed point free involution, and only afterwards will we treat the case where η and $\sigma(\eta)$ are disjoint.

Let S a surface with at least two boundary curves with a fixed point free involution σ. This is the first appearance of the main trick of our proof which we shall use repeatedly. Among all paths between distinct boundary components there is (at least) one that is shortest, and by proposition 2.1, it is simple and meets the two boundary curves it joins in right-angles. Denote this curve by c, and consider its image $\sigma(c)$ by the involution. Note that $c\cap\sigma(c)=\emptyset$, for otherwise we either have $c=\sigma(c)$, in which case the midpoint of c is a fixed point of σ, or c and $\sigma(c)$ intersect. If they intersect once, then their intersection point is a fixed point of σ, a contradiction. Now if they intersect more than once, then one can always find a shorter path between boundary components by the following process. Between two consecutive intersection points p and q, up to an exchange between c and $\sigma(c)$, we can suppose that $\ell(c)\mid_{pq}\geq\ell(\sigma(c))\mid_{pq}$. Now we can replace the segment of c between p and q by the segment of $\sigma(c)$ between p and q, we get a non-geodesic path between the boundary curves joined by c of less or equal length to the length of c, a contradiction. We have shown that $c\cap\sigma(c)=\emptyset$. Consider the midpoint M_c of c, and a shortest path d between the points M_c and $\sigma(M_c)$. By the same reasoning as above, the paths d and $\sigma(d)$ are geodesic, simple and do not intersect other than in M_c and $\sigma(M_c)$. It follows that they are homotopically distinct, and that $d\cup\sigma(d)$ is a simple closed curve not homotopic to a point. This curve is also non-homotopic to boundary, otherwise it would either bound an invariant cylinder with a boundary curve, which is impossible because σ is orientation reversing, or be homotopic to two distinct boundary geodesics, which is impossible because the boundary curves would then bound a cylinder, which is impossible on a hyperbolic surface. We now take γ to be the unique geodesic representative of $d\cup\sigma(d)$, and because $\sigma(d\cup\sigma(d))=\sigma(d)\cup d$, we have that $\sigma(\gamma)=\gamma$ as required.

We now consider that η and $\sigma(\eta)$ are disjoint (and not equal). Now consider a shortest path c between η and $\sigma(\eta)$ on S. Note that the path c is simple, geodesic and forms a right angle with both η and $\sigma(\eta)$. The path $\sigma(c)$ is also a shortest path between η and $\sigma(\eta)$, as σ is an involution. We claim that these paths satisfy

$c \cap \sigma(c) = \emptyset$. To see this, note that if they intersect more than once, then there would be a shorter path between η and $\sigma(\eta)$, and if they intersect once, then the point of intersection is a fixed point of σ, a contradiction. Now if $\sigma(c) = c$, then the midpoint of the path would be a fixed point of σ, yet another contradiction obtained using the same trick.

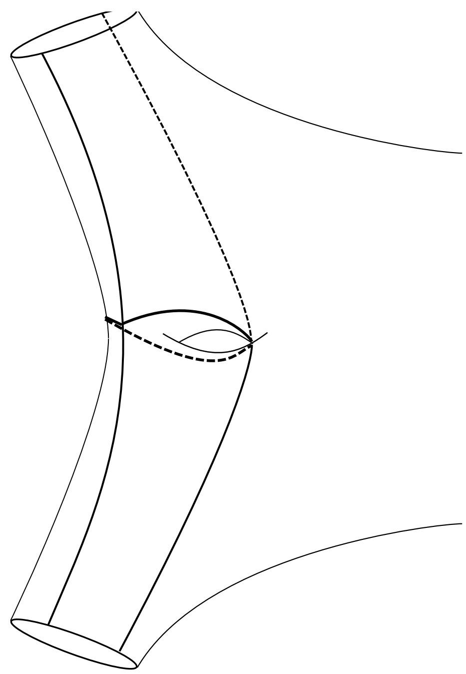

FIGURE 1. The systole η and its copy $\sigma(\eta)$ used to construct γ

Denote by M_c, resp. $M_{\sigma(c)}$, the midpoint of c, resp $\sigma(c)$. Let d be the (or a) shortest path between M_c and $M_{\sigma(c)}$. Now $\sigma(d)$ is also a path between the same two points. If $\sigma(d) = d$, then the midpoint of d is a fixed point of σ, a contradiction. In fact, d and $\sigma(d)$ do not transversally intersect, for the same reason that c and $\sigma(c)$ do not transversally intersect. Thus $d \cap \sigma(d) = \{M_c, M_{\sigma(c)}\}$. The closed curve $d \cup \sigma(d)$ is thus a simple non-trivial closed curve, and by the same reasoning as in the case of the surface with at least two boundary curves, is not homotopic to either η or $\sigma(\eta)$. We now take γ to be its unique geodesic representative. Clearly, $\sigma(\gamma) = \gamma$.

The process that follows is very similar to our construction of γ in what precedes. For that reason, we reset all our notations, with the obvious exceptions of S and σ, and with the exception of γ. We now construct $\overline{\Gamma} := \mathrm{Fix}(\sigma) \cup \{\gamma_0, \dots, \gamma_n\}$ recursively. We set $\gamma_0 := \gamma$. For $k > 0$, take γ_{k-1}^1 and γ_{k-1}^2, the two copies of γ_{k-1} on $S_{k-1} := S \setminus \mathrm{Fix}(\sigma) \setminus \{\gamma_0, \dots, \gamma_{k-1}\}$. Consider c_k the shortest path between γ_{k-1}^1

and γ_{k-1}^2 on S'. Now $\sigma(c_k)$ is also a shortest path between γ_{k-1}^1 and γ_{k-1}^2, which for reasons which should be apparent, satisfies $c_k \cap \sigma(c_k) = \emptyset$. Now consider M_{c_k} and $M_{\sigma(c_k)}$ the midpoints of c_k and $\sigma(c_k)$. Take d_k to be a shortest path between M_{c_k} and $M_{\sigma(c_k)}$. As before, we obtain a globally invariant non-trivial simple closed curve $d_k \cup \sigma(d_k)$, and we take γ_k to be its unique geodesic representative. Note that γ_k is both disjoint and distinct from the curves $\gamma_0, \ldots, \gamma_{k-1}$. (The distinctness follows once again from the fact that σ is orientation reversing, and $d_k \cup \sigma(d_k)$ cannot bound an invariant cylinder or cannot be interior to an invariant cylinder). This process is continued until the set $\overline{\Gamma}$ is separating.

Denote by n the cardinality of the set $\tilde{\Gamma}$ and consider the surface $\tilde{S} := S \setminus \mathrm{Fix}\sigma$. By proposition 3.5, the set $\tilde{\gamma}$ separates $\tilde{S}$ into two isometric, and thus homeomorphic surfaces $\tilde{S}_1$ and $\tilde{S}_2$. Denote their underlying genus by $\tilde{g}$. Now clearly $g = 2\tilde{g} + n - 1$, so g and n are necessarily of different parity, and $n \leq g + 1$.

Let us now explain the procedure to replace an odd number $l \geq 3$ of curves, subset of $\tilde{\Gamma}$, by a single curve δ such that the new set is still both separating and each curve in the set is globally invariant (see also the proof of Lemma 1.2 and the Figure 2.1.2 in [9]). Denote the set of curves we aim to replace by $\gamma_1, \ldots, \gamma_l$. These set of curves separate a subsurface (which we denote $\overline{S}$) of $\tilde{S}$ obtained by cutting along $\tilde{\Gamma} \setminus \{\gamma_1, \ldots, \gamma_l\}$. Denote by $\overline{S}_1$ and $\overline{S}_2$ the two subsurfaces of $\overline{S}$ separated by $\{\gamma_1, \ldots, \gamma_l\}$. For each γ_k, fix two diametrically opposite points, say p_k and q_k. The goal is now to construct a set of disjoint simple paths $c_1, \ldots, c_l \subset \overline{S}_1$ between the points $\{p_1, q_1, \ldots, p_l, q_l\}$. On $\overline{S}_1$, consider some simple path c_1 between q_1 and p_2. By cutting $\overline{S}_1 \setminus c_1$ is still connected and has one less boundary curve than $\overline{S}_1$. On this new surface, consider a path c_2 between q_2 and p_3. Continue this process until c_{l-1} is constructed between q_{l-1} and p_l. This is of course possible because at step k, the surface $\overline{S}_1 \setminus \{c_1, \ldots, c_k\}$ is still connected. Finally, we construct c_l between q_l and p_1 on $\overline{S}_1 \setminus \{c_1, \ldots, c_{l-1}\}$. The path c_l is different from the others, in that it is a path that joins a common boundary curve on $\overline{S}_1 \setminus \{c_1, \ldots, c_{l-1}\}$. It follows that $\overline{S}_1 \setminus \{c_1, \ldots, c_l\}$ is no longer connected. Denote by $\overline{S}_{11}$ and $\overline{S}_{12}$ the two subsurfaces of $\overline{S}_1$ separated by $\{c_1, \ldots, c_l\}$. Now consider the set of paths $\{\sigma(c_k)\}_{k=1,\ldots,l\}}$ of $\overline{S}_2$. This set of paths separated $\overline{S}_2$ into two subsurfaces $\overline{S}_{21} := \sigma(\overline{S}_{11})$ and $\overline{S}_{22} := \sigma(\overline{S}_{12})$. Furthermore, for $k \in \{1, \ldots, l-1\}$, the path $\sigma(c_k)$ joins the points p_k and q_{k+1}, and $\sigma(c_l)$ joins q_1 and p_l. Consider the point set $\psi := \{c_1, \ldots, c_l\} \cup \{\sigma(c_1), \ldots, \sigma(c_l)\}$. Because l is odd, it is not difficult to see that ψ is a simple closed curve, and by construction ψ is separating and invariant by σ.

In the event where $\tilde{g} > 0$, the set $\tilde{\Gamma}$ can be replaced by a set with greater cardinality by using the following procedure. For any choice of $\gamma \in \tilde{\Gamma}$, consider a simple non-trivial path c contained in $\tilde{S}_1$ between two diametrically opposite points such that $\tilde{S}_1 \setminus c$ is still connected. (This is possible because $\tilde{g} > 0$.) Now $c \cup \sigma(c)$ is a simple closed curve, invariant by σ. Denote by $\tilde{\gamma}$ its geodesic representative. As both $\tilde{S}_1 \setminus c$ and $\tilde{S}_2 \setminus \sigma(c)$ are connected, it follows that the set $\tilde{\Gamma} \cup \tilde{\gamma} \setminus \gamma$ no longer

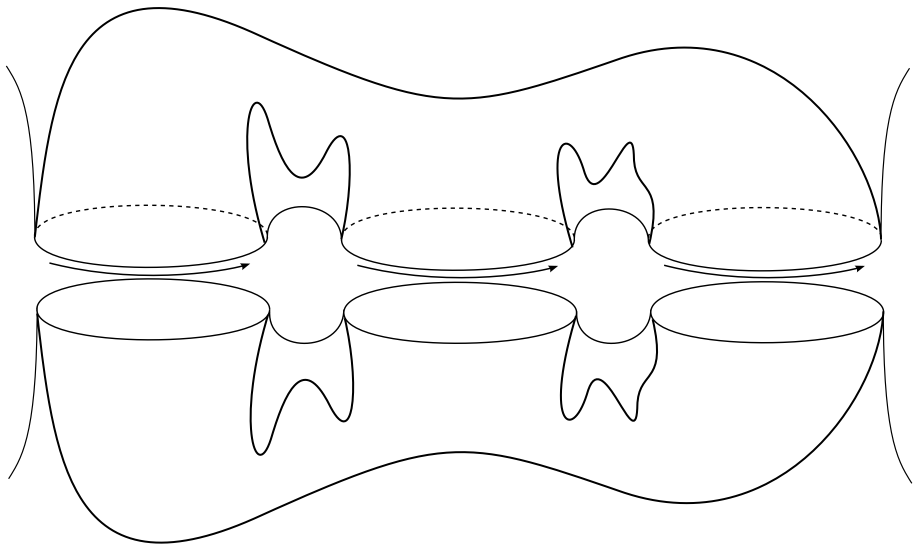

FIGURE 2. The procedure for reducing the number of disjoint invariant curves. The arrows around the curves γ_k represent half-twists.

disconnects $\tilde{S}$. Therefore this new set can be completed by the technique used to construct $\tilde{\gamma}$. As the process increases cardinality, this proves the claim.

Now as any n can be decreased with precision to any number of correct parity, and can be increased in the event where $n < g+1$, this proves the second part of the theorem. □

4. Application to the study of non-euclidean crystallographic groups

Let $\mathcal{D}$ be the complex unit disc and $\mathcal{G} = \mathrm{Aut}(\mathcal{D})$ be the group of conformal and anticonformal automorphisms of $\mathcal{D}$. A discrete, cocompact subgroup Γ of $\mathrm{Aut}(\mathcal{D})$ is called an *NEC (Non-Euclidean Crystallographic) group*. The subgroup of Γ consisting of the orientation-preserving elements is called the *canonical Fuchsian subgroup of* Γ.

The algebraic structure of an NEC group and the geometric and topological structure of its quotient orbifold are given by the signature of Γ:

$$s(\Gamma) = (h; \pm; [m_1, ..., m_r]; \{(n_{11}, ..., n_{1s_1}), ..., (n_{k1}, ..., n_{ks_k})\}). \tag{1}$$

The orbit space $\mathcal{D}/\Gamma$ is an orbifold with underlying surface of genus h, having r cone points and k boundary components, each with $s_j \geq 0$ corner points, $j = 1, ..., k$. The signs $''+''$ and $''-''$ correspond to orientable and non-orientable orbifolds respectively. The integers m_i are called the proper periods of Γ and they are the orders of the cone points of $\mathcal{D}/\Gamma$. The brackets $(n_{i1}, ..., n_{is_i})$ are the period cycles of Γ. The integers

n_{ij} are the link periods of Γ and the orders of the corner points of $\mathcal{D}/\Gamma$. The group Γ is called the *fundamental group* of the orbifold $\mathcal{D}/\Gamma$.

For an NEC group of signature (1), a *canonical presentation* with four types of generators is given in [8] and [13]:

1.*Hyperbolic generators* $a_1, b_1, ..., a_h, b_h$ if $\mathcal{D}/\Gamma$ is orientable or *glide reflection generators* $d_1, ..., d_h$ if $\mathcal{D}/\Gamma$ is non-orientable,

2. *Elliptic generators*: $x_1, ..., x_r$,

3. *Connecting generators* (hyperbolic or elliptic transformations): $e_1, ..., , e_k$,

4. *Reflection generators*: c_{ij}, $1 \leq i \leq k, 1 \leq j \leq s_i + 1$.

And relators:

1. $x_i^{m_i}, i = 1, ..., r$,
2. c_{ij}^2,
3. $(c_{ij-1}c_{ij})^{n_{ij}}$,
4. $c_{i0}e_i^{-1}c_{is_i}e_i, i = 1, ..., k, j = 2, ..., s_i + 1$,
5. The long relation: $x_1 \cdots x_r e_1 \cdots e_k a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_h b_h a_h^{-1} b_h^{-1}$ or $x_1 \cdots x_r e_1 \cdots e_k d_1^2 \cdots d_h^2$, according to whether $\mathcal{D}/\Gamma$ is orientable or not.

The canonical presentation is a very useful tool in the study of NEC groups. Using Theorem 3.9, we can modify the canonical presentation to obtain some other presentations that can be used in the study of automorphisms of Riemann and Klein surfaces.

Proposition 4.1. *Let Γ be an NEC group with signature*

(2) $$s(\Gamma) = (h; -; [m_1, ..., m_r]; \{(n_{11}, ..., n_{1s_1}), ..., (n_{k1}, ..., n_{ks_k})\}).$$

Let l be an integer such that $1 \leq l \leq h + 1$ and $h + l \equiv 1 \mod 2$. The group Γ has a presentation with generators:

1. Glide reflection generators: $d_1, ..., d_l$,

2. Hyperbolic generators $a_1, b_1, ..., a_{h-l-1}, b_{h-l-1}$,

3. Elliptic generators: $x_1, ..., x_r$,

4. Connecting generators (hyperbolic or elliptic transformations): $e_1, ..., , e_k$,

5. Reflection generators: c_{ij}, $1 \leq i \leq k, 0 \leq j \leq s_i$.

And relations:

1. $x_i^{m_i}, i = 1, ..., r$,

2. c_{ij}^2,

3. $(c_{ij-1}c_{ij})^{n_{ij}}$,

4. $c_{i0}e_i^{-1}c_{is_i}e_i, i = 1, ..., k, j = 2, ..., s_i + 1$,

5. The long relation: $x_1 \cdots x_r e_1 \cdots e_k d_1^2 \cdots d_l^2 a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_{h-l-1} b_{h-l-1} a_{h-l-1}^{-1} b_{h-l-1}^{-1}$.

Proof. We shall do the proof only in the case when Γ has signature:

$$(h; -; [-]; \{(-), \overset{k}{\dots}, (-)\}),$$

because this situation follows easily from Theorem 3.9. The general case has several technical complications which would make the proof much longer (although a similar proof would be applicable in the remaining cases as well) and as this is not the main goal of the article, we leave this to the motivated reader. Let Γ^+ be the canonical Fuchsian subgroup of Γ, then $\mathcal{D}/\Gamma^+$ is a Riemann surface admitting an anticonformal involution σ. By Theorem 3.9, we can obtain a set Γ_l of l closed geodesics invariant by σ, and such that $\mathrm{Fix}(\sigma) \cup \Gamma_l$ separates $\mathcal{D}/\Gamma^+$. Then we can consider the quotient $(\mathcal{D}/\Gamma^+ - \Gamma_l)/\langle\sigma\rangle$, that is an orientable surface of genus $h-l-1$ with $k+l$ boundary components. Then $(\mathcal{D}/\Gamma^+ - \Gamma_l)/\langle\sigma\rangle$ can be uniformized by an NEC group with signature $(h-l-1; -; [-]; \{(-), \overset{k+l}{\dots}, (-)\})$. A canonical set of generators is

$$a_1, b_1, ..., a_{h-l-1}, b_{h-l-1},\ e_1, ..., , e_{k+l},\ c_{i0},\ 1 \le i \le k+l,$$

satisfying the relations:
$c_{i0}^2,\ c_{i0}e_i^{-1}c_{i0}e_i,\ 1 \le i \le k+l$ and
$e_1...e_{k+l}a_1b_1a_1^{-1}b_1^{-1}...a_{h-l-1}b_{h-l-1}a_{h-l-1}^{-1}b_{h-l-1}^{-1}.$

By the first type of relations, the hyperbolic transformation e_i has as axis the set of fixed points of c_{i0}. Let e'_j, $j = k+1, ..., k+l$, be the hyperbolic transformations with the same axis as e_j and such that $e'^2_j = e_j$. We define the glide reflection $d_j = e'_j c_{j0}$, $j = k+1, ..., k+l$.

The group:

$$\Gamma' = \left\langle \begin{array}{c} a_1, b_1, ..., a_{h-l-1}, b_{h-l-1}, e_1, ..., , e_k, d_1, ..., d_l, c_{10}, ..., c_{k0} : \\ c_{i0}^2, c_{i0}e_i^{-1}c_{i0}e_i, 1 \le i \le k+l, \\ e_1 \cdots e_k d_1^2 \cdots d_l^2 a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_{h-l-1} b_{h-l-1} a_{h-l-1}^{-1} b_{h-l-1}^{-1} \end{array} \right\rangle$$

uniformizes $\mathcal{D}/\Gamma$ and thus Γ and Γ' are conjugate, and thus Γ admits a presentation as the one defining Γ'. □

5. Application to the study of moduli space of real algebraic curves

A *real* Riemann surface is a Riemann surface with an anticonformal involution σ. Real Riemann surfaces and real algebraic curves are equivalent objects. Given a real Riemann surface X of genus g, we shall denote by $\pm k$ the topological type of the action of the anticonformal involution σ on X, where k is the number of connected components of the fixed point set $\mathrm{Fix}(\sigma)$ and $\pm$ tell us the fact that $\mathrm{Fix}(\sigma)$ disconnects or not X.

Let $\mathcal{M}_g$ be the moduli space of Riemann surfaces of genus g. Let $\mathcal{M}_g^{\pm k} \subset \mathcal{M}_g$ be the set of points corresponding to real Riemann surfaces with a fixed topological type $\pm k$. It is well-known that the space $\mathcal{M}_g^{\pm k}$ is connected (see for example [3], [4] and [1]). We can now show the following proposition which is implicit in the proof of Theorem 7.5 of [11].

Proposition 5.1. *$\bigcap_{k=0}^{g} \overline{\mathcal{M}_g^{-k}} \cap \overline{\mathcal{M}_g^{+(g+1)}} \neq \emptyset$ and $\overline{\mathcal{M}_g^{-k}} \cap \overline{\mathcal{M}_g^{+(k+l)}} \neq \emptyset$, where $l > 0$ and $k + l + g \equiv 1 \mod 2$.*

Proof. Let S be a Riemann surface in $\overline{\mathcal{M}_g^{-k}}$, $k = 0, ..., g$, let σ be the anticonformal involution of S. Applying Theorem 3.9 we can construct a set Γ_{k-g-1} of $k-g-1$ closed geodesics that are invariant by σ and such that $\text{Fix}(\sigma) \cup \Gamma_{k-g-1}$ separates S in two planar surfaces S' and S''. Let Γ' be a set of closed geodesics such that $\text{Fix}(\sigma) \cup \Gamma_{k-g-1} \cup \Gamma'$ produces a pants decomposition of S'. Now, we can collapse continuously each curve in $\text{Fix}(\sigma) \cup \Gamma_{k-g-1} \cup \Gamma'$ to a point and, using σ, extend such a deformation to a deformation of a pants decomposition of S invariant by σ. The limit surface $\overline{S}$ by this deformation is independent of k. Then $\overline{S} \in \bigcap_{k=0}^{g} \overline{\mathcal{M}_g^{-k}}$. By the same construction $\overline{S} \in \overline{\mathcal{M}_g^{+(g+1)}}$.

Similarly, using Theorem 3.9 again, it is possible to construct $\overline{S_k} \in \overline{\mathcal{M}_g^{-k}} \cap \overline{\mathcal{M}_g^{+(k+l)}}$. □

Now let C_1, C_2 be two (smooth) real algebraic curves, both represented as zeros of a set of real polynomials. We can continuously modify the real coefficients of such polynomials, preserving the real character of the corresponding curve, to arrive to a set of polynomials defining C_2, in the way it is possible that the set of polynomials define for some values singular curves. As an immediate consequence of proposition 5.1 we obtain the following.

Corollary 5.2. *If the topological types of the two curves C_1 and C_2 are contained in the set $\{-g, ..., -1, 0, g+1\}$, it is possible to perform a real deformation from C_1 to C_2 passing through only one singular curve. In the general case it is possible to construct the deformation passing through at most two singular curves.*

□

References

[1] P. Buser, M. Seppälä, and R. Silhol. Triangulations and moduli spaces of Riemann surfaces with group actions. *Manuscripta Math.*, 88(2):209–224, 1995.

[2] Peter Buser. *Geometry and spectra of compact Riemann surfaces*, volume 106 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1992.

[3] A. F. Costa and M. Izquierdo. On the locus of real algebraic curves. *Atti Sem. Mat. Fis. Univ. Modena*, 49(suppl.):91–107, 2001. Dedicated to the memory of Professor M. Pezzana (Italian).

[4] Clifford J. Earle. On the moduli of closed Riemann surfaces with symmetries. In *Advances in the theory of riemann surfaces (Proc. Conf., Stony Brook, N.Y., 1969)*, pages 119–130. Ann. of Math. Studies, No. 66. Princeton Univ. Press, Princeton, N.J., 1971.

[5] Axel Harnack. Ueber die Vieltheiligkeit der ebenen algebraischen Curven. *Math. Ann.*, 10(2):189–198, 1876.

[6] Felix Klein. Über eine neue Art von Riemannschen Flächen, Zweite Mitteilung. *Math. Ann.*, 10(2):136–155, 1876.

[7] Irwin Kra and Bernard Maskit. Bases for quadratic differentials. *Comment. Math. Helv.*, 57(4):603–626, 1982.

[8] A. M. Macbeath. The classification of non-euclidean plane crystallographic groups. *Canad. J. Math.*, 19:1192–1205, 1967.

[9] S. M. Natanzon. *Moduli of Riemann surfaces, real algebraic curves, and their superanalogs*, volume 225 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2004. Translated from the 2003 Russian edition by Sergei Lando.

[10] P. Schmutz. Riemann surfaces with shortest geodesic of maximal length. *Geom. Funct. Anal.*, 3(6):564–631, 1993.

[11] M. Seppälä. Moduli spaces of stable real algebraic curves. *Ann. Sci. École Norm. Sup. (4)*, 24(5):519–544, 1991.

[12] G. Weichold. *Ueber symmetrische Riemann'sche Flchen und die Periodicittsmoduln der zugehrigen Abel'schen Normalintegrale erster Gattung*. PhD thesis, Diss. Leipzig. Schlömilch Z. XXVIII. 321-352. , 1883.

[13] H. C. Wilkie. On non-Euclidean crystallographic groups. *Math. Z.*, 91:87–102, 1966.

Departamento de Matemáticas Fundamentales, Facultad de Ciencias, Universidad Nacional de Educación a Distancia, Madrid 28040, SPAIN

E-mail address: acosta@mat.uned.es

Section de Mathématiques, École Polytechnique Fédérale de Lausanne, SB-IGAT, BCH, CH-1015 Lausanne, Switzerland

E-mail address: hugo.parlier@epfl.ch