university of groningen

faculty of science
and engineering

The Representation of Hyperelliptic Surfaces by Admissible Polygons

Master Project Mathematics

November 2020

Student: T.M. van Haeringen

First supervisor: Prof.dr. G. Vegter

Second supervisor: dr. A.E. Sterk

ABSTRACT

Closed hyperelliptic Riemann surfaces can be represented by a class of polygons in the hyperbolic plane, called admissible or hyperelliptic polygons. In this thesis we study this relation, and derive conditions on the vertices of a hyperbolic polygon as to be admissible. Furthermore, we apply our representation of a closed hyperelliptic Riemann surface M by an admissible polygon to solve a specific problem: given a non-contractible closed curve α on M, we provide algorithms to calculate the unique closed geodesic β homotopic to α, as well as a homotopy from α to β.

CONTENTS

Part I

INTRODUCTION AND STUDY

1 INTRODUCTION

One of the core concepts of topology and geometry is that we may construct surfaces by gluing together sides of a polygon in the plane. A basic example is that of the torus, which we can construct by identifying opposite sides of a rectangle in the Euclidean plane. This allows us, for instance, to study curves on the torus by studying its corresponding segments in this rectangle.

The main topic of this thesis is the description of a particular class of surfaces, the closed hyperelliptic (Riemann) surfaces. We may describe each hyperelliptic surface by a polygon in the hyperbolic plane. Such a polygon will belong to a subset of hyperbolic polygons that are called *admissible* or *hyperelliptic* (see figure 1. The relation between hyperelliptic surfaces and admissible polygons was first shown by Schaller [7]. If a hyperelliptic surface is provided with a bit of extra structure, a so-called marking, this relation is one-to-one. The coordinates of $2g-1$ vertices of such polygons can be used to parametrize the space of marked hyperelliptic surfaces. In this thesis we will describe the space of possible vertices in the hyperbolic disk for admissible polygons. This will build on results from [1] on closed Riemann surfaces of genus 2, which are all hyperelliptic.

Furthermore, we will look at a specific computational problem on hyperelliptic Riemann surfaces. Each non-contractible closed curve on a Riemann surface is homotopic to a unique closed geodesic. We provide an algorithm that, given a non-contractible closed curve, computes this closed geodesic, as well as an algorithm that computes a homotopy from the closed curve to the closed geodesic explicitly.

In Chapter 2 we recall some necessary background. We will briefly describe some of the basics on hyperbolic geometry, Riemann surfaces and Fuchsian groups, and spend some time on markings and Teichmüller space. In Chapter 3 we will move on to hyperelliptic Riemann surfaces. We will define admissible polygons and hyperellipticity, and provide an exposition of

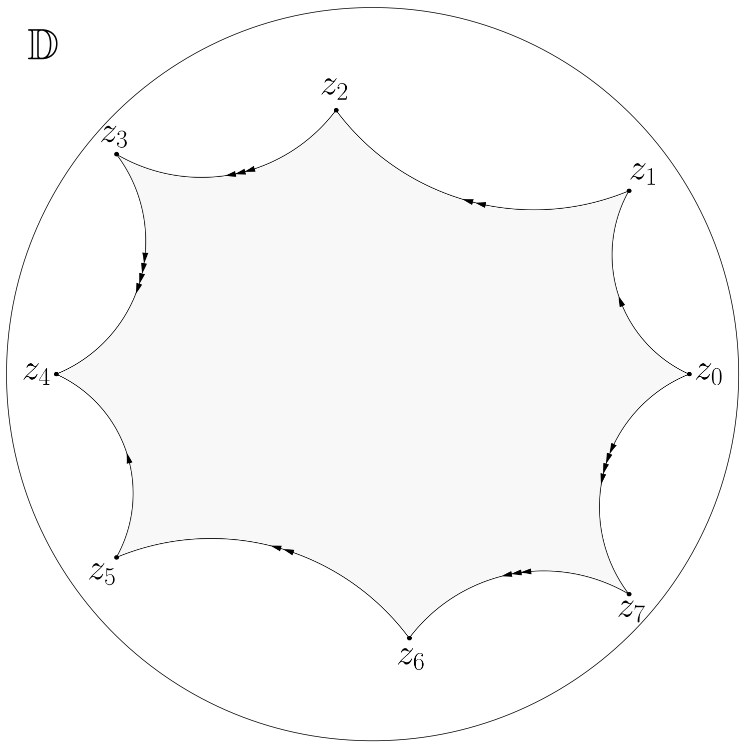

Figure 1: An admissible octagon in the Poincaré disk model the hyperbolic plane. By identifying its opposite sides as indicated by the arrows, we obtain a closed hyperelliptic surface of genus 2.

the results of Schaller which relate the two. In Chapter 4, we investigate the space of admissible polygons, and find a condition on the vertices of a symmetric polygon as to be admissible. In Chapter 5 we will construe algorithms for finding the homotopic closed geodesic of a curve on a hyperelliptic surface, and moreover find a homotopy. Finally, we present a discussion of the thesis and its results in Chapter 6.

2 BACKGROUND

We will first briefly describe the setting in which the questions of this thesis appear, that of hyperbolic geometry. For a similar brief overview see [1] as well, and for a more detailed discussion see for instance [2] and [3].

2.1 THE HYPERBOLIC PLANE

We use the Poincaré disk model of two dimensional hyperbolic space. It consists of the points in the unit disk

$$\mathbb{D} = \{z \in \mathbb{C} | |z| < 1\}$$

together with the metric

$$ds^2 = \frac{4(dx^2 + dy^2)}{(1 - x^2 - y^2)^2}.$$

The geodesics in this space consist of circular arcs which meet the boundary of $\mathbb{D}$ perpendicularly. The isometries of the Poincaré disk which preserve orientation are the maps

$$\left\{ z \mapsto \frac{az + b}{\bar{b}z + \bar{a}} \mid a, b \in \mathbb{C}, |a|^2 - |b|^2 = 1 \right\},$$

on $\mathbb{D}$, where we let $\bar{c}$ denote the complex conjugate of a complex number c. These transformations form a subgroup of the group of Möbius transformations of the extended complex plane.

We can classify the orientation preserving isometries which are not the identity by their fixed points in $\mathbb{D}$. Those with one fixed point in the interior of $\mathbb{D}$ (and hence none on the boundary) are called *elliptic*. Those with a double fixed point on the boundary are called *parabolic*. Those with two distinct fixed points on the boundary interior are called *hyperbolic*.

A hyperbolic transformation ϕ has an attracting fixed point ξ_+ and a repelling fixed point ξ_-, in the sense that $\lim_{n \to \pm\infty} \phi^n(z) = \xi_\pm$ for any $z \in \mathbb{D}$. The geodesic between the two fixed points of a hyperbolic transformation ϕ is called the *axis* A_ϕ of the transformation. Applying a hyperbolic transformation to a point in $\mathbb{D}$ will move it along an equidistant of the axis. In particular, the distance $d(z, \phi(z))$ is minimal if and only if z lies on A_ϕ.

2.2 HYPERBOLIC POLYGONS

We define an n-*sided hyperbolic polygon* P, or *hyperbolic* n-*gon*, to be a region in $\mathbb{D}$ bounded by n geodesic segments (called *sides* or *edges*), which do not mutually intersect except at their endpoints (called *vertices*).

Unlike the Euclidean case, the area of a hyperbolic polygon is fully determined by the interior angles at its interior vertices. For an n-sided hyperbolic polygon P with internal angles $\alpha_0, \dots, \alpha_{n-1}$, we have the following:

$$\text{area}\, P = (n-2)\pi - \sum_{i=0}^{n-1} \alpha_i. \tag{1}$$

We now consider a hyperbolic polygon P with $4g$ sides $b_0, \dots, b_{4g-1}$ and internal angles $\alpha_0, \dots, \alpha_{4g-1}$, $g \in \mathbb{N}, g \geqslant 2$ which satisfies

- The lengths of opposite sides are equal, i.e. $\ell(b_i) = \ell(b_{i+2g})$, $i = 0, \dots, 2g-1$,
- The internal angles of P sum to 2π, i.e. $\sum_{i=0}^{4g-1} \alpha_i = 2\pi$.

If we now glue together its opposite sides as in figure 2, we will obtain a closed Riemann surface M of genus g. The hyperbolic isometries that send a side to its opposite side, called the *side-pairings*, generate a discrete subgroup of the group of isometries on $\mathbb{D}$, and are called the *Fuchsian group* of M, which we will denote by Γ. By Poincaré's polygon theorem [6], the Riemann surface M is then quotient space of $\mathbb{D}$ under the action of the Fuchsian group Γ, i.e. $M = \mathbb{D}/\Gamma$.

Conversely, on every compact Riemann surface M of genus $g \geqslant 2$ we can construct $2g$ simple geodesic loops with the same base point, and obtain a hyperbolic $4g$-gon if we cut open M along them.

Throughout the rest of this thesis, we will use the term surface for a compact Riemann surface of genus $g \geqslant 2$.

2.3 MARKINGS AND TEICHMÜLLER SPACE

We have noted that we can construct any closed surface of genus $g \geqslant 2$ by gluing the sides of a hyperbolic polygon. This

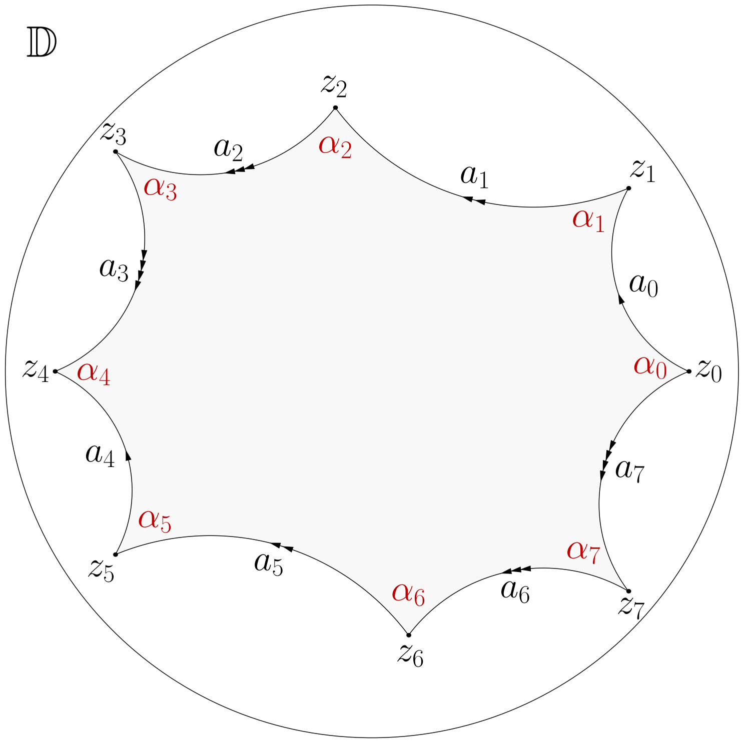

Figure 2: A hyperbolic polygon P whose internal angles sum to 2π and whose opposite sides have equal (hyperbolic) length.

correspondence is not one-to-one however; many hyperbolic polygons correspond to the same surface. The (moduli) space of all closed surfaces of genus g turns out to be difficult to study directly. We will find it easier to study the space of *marked* (Riemann) surfaces.

Definition 1 *Let* F *be a fixed orientable compact smooth surface of genus* g. *A marked (Riemann) surface* (M, ϕ) *is a compact Riemann surface* M *together with a homeomorphism* $\phi : F \to M$, *called a marking homeomorphism.*

Definition 2 *Two marked surfaces* (M_1, ϕ_1) *and* (M_2, ϕ_2) *are said to be marking equivalent if there exists an isometry* $\chi : M_1 \to M_2$ *such that* $\chi \circ \phi_1$ *and* ϕ_2 *are isotopic. Teichmüller space* $\mathcal{T}_g$ *is the set of all equivalence classes of marked surfaces of genus* g.

Here isotopy is a homotopy between homeomorphisms whose "intermediate maps" are also homeomorphisms, in the following sense:

Definition 3 *Let* M, M' *be two surfaces, and* ϕ_0, ϕ_1 *homeomorphisms from* M *to* M'. *Then* ϕ_0 *and* ϕ_1 *are isotopic if there exists a continuous mapping* $J : [0, 1] \times M \to M'$ *such that*

1. *J is a homotopy from ϕ_0 to ϕ_1,*
2. *For any $s \in [0,1]$, the map $J_s : M \to M'$ defined by $J_s(t) = J(s,t)$ is a homeomorphism.*

A different way to define a marking on a Riemann space is through a set of labeled curves, sometimes called a *star set*.

Definition 4 *Let M be closed surface of genus g. A star set on M is an ordered set $S = \{\gamma_1, \ldots, \gamma_{2g}\}$ of geodesic loops on M which all mutually intersect at a single point $p \in M$, and only at that point. That is $\gamma_i \cap \gamma_j = \{p\}$ for all $i \neq j$. We will call S a geodesic star set if γ_1 and γ_2 are closed geodesics.*

Note that there is a distinction between a geodesic loop and a closed geodesic. A geodesic loop need not be smooth at its base point, while a closed geodesic is. Zieschang, Vogt and Coldewey showed that a marking equivalence class on a marked surface M corresponds uniquely to a geodesic star set defined on M (see [3], Section 6.7). This gives us the following characterization of Teichmüller space:

Theorem 5 *Let $\mathcal{S}_g$ be the space of closed surfaces of genus $g \geqslant 2$ with a geodesic star set defined on it. Then there is a one-to-one correspondence between $\mathcal{S}_g$ and Teichmüller space $\mathcal{T}_g$.*

Teichmüller space $\mathcal{T}_g$ has a natural real analytic structure, homeomorphic to $\mathbb{R}^{6g-6}$ (see [3], Section 6.3). So in particular, we need $6g-6$ real parameters to specify a marked surface in $\mathcal{T}_g$.

For a more in depth look at Teichmüller spaces, see for instance [3], [4] and [5].

3 A STUDY OF THE RELATION BETWEEN ADMISSIBLE POLYGONS AND HYPERELLIPTIC SURFACES

Now that we have recalled the necessary background, it is time to delve into the main topic of this thesis.

3.1 ADMISSIBLE POLYGONS

We have introduced Teichmüller space, and seen a few ways in which we can describe it. In their article 'Hyperbolic octagons and Teichmüller space in genus 2' [1], Aigon-Dupuy et al. investigate a parametrization of Teichmüller space of genus 2 by a certain set of octagons, those that they call *admissible*:

Definition 6 *Let* P *be a hyperbolic octagon in* $\mathbb{D}$ *with vertices* $z_0, \ldots, z_7$*, which satisfies*

1. z_0 *lies in the interval* $(0, 1)$ *on the real axis,*
2. $z_{k+4} = z_k$*,* $k = 0, 1, 2, 3$*, i.e. the polygon is symmetric.*

We say P *is admissible if the sum of its internal angles is* 2π*.*

See figure 3. Note that since P is symmetric, its opposite sides have equal length. So if it is admissible, we can glue together opposite sides as in figure 3 to obtain a (compact Riemann) surface of genus 2. If we label the sides of the polygon, they will give us a marking on this surface as well.

It has long been known that such admissible octagons are in one-to-one correspondence with Teichmüller space of genus 2. Aigon-Dupuy et al. show how we may choose vertices in $\mathbb{D}$ in such a way that the resulting octagon is admissible, and how this description relates to a few other descriptions of Teichmüller space. Their main goal in doing so is to provide tools for computation and numerical experimentation on compact Riemann surfaces of genus 2.
Our question is the following: can we extend the results of Aigon-Dupuy to closed surfaces higher genus? We look at the straightforward generalization of admissible polygons in higher genus:

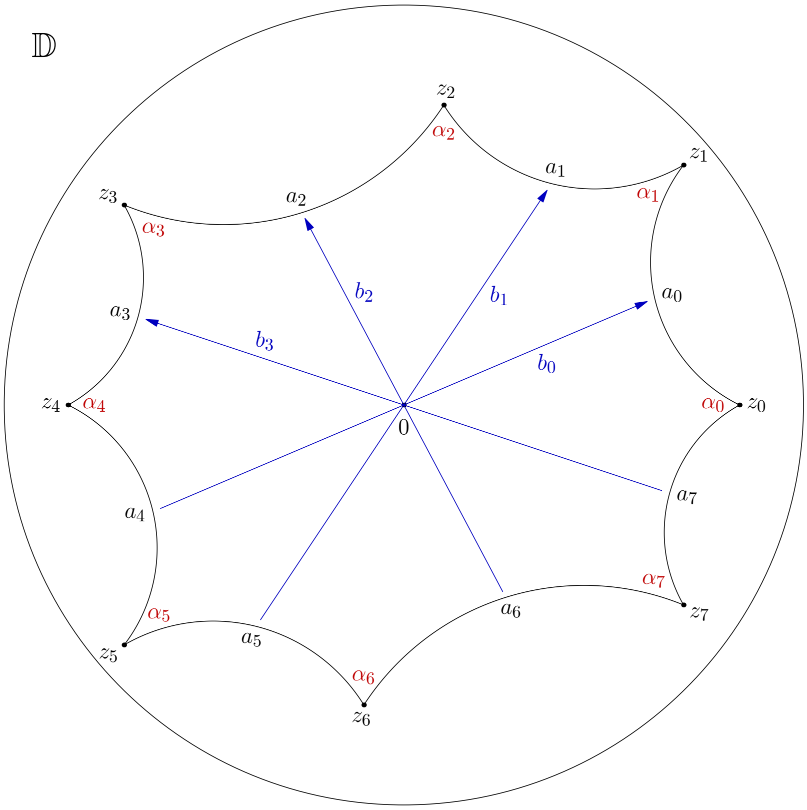

Figure 3: An admissible octagon.

Definition 7 *Let* $g \in \mathbb{N}, g \geqslant 2$*, and let* P *be a* $4g$*-gon in the hyperbolic disk with vertices* $z_0, \ldots, z_{4g-1}$*, which satisfies*

1. z_0 *lies in the interval* $(0, 1)$ *on the real axis,*
2. $z_{k+2g} = z_k$, $k = 0, \ldots, 2g-1$*, i.e. the polygon is symmetric.*

We say P *is admissible if the sum of its internal angles is* 2π*.*

The first thing to note is that these polygons cannot parametrize Teichmüller space in general: each such polygon is fully determined by its $2g$ vertices, which together with the restriction on z_0 and the internal angles gives us only $4g-2$ real parameters, while Teichmüller space needs $6g-6$. Only in the case that $g = 2$ these coincide.

Instead, what we will find is that these admissible polygons correspond to the hyperelliptic surfaces[1]. This was first shown by Schaller [7]. All closed (Riemann) surfaces of genus 2 are hyperelliptic, which is why admissible octagons can be used to parametrize the entirety of Teichmüller space for genus two. We will present Schaller's result in detail in the next section.

1 For this reason, they are sometimes referred to as *hyperelliptic polygons*.

3.2 CORRESPONDENCE TO HYPERELLIPTIC SURFACES

As is the case for Riemann surfaces in general, there exist many equivalent definitions of hyperellipticity. Some are more algebraic in nature, some more geometric. As we take a geometric approach to surfaces in this thesis, it makes sense to take a geometric definition. We use the following, for compact Riemann surfaces:

Definition 8 *Let* M *be a compact Riemann surface of genus* $g \geqslant 2$*. We say* M *is hyperelliptic if there exists an isometry* $\phi : M \to M$ *which has exactly* $2g+2$ *fixed points on* M *and* $\phi^2 = \mathrm{id}$ *(the identity on* M*).*

Such an isometry ϕ is called an *involution*. We will look at the following result of Schaller [7]:

Theorem 9 *Let* M *be a closed Riemann surface of genus* g*. The following are equivalent:*

1. M *is hyperelliptic.*
2. *There exists a set of at least* $2g-2$ *simple closed geodesics that all meet in one point, and nowhere else.*
3. *There exists a set of* $2g$ *simple closed geodesics on* M *that all meet in one point, and nowhere else.*
4. M *corresponds to an admissible polygon, i.e. there is an admissible polygon* P *such that identifying its opposite sides gives us* M*.*

Proof. We will prove $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1$.
$(1 \Rightarrow 2)$: Suppose M is hyperelliptic. Then by definition there exists an isometry $\phi : M \to M$ with fixed points $A_1, \ldots, A_{2g+2}$ and $\phi^2 = \mathrm{id}$. Define h_1 to be a simple geodesic from A_1 to A_2. Since ϕ is smooth and A_1, A_2 are fixed points, $u_1 = h_1 \cap \phi(h_1)$ is a closed geodesic. We note the following:

- u_1 does not contain any fixed points of ϕ other than A_1 and A_2, as this would imply there exists a non-trivial Möbius transformation with 3 fixed points. Note that in particular, this implies that u_1 is simple.
- $M \setminus u_1$ is connected; if it were not, then ϕ would have to map its two connected components either both to themselves or both to the other component. Since ϕ is locally

a π-rotation at its fixed points, it must be that it maps each component to the other. However, since there are still other fixed points of ϕ which do not lie on u_1, this cannot be the case.

So we have constructed a simple closed geodesic on M using fixed points A_1 and A_2. We will continue in the same vein, using the other fixed points.
First, we cut open the surface M along u_1 to obtain $M_1 = M \setminus u_1$, which will be a $(g-1, 2)$-surface. Now we take h_2 to be a simple geodesic from a copy of A_1 on one of the boundaries to A_3. Then $u_2 = h_2 \cap \phi(h_2)$ is again a closed geodesic on M. Moreover, since h_2 only meets u_1 in A_1, $\phi(h_2)$ will only meet $\phi(u_1) = u_1$ in A_1, so $u_1 \cap u_2 = A_1$.

If we cut open M_1 across u_2, we obtain a $(g-1, 1)$-surface M_2. We can repeat what we did above; we use the fixed point A_{i+1} to create a new closed geodesic u_i and cut along it to obtain M_i. The surface M_i will be connected by exactly the same argument as before. We do this until we have $2g-2$ closed geodesics and obtain the desired result.

$(2 \Rightarrow 3)$: Suppose we have a set $S_{2g-2} = \{u_1, \dots, u_{2g-2}\}$ of simple closed geodesics that all meet, and only meet, in some point $p \in M$. We cut open the surface M along u_1 to obtain the surface $M_1 \setminus u_1$. Since the remaining $u_2, \dots, u_{2g-2}$ did not meet u_1 but in p, and there intersected u_1 transversally, they are now geodesics between the two distinct copies of u_1 on M_1. Hence it is clear M_1 is connected and a $(g-1, 2)$-surface. If we now cut M_1 open along u_2 to obtain $M_2 = M_1 \setminus u_2$, we obtain a $(g-1, 1)$-surface. If $g > 2$, we can cut open M_2 along $u_3, \dots, u_{2g-2}$, and by the same arguments as above, we will end up with a $(1, 1)$-surface M_{2g-2}.

Now, we note that the boundary of M_{2g-2} consists of $4g-4$ geodesic segments, corresponding to copies of $u_1, \dots, u_{2g-2}$. Each u_i has two copies on this boundary, and they are opposite to each other on the boundary (this follows from the fact that the geodesics meet transversally in p). We may assume without loss of generality that the $u_1, \dots, u_{2g-2}$ were labeled such that the clockwise order of the geodesic segments on the boundary is $u_1, \dots, u_{2g-2}, u_1, \dots, u_{2g-2}$.

Let q and q' be the two distinct points on the boundary of

M_{2g-2} where a copy of u_1 and a copy of u_{2g-2} meet. Then we can define u_{2g-1} to be a geodesic from q to q' which is not homotopic to the boundary. We cut M_{2g-2} open along our new u_{2g-1} to obtain M_{2g-1}, which is a cylinder. The way in which we chose q and q' now ensures that each of it boundary components consists of copies of $u_1, \dots, u_{2g-1}$, and in order when going around the boundary. We repeat the same trick by letting r and r' be the two distinct points where u_1 and u_{2g-1} meet on M_{2g-1}, and let u_{2g} be a geodesic between r and r'. We cut open M_{2g-1} along u_{2g}, and see that we have now obtained a hyperbolic polygon P with $4g$ sides. Its opposite sides correspond to the same geodesic u_i.

From the construction it is clear that u_{2g-1} and u_{2g} become simple closed loops on M that only meet each other, and each of the other $u_1, \dots, u_{2g-2}$, in p. What remains is to prove that they are smooth at p. We prove this on the polygon P. For this we will need the following lemma:

Lemma 10 *Let T be a hyperbolic triangle with sides of length a, b, c, opposite to angles α, β, γ respectively. Let T' be another hyperbolic triangle with sides of length a', b', c' opposite to angles α', β', γ' which satisfies $a' = a$ and $b' = b$. Then*

$$c' > c \iff \gamma' > \gamma \iff \alpha' + \beta' < \alpha + \beta.$$

Proof. We use the fact that

$$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos \gamma. \tag{2}$$

We can then already see that

$$\begin{aligned} c' > c &\iff \cosh c' > \cosh c, \\ &\iff \cos \gamma' < \cos \gamma, \\ &\iff \gamma' > \gamma. \end{aligned}$$

Now let m be the midpoint of the side of length c opposite to γ, and let u be the geodesic segment between m and the vertex associated to γ (see figure). Write $\ell(u) = \frac{d}{2}$. If we now attach two copies of the triangle T as in figure 4, we see that we obtain a quadrangle. The two copies of u form a diagonal which splits the quadrangle into two triangles. Again applying formula (2), we see that

$$\begin{aligned} \cosh a &= \cosh \frac{c}{2} \cosh \frac{d}{2} - \sinh \frac{c}{2} \sinh \frac{d}{2} \cos \delta, \\ \cosh b &= \cosh \frac{c}{2} \cosh \frac{d}{2} + \sinh \frac{c}{2} \sinh \frac{d}{2} \cos \delta. \end{aligned}$$

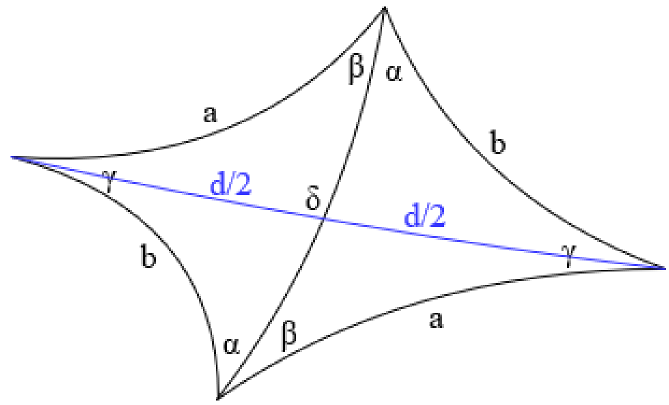

Figure 4: Two copies of T, attached at the side with length c.

So

$$\cosh a + \cosh b = 2 \cosh \frac{c}{2} \cosh \frac{d}{2}.$$

If we keep a and b fixed we see that, as c increases, d decreases. If d decreases, then so does $\alpha + \beta$ by the first part of the proof. ∎

Corollary 11 *Let Q and Q' be two hyperbolic quadrilaterals with the sides a, b, c, d and a', b', c', d' respectively, which follow in the same order (say, clockwise) around the quadrilaterals. Suppose*

$$(\ell(a), \ell(b), \ell(c), \ell(d)) = (\ell(a'), \ell(b'), \ell(c'), \ell(d')),$$

and let $\alpha, \beta, \gamma, \delta$ and $\alpha', \beta', \gamma', \delta'$ be the internal angles, where corresponding symbols (e.g. α, α') denote angles between sides of the same length. Then

$$\alpha + \gamma > \alpha' + \gamma' \iff \beta + \delta < \beta' + \delta'.$$

Proof. We draw diagonals e, e' in Q and Q'. By the first equivalence of the lemma,

$$\begin{aligned} \alpha > \alpha' &\iff \ell(e) > \ell(e'), \\ &\iff \gamma > \gamma', \end{aligned}$$

and the result follows from the second equivalence of the lemma. ∎

We go back to our polygon P. Let z_0 be a vertex where two sides corresponding to u_1 and u_{2g} meet, and label the other vertices

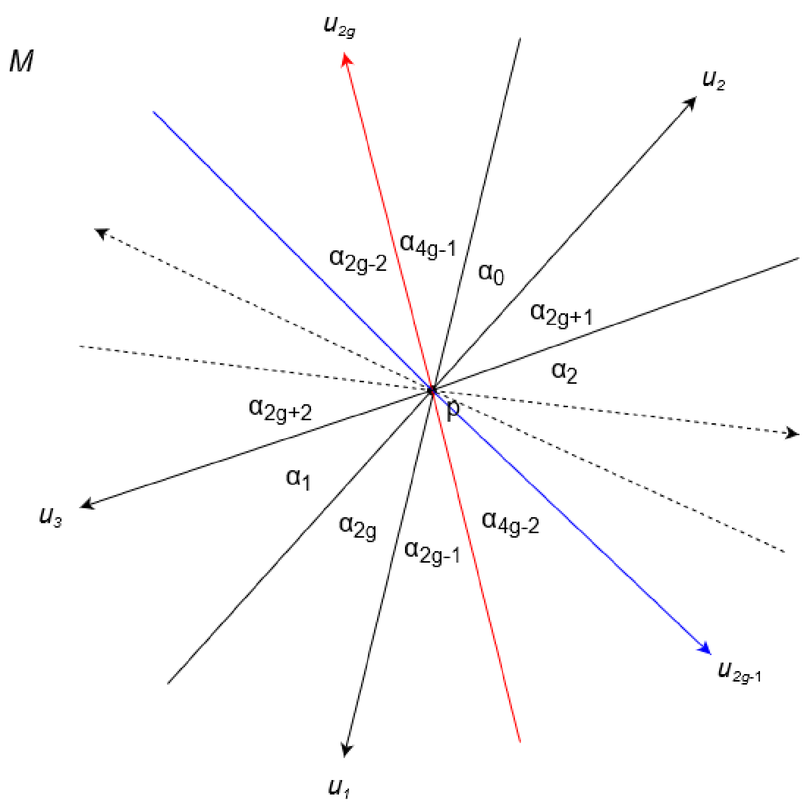

Figure 5: The geodesic loops $u_1, \ldots, u_{2g}$ meeting at p on the surface M

such that z_i and z_{i+2g} are the vertices where u_i and u_{i-1} meet, and going around the polygon gives the vertices $z_0, \ldots, z_{4g-1}$ in order. Let α_i be the internal angle of the polygon at the vertex z_i. Note that the α_i and α_{i+2g} correspond to the angles u_i and u_{i+1} make on the surface, as in figure 5, where we let $u_0 = u_{2g}$. Since $u_1, \ldots, u_{2g-2}$ at least are closed geodesics, we see that $\alpha_i = \alpha_{i+2g}$ for $i = 0, \ldots 2g-3$. Also,

$$\alpha_{2g-2} + \alpha_{4g-1} = \alpha_{2g-1} + \alpha_{4g-2}.$$

If we can prove $\alpha_i = \alpha_{i+2g}$ for $i = 2g-2, 2g-1$ as well, we can can conclude u_{2g-1} and u_{2g} are closed geodesics as well.

Let d be the geodesic between z_{2g-1} and z_{4g-1}, d_1 the geodesic between z_{2g-3} and z_{4g-3}. See figure 6.
The geodesics d, d_1 and d_2 divide the polygon in 4 components. Let V_i be the component with d_i among its sides, but not d, and let W_i be the component with both d_i and d among its sides. Now, since $\alpha_i = \alpha_{i+2g}$ for $i = 0, \ldots 2g-3$, V_1 and V_2 are isometric and $\ell(d_1) = \ell(d_2)$. Therefore, W_1 and W_2 must

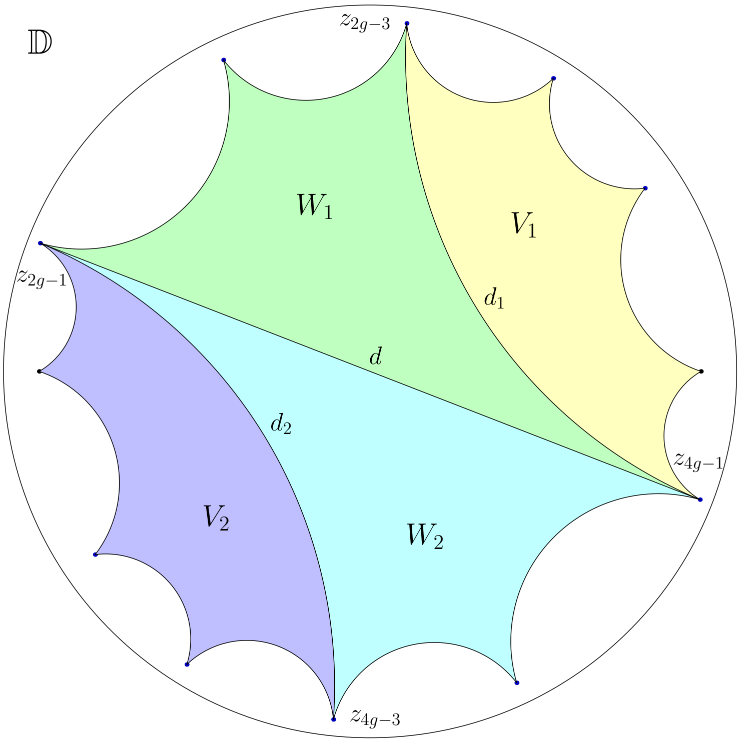

Figure 6: The polygon P and geodesics d, d_1, d_2.

be quadrilaterals with the same side lengths. Now suppose W_1 and W_2 are isometric. Then we certainly have $\alpha_{2g-2} + \alpha_{4g-1} = \alpha_{2g-1} + \alpha_{4g-2}$. If we now try to keep W_1 fixed and try to vary W_2 by varying its angles, corollary 11 implies that we cannot do so while keeping $\alpha_{2g-2} + \alpha_{4g-1} = \alpha_{2g-1} + \alpha_{4g-2}$. For instance, increasing α_{2g-2} increases α_{4g-1} as well, while decreasing α_{2g-1}. Therefore W_1 and W_2 must be isometric, and $\alpha_i = \alpha i + 2g$ for all $i = 0, \dots, 4g-1$. With this, we have then proved $u_1, \dots, u_{2g}$ are all closed geodesics.

(3 ⇒ 4) Suppose we have a set $S_{2g} = \{u_1, \dots, u_{2g}\}$ of simple closed geodesics that all meet, and only meet, in some point $p \in M$. If we look at the proof of (2 ⇒ 3) above, we see that most of the work is already done: we see that by cutting along M open along all u_i results in a hyperbolic polygon with opposite sides of equal length and equal opposite angles that sum to 2π. We translate and rotate the polygon by isometries in the hyperbolic disk such that z_0 lies on the real axis, and $z_{2g-1} = -z_0$. The upper half of the polygon defined by the vertices $z_0, z_1 \dots, z_{2g-1}$ is isometric to its lower half defined by the vertices $z_{2g-1}, z_{2g}, \dots, z_0$, and we see that $z_{i+2g} = -z_i$ for $i = 0, \dots, 2g-1$. Hence we M corresponds to an admissible

polygon.

($4 \Rightarrow 1$) Suppose we have an admissible polygon P that corresponds to the surface M. The projection to M of the map $\phi : z \mapsto -z$ on P has exactly $2g+2$ fixed points: $2g$ coming from the midpoints of the sides, one from the point corresponding to the vertices, and one from the origin. Moreover, it is not the identity, but ϕ^2 is the identity on P, and hence its projection to M is the identity on M. We conclude M is hyperelliptic, and the proof is finished. ■

Note that this theorem and its proof tell us that an admissible polygon gives us a closed hyperelliptic surface, on which the curves corresponding to its sides form what we defined as a geodesic star set (see section 2). Hence we see that an admissible polygon with labeled sides corresponds to a unique marked hyperelliptic surface, and we have the following:

Corollary 12 *The marked closed hyperelliptic surfaces of genus* g *are in one-to-one correspondence to the set of admissible* $4g$*-gons.*

Note again that in the case of genus 2, all surfaces are hyperelliptic, so that the admissible octagons correspond to the entirety of Teichmüller space, as we noted earlier. For higher genus this is however not the case, and the results we will obtain will not pertain to Teichmüller space in general.

Part II

RESEARCH AND DISCUSSION

4 CONSTRUCTING ADMISSIBLE POLYGONS

We have established that we may use admissible polygons to describe hyperelliptic Riemann surfaces. We now wish to see what the space of admissible polygons looks like. We can construct a symmetric $4g$-gon by choosing $2g$ vertices in the upper half of the hyperbolic disk, $\mathbb{D}^+ = \{z \in \mathbb{D} | \operatorname{Im} z \geqslant 0\}$, and letting $z_{i+2g} = z_i$, $i = 0, \dots, 2g-1$. But clearly, such a polygon need not be admissible; we will want to require that z_0 lies on $(0, 1)$, and that its internal angles add up to π. So we will pursue the following question: given $z_1, \dots, z_{2g-1} \in \mathbb{D}^+$, can we find a suitable z_0 so that the polygon with vertices $z_0, \dots, z_{2g-1}, -z_0, \dots, -z_{2g-1}$ is admissible?

Let us first fix some notation and definitions for ease of exposition. We denote by $P[z, \dots, z_{2g-1}]$ the symmetric hyperbolic polygon with vertices $z_0, \dots, z_{2g-1}, -z_0, \dots, -z_{2g-1}$. See figure 7. It will be convenient to work with just the upper half of the symmetric polygon as well, so for $z \in [0, 1]$ we will denote by $D[z, z_1, \dots, z_{2g-1}]$ the hyperbolic polygon with vertices $z, \dots, z_{2g-1}, -z$. This is a $(2g+1)$-gon if z is not 0 and a $2g$-gon if it is. We now define what it means for a set of $2g$ points in $\mathbb{D}^+$ to be admissible:

Definition 13 *Let $z_1, \dots, z_{2g-1} \in \mathbb{D}$ be such that*

$$0 < \arg z_1 < \cdots < \arg z_{2g-1} < \pi.$$

Then $z_1, \dots, z_{2g-1}$ are called admissible if there exists a $z_0 \in (0, 1)$ such that $P[z_0, \dots, z_{2g-1}]$ is admissible.

Our main question becomes: when are $z_1, \dots, z_{2g-1} \in \mathbb{D}^+$ admissible? We will deduce the following criterion:

Theorem 14 *Let $z_1, \dots, z_{2g-1} \in \mathbb{D}$ be such that*

$$0 < \arg z_1 < \cdots < \arg z_{2g-1} < \pi. \tag{3}$$

Then $z_1, \dots, z_{2g-1}$ are admissible if and only if

$$\arg\left((1 - \bar{z}_1)(1 + z_{2g-1})\right) + \sum_{k=1}^{2g-2} \arg\left(1 - z_k \bar{z}_{k+1}\right) > (g-1)\pi. \tag{4}$$

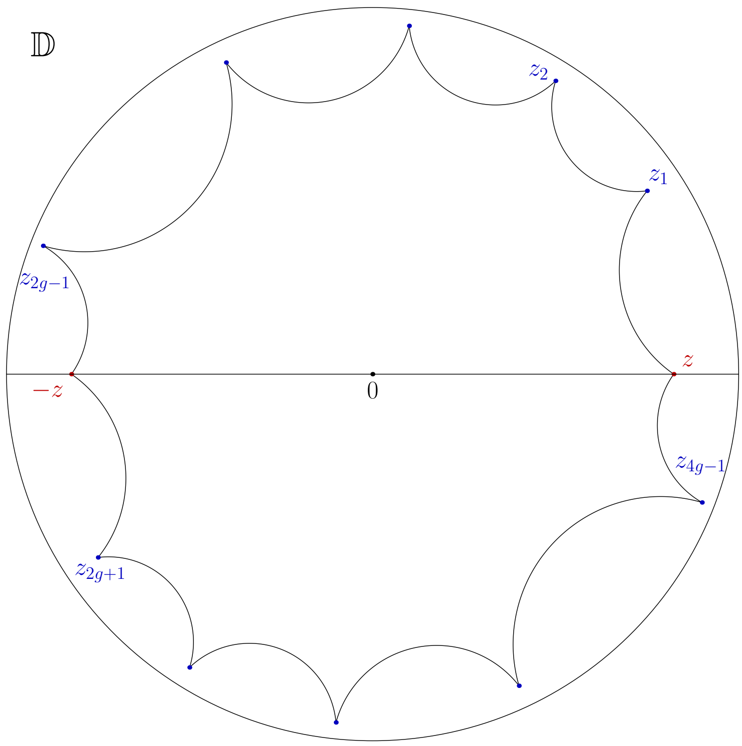

Figure 7: A symmetric hyperbolic polygon which we denote by $P[z, z_1, \ldots, z_{2g-1}$. The upper half we denote by $D[z, z_1, \ldots, z_{2g-1}]$.

If they are, then there exists a unique z_0 *such that* $P[z_0, \ldots, z_{2g-1}]$ *is admissible. This* z_0 *is given by*

$$z_0 = \frac{2c}{-b \pm \sqrt{b^2 - 4ac}}, \tag{5}$$

where

$$\begin{aligned} a &= \operatorname{Im} -u\bar{z}_1 z_{2g-1}, \\ b &= \operatorname{Im}((-\bar{z}_1 + z_{2g-1})u), \\ c &= \operatorname{Im} u, \\ u &= (-1)^g \prod_{k=1}^{2g-2} (1 - z_k \bar{z}_{k+1}). \end{aligned}$$

Note that the theorem not only gives a criterion for admissibility for vertices $z_1, \ldots, z_{2g-1}$, but also tells that there is a only one admissible polygon that incorporates these vertices. Moreover, it shows us how we may calculate the missing vertex, z_0.

Proof. Suppose we have $z_1, \ldots, z_{2g-1}$ such that (3) is satisfied. We wish to find a z_0 on $(0, 1)$ such that the internal angles of

$P[z_0, \ldots, z_{2g-1}]$ sum to 2π. Equivalently, we want that the internal angles of its upper half $D[z_0, \ldots, z_{2g-1}]$ sum to π. By the area formula for hyperbolic polygons, this is true if the area of $D[z_0, \ldots, z_{2g-1}]$ is $(2g-2)\pi$.

We note that the area of $D = D[z, \ldots, z_{2g-1}]$ becomes that of a $2g$-gon as z goes to 0, which is smaller than $(2g-2)\pi$. Furthermore, the area of D is strictly increasing as z increases along the real axis. So taking the limit of z to 1, we see that $z_1, \ldots, z_{2g-1}$ are admissible if and only if the area of $D[1, z_1, \ldots, z_{2g-1}]$ is bigger than $(2g-2)\pi$. It also shows us that a suitable z_0 for admissible $z_1, \ldots, z_{2g-1}$ will be unique.

To derive the criterion given in (4), we calculate the area of $D[1, z_1, \ldots, z_{2g-1}]$. For this we use the following lemma:

Lemma 15 *Let $z_1, z_2 \in \mathbb{D}$ s.t. $0 < \arg z_1 < \arg z_2 < \pi$. Then the area of the hyperbolic triangle with vertices 0, z_1 and z_2 is given by*

$$\operatorname{area} T = 2\arg(1 - z_1\bar{z}_2).$$

A proof is given in the appendix of [1]. Using the lemma, we can find

$$\operatorname{area} D = 2\arg(1-\bar{z}_1) + 2\sum_{i=1}^{2g-2} \arg(1 - z_k\bar{z}_{k+1}) + 2\arg(1 + z_{2g-1}).$$

So we find the criterion for admissibility is

$$\arg(1-\bar{z}_1) + \sum_{i=1}^{2g-2} \arg(1 - z_k\bar{z}_{k+1}) + \arg(1 + z_{2g-1}) > (g-1)\pi.$$

As the first and last term are both smaller than $\frac{\pi}{2}$ (as they represent half the area of an hyperbolic triangle), we may rewrite this as

$$\arg\left((1-\bar{z}_1)(1+z_{2g-1})\right) + \sum_{k=1}^{2g-2} \arg\left(1 - z_k\bar{z}_{k+1}\right) > (g-1)\pi.$$

All that is left is to calculate z_0, given that $z_1, \ldots, z_{2g-1}$ are admissible. We need that

$$\begin{aligned}(g-1)\pi &= \sum_{k=0}^{2g-1} \arg\left(1 - z_k\bar{z}_k\right), \\ &= \arg((1-z_0\bar{z}_1)(1+z_{2g-1}z_0)) + \sum_{k=1}^{2g-2} \arg\left(1 - z_k\bar{z}_{k+1}\right).\end{aligned}$$

Since we assume $z_1, \ldots, z_{2g-1}$ are admissible and the first term is between 0 and π, the second term must be between $(g-2)\pi$ and $(g-1)\pi$. We note that if g is even, this implies that the second term can be written as

$$\begin{aligned}\sum_{k=1}^{2g-2} \arg\left(1 - z_k\bar{z}_{k+1}\right) &= (g-2)\pi + \arg\left(\prod_{k=1}^{2g-2}(1 - z_k\bar{z}_{k+1})\right), \\ &= (g-2)\pi + \arg(u),\end{aligned}$$

while if g is odd,

$$\begin{aligned}\sum_{k=1}^{2g-2} \arg(1 - z_k\bar{z}_{k+1}) &= (g-2)\pi + \arg\left(-\prod_{k=1}^{2g-2}(1 - z_k\bar{z}_{k+1})\right), \\ &= (g-2)\pi + \arg(u).\end{aligned}$$

With this, our requirement for z_0 becomes

$$\pi = \arg\left((1 - z_0\bar{z}_1)(1 + z_{2g-1}z_0)u\right)$$

where we know the right hand side is not 0. Hence we may also write this as

$$\begin{aligned}0 &= \operatorname{Im}\left((1 - z_0\bar{z}_1)(1 + z_{2g-1}z_0)(-1)^g \prod_{k=1}^{2g-2}(1 - z_k\bar{z}_{k+1})\right), \\ &= az_0^2 + bz_0 + c.\end{aligned}$$

We deduce that $z_0 = \pm\frac{2c}{-b\pm\sqrt{b^2-4ac}}$. To see that it is the solution with the plus sign that we are looking for, note $0 < \arg(u) < \pi$, and hence $c = \operatorname{Im} u > 0$. We now look at the cases $a > 0$ and $a < 0$.

If $a < 0$, $b^2 - 4ac > b^2$, so that the solution with the minus sign lies in $(-\infty, 0)$.

For the case $a > 0$, first note that (4) implies

$$\pi < \arg\left((1 - \bar{z}_1)(1 + z_{2g-1})\right) + \arg(u).$$

We see that $a + b + c < 0$. Hence $a > 0$ and $c > 0$ implies that $b < 0$. Now, $a + b + c < 0$ also implies $-a > b + c$. Therefore, $b^2 - 4ac > b^2 + 4bc + 4c^2 = (-b - 2c)^2$. We conclude that $\sqrt{b^2 - 4ac} > -b - 2c$, and $-b - \sqrt{b^2 - 4ac} < 2c$. So we find that $\frac{2c}{-b-\sqrt{b^2-4ac}} > 1$, and the solution with the minus sign again does not lie in (0,1). ∎

Let us compare this to the case of genus 2, as described in Aigon-Dupuy et al. We note that in the case of genus 2, the $(-1)^g$-term plays no role, and the coefficients a, b, c are those given in Aigon-Dupuy et al. The criterium (4) for admissibility becomes a sum of only 4 terms, which are each smaller than $\frac{\pi}{2}$. It can therefore be restated more neatly as

$$a + b + c = \operatorname{Im}((1 - \bar{z}_1)(1 + z_3) \prod_{k=1}^{2} (1 - z_k \bar{z}_{k+1})) < 0,$$

as in Aigon-Dupuy. This condition, $a + b + c < 0$, is necessary but not sufficient for admissibility in higher genus, because it includes the $z_1, \ldots, z_{2g-1}$ that can form a symmetric polygon of area a multiple of 4π less than that which is required.

5 ALGORITHMS FOR CLOSED GEODESICS

We now turn to apply our description of hyperelliptic surfaces by admissible polygons to a specific computational problem. Given any non-contractible closed curve on a Riemann surface, there exists a unique closed geodesic that is homotopic to it. In this section we will see how we may compute this homotopic closed geodesic on a hyperelliptic surface, by computing it on the admissible polygon corresponding to the surface.

5.1 FINDING THE HOMOTOPIC CLOSED GEODESIC

To tackle the problem we will first introduce some notation. Let P be an admissible $4g$-gon corresponding to a hyperelliptic surface M of genus $g \geqslant 2$. We will label the sides counter-clockwise by $a_0, \ldots, a_{4g-1}$, as in figure 8. We denote the hyper-

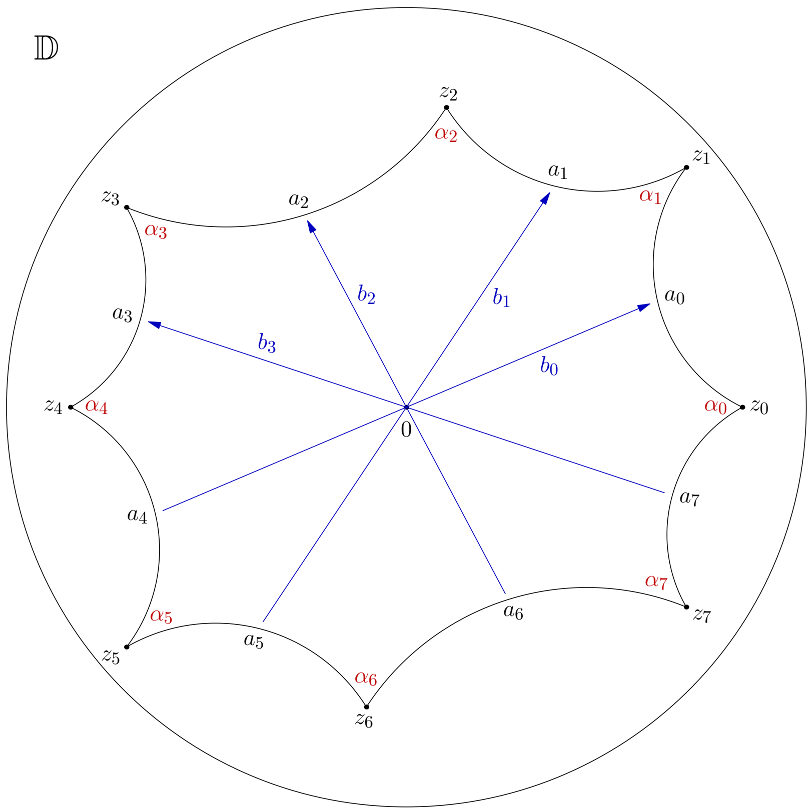

Figure 8: An admissible octagon.

bolic transformation that maps the side a_k to a_{k+2g} by b_k, for $k = 0, \ldots, 4g-1$. Note that $b_k^{-1} = b_{k+2g}$, $k = 0, \ldots 2g-1$. We will abuse notation and take our indices of a and b modulo $4g$, so $a_{i+2g} := a_{(i+2g) \bmod 4g}$. We denote the Fuchsian group of the surface by Γ. Note that

$$\Gamma = \langle b_0, \ldots, b_{2g-1} | b_0 b_1^{-1} b_2 b_3^{-1} \ldots b_{2g-2} b_{2g-1}^{-1} = \mathrm{id} \rangle.$$

Since we the denote the inverses b_i^{-1} by b_{i+2g}, we may also write the relation defining Γ as

$$b_0 b_{1+2g} b_2 b_{3+2g} \cdots b_{2g-2} b_{4g-1} = \mathrm{id}\,.$$

5.1.1 *Representing curves on an admissible polygon*

We wish to study curves on a hyperelliptic surface M through its representation by an admissible polygon P. Throughout this section we will assume all curves have a constant speed parametrization with domain $[0, 1]$. We must note though that the projection of P to M will not give us a chart on M; the interior of P is in one-to-one correspondence with an open subset of M, but points on M corresponding to the sides of P each correspond to two points in P, and all $4g$ vertices of P even project to a single point on M. This implies that we may represent any (continuous) curve $\hat{\alpha}$ in M on P by a sequence $\alpha_0, \ldots, \alpha_{m-1}$ of continuous segments on P, $m \in \mathbb{N}$, $M > 1$ such that:

1. The begin- and endpoints of the segments $\alpha_1, \ldots, \alpha_{m-1}$ lie on the boundary of P, as do the endpoint $\alpha_0(1)$ of the first segment and the begin point $\alpha_{m-1}(0)$ of the last.

2. The begin- and endpoints of subsequent segments are such that $\alpha_i(1) = \alpha_{i+1}(0)$, $i = 0, \ldots, m-2$, so that they project to the same point on surface M.

3. The begin- and endpoint of subsequent segments are not the same, so $\alpha_i(1) \neq \alpha_{i+1}(0)$, $i = 0, \ldots, m-2$ [1].

See Figure 9 for an example. The segments $\alpha_0, \ldots, \alpha_{m-1}$ together define a piecewise continuous curve that represents the curve α in M on P. Throughout this section, we will denote this representation on P by α as well. To distinguish the curve on M,

1 This we require so that α truly "crosses" a side of P at the endpoint of each of its segments, and this will be beneficial to the clarity of the algorithms presented later on.

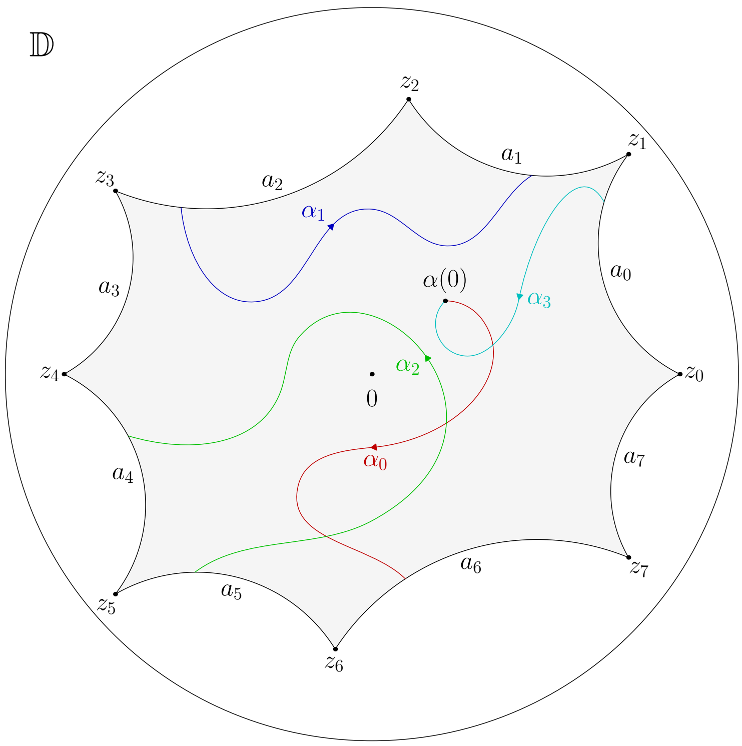

Figure 9: A sequence of 4 segments $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ in an admissible polygon P that project to a closed loop α on the surface M. Note also that two points on opposite sides that project to the same point on M may not appear to be placed equally far along their respective sides. This is due to the hyperbolic metric differing from the Euclidean metric.

we add a hat; that is, α will now denote a piecewise continuous curve on P that projects to a continuous curve on M, which we denote by $\hat{\alpha}$.

5.1.2 *The lift of a closed curve*

Now, we are given a curve α in P that represents a closed curve $\hat{\alpha}$ on the surface M corresponding to P, and let α consist of the segments $\alpha_0, \dots, \alpha_{m-1}$ (in the manner of the previous paragraph). To facilitate our study of $\hat{\alpha}$, we will study its lift to $\mathbb{D}$, which is the universal cover of M. Let $\pi : \mathbb{D} \to \mathbb{D}/\Gamma = M$ be the projection from $\mathbb{D}$ to M. The fact that $\mathbb{D}$ is a universal cover implies that, given a point $z \in \pi^{-1}(\hat{\alpha}(0)) \subset \mathbb{D}$, there exists a unique curve $\tilde{\alpha}$ in $\mathbb{D}$ starting at z that projects to $\hat{\alpha}$. In particular, we may pick $\tilde{\alpha}$ to be the curve that starts in P at $\alpha(0)$ (where α is again the piecewise continuous curve on P). We will abuse terminology and refer to this curve $\tilde{\alpha}$ as the "lift of α (to $\mathbb{D}$)". In the same vein, we will call α a "representation of $\tilde{\alpha}$ in P", so we need not write "a representation in P of the projection $\hat{\alpha}$ on M of $\tilde{\alpha}$ that starts at $\tilde{\alpha}(0)$". See figure 10 for an example of a piecewise continuous curve α representing a curve on the surface.

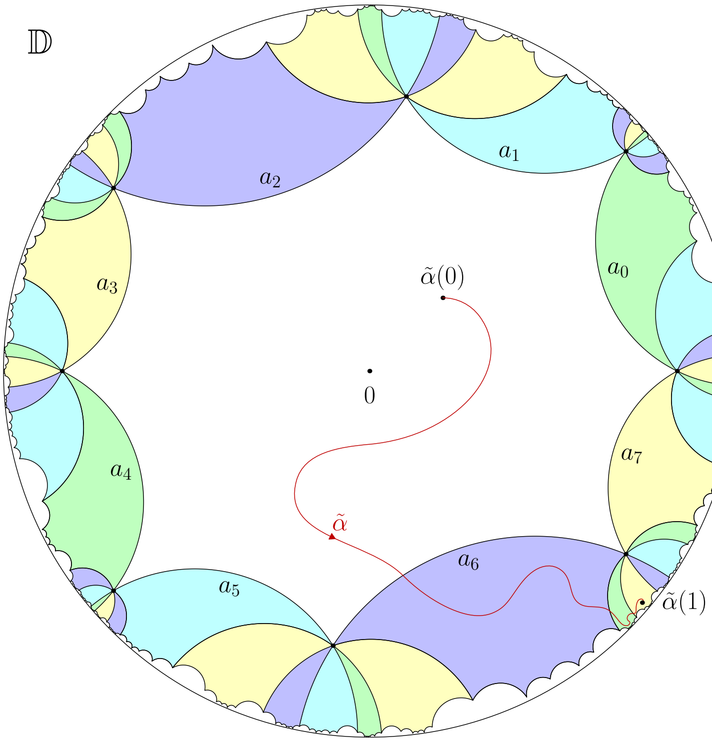

Figure 10: The lift $\tilde{\alpha}$ of the curve α in figure 9. Note that $\tilde{\alpha}$ is a continuous curve traversing multiple copies of P in $\mathbb{D}$.

Since $\hat{\alpha}$ is a closed loop on M, its endpoint $\hat{\alpha}(1)$ must be lifted to $g\hat{\alpha}(1)$ for some unique $g \in \Gamma$. Now we have introduced the necessary notation and terminology, we specify the goal of our algorithm. Its goal is the following: we take as input a sequence of segments $\alpha_0, \ldots, \alpha_{m-1}$ in P, that represent a closed curve $\hat{\alpha}$ on the surface M. As output we wish to produce a sequence of segments $\beta_0, \ldots, \beta_{m-1}$ in P that represent the closed geodesic $\hat{\beta}$ homotopic to $\hat{\alpha}$ on M.

The general idea is this: any homotopy of $\hat{\alpha}$ lifts to a unique homotopy of $\tilde{\alpha}$. This implies that this g will be invariant under homotopy. But we know that g has an axis, A_g, along which the displacement length is minimal, i.e. $d(z, gz)$ is minimal for z on A_g. Then for each z on the axis, the segment from z to gz of the axis projects to the closed geodesic on M that is homotopic to α.

We now have a general outline for our algorithm:

1. We find g. We do this by determining what sides of the copies of P are crossed by $\tilde{\alpha}$.
2. We choose a point z on the axis A_g, and calculate gz on the axis.
3. We calculate the sequence $P_0, \ldots, P_{r-1}$ of copies of P along the segment $\tilde{\beta}$ of A_g between z and gz, and find the representation in P of the sequence of segments of $\tilde{\beta}$ defined by the copies $P_0, \ldots, P_{r-1}$.

During the algorithm we will assume that both $\tilde{\alpha}$ and the calculated $\tilde{\beta}$ do not pass through vertices of copies of P (though their endpoints may be vertices). This ensures that the segments of $\tilde{\alpha}$ and $\tilde{\beta}$ lie in neighboring copies of P (copies of P sharing a side) in $\mathbb{D}$. Though it is not difficult to deal with the case that $\tilde{\alpha}$ or $\tilde{\beta}$ does pass through vertices, it complicates the exposition of the algorithm. Also, given such a curve, it is easy to see how we may apply a homotopy that perturbs the curve slightly as to avoid the vertices crossed.

We will now present the algorithm in more detail. We make use of three subroutines, which will be given after the algorithm to reduce clutter.

Algorithm 1

Input: A piecewise-continuous curve α given by a sequence of segments $\alpha_0, \ldots, \alpha_{m-1}$ in an admissible polygon P, that represents a closed curve $\hat{\alpha}$ on the hyperelliptic surface M corresponding to P.

Output: A sequence of segments $\beta_0, \ldots, \beta_{r-1}$ in P that represents the closed geodesic $\hat{\beta}$ homotopic to $\hat{\alpha}$.

1. *Finding* g*:*
   a) We first determine which sides of copies of P the lift $\tilde{\alpha}$ of α will cross. These we can determine by the begin and endpoints of the segments $\alpha_0, \ldots, \alpha_{m-1}$, using subroutine 1 given below. The result is a sequence of indices $i_0, i_1, \ldots, i_{m-2}$, that are such that $\tilde{\alpha}$ crosses the sides $a_{i_0}, \ldots, a_{i_{m-2}}$ of copies of P before arriving at $\tilde{\alpha}(1)$.

b) We now calculate the sequence $(P_0, P_1, \dots, P_{m-1})$ of copies of P that $\tilde{\alpha}$ crosses in $\mathbb{D}$, where we denote $P_0 := P$. Moreover, we calculate the transformations h_k such that $P_k = h_k P$, $k = 0, \dots, m-1$. For this we use the indices from the previous step, and subroutine 2 (given below), which, given a copy P' of P and the index k of a side, calculates the copy of P sharing the side $h_k a_k$ with P'. This allows us to calculate the sequence $(P_0, P_1, \dots, P_{m-1})$ and the sequence of transformations h_k inductively.

c) The transformations h_{m-1} sends P_0 to P_{m-1}, and thus $\tilde{\alpha}(0)$ to $\tilde{\alpha}(1)$. So we find that $g = h_{m-1}$.

2. *Finding the axis of* g *and picking a segment of it:*

a) We solve $gz = z$ to find the fixed points of g on the boundary of $\mathbb{D}$.

b) We calculate the axis A_g of g by finding the geodesic between g's fixed points, i.e. the circular arc that meets the boundary perpendicularly at g's fixed points.

c) We pick a point $\tilde{p}_0$ on A_g, and calculate $g\tilde{p}_0$. We define $\tilde{\beta}$ to be the segment of A_g from $\tilde{p}_0$ to $g\tilde{p}_0$.

3. *Finding the representation of* $\tilde{\beta}$ *in* P*:*

a) Since the transformations $b_0, \dots, b_{4g-1}$ are used repeatedly, we first calculate them and put them in a table for later use. Each transformation b_k is given by a π-rotation around the origin, followed by a π-rotation around the midpoint of the side a_k [2].

b) We find a copy of P that contains $\tilde{p}_0$. If $\tilde{p}_0$ is not already in P, we calculate the geodesic segment $\tilde{\gamma}$ from 0 to $\tilde{p}_0$. We then apply the subroutine 3 given below to obtain the sequence of copies of P that $\tilde{\gamma}$ crosses. We only need the last one, which we will call P'_0. The subroutine also determines the transformation h'_0 such that $h'_0 P = P'_0$.

c) We now apply subroutine 3 to find the sequence $P'_0, P'_1, \dots, P'_{r-1}$ of copies of P that $\tilde{\beta}$ crosses, and define $\tilde{\beta}_i$ to be the segment of $\tilde{\beta}$ that lies in P'_i, $i = 0, \dots, r-1$. The subroutine also provides the transformations $h'_0, \dots, h'_{r-1}$ such that $h'_k P = P_k$.

2 See [1] section IV for an explicit expression for b_k in terms of the vertices of P.

d) We calculate the representation β_i in P of each segment $\tilde{\beta}_i$ by applying the inverses of the transformations we found to map P to P'_i, i.e. $\beta_k = h_k'^{-1}\tilde{\beta}_k$, $k = 0, \dots, r-1$. This gives us the segments $\beta_0, \dots, \beta_{r-1}$ that make up β, and we are done. ■

These are the subroutines referred to in the algorithm:

Subroutine 1

Goal: to find the indices of sides crossed by a sequence of segments.

Input: a sequence of segments $\alpha_0, \dots, \alpha_{m-1}$ in an admissible polygon P, that represents a closed curve $\hat{\alpha}$ on the hyperelliptic surface M corresponding to P.

Output: a sequence $i_0, \dots, i_{m-2}$ of indices in $\{0, \dots, 4g-1\}$ with the property that α "crosses" the sides $a_{i_0}, \dots, a_{i_{m-2}}$ of P in order.

We repeat the following step for $k = 0, \dots, m-2$:

1. We check which side the endpoint of α_k lies on. If it lies on the side a_j, we have determined $i_k = j$ [3]. ■

Subroutine 2

Goal: to calculate a neighboring copy of P, from the shared side.

Input:

- the transformation $h_0 \in \Gamma$ such that $P'_0 := h_0 P$ is the copy of P we wish to find a neighbor of.
- the index $k \in \{0, \dots, 4g-1\}$ such that $h_0 a_k$ is the side of P'_0 shared by the neighboring copy we wish to construct.
- the transformations $b_0, \dots, b_{4g-1}$ (the side-pairing transformations of P).

Output:

- the transformation $h_1 \in \Gamma$ such that $P'_1 := h_1 P$ is the copy of P that shares the side $h_0 a_k$ with P'_0.

3 Note that the endpoint of α_k, $k = 0, 1, \dots, m-2$ is by assumption not a vertex, and hence lies on one side only.

- P_1' itself.

Steps:

1. We calculate the transformation $h_1 = h_0 b_k h_0^{-1}$.
2. We determine $P_1' = h_1 P$. ■

For subroutine 3 (given below), we assume we may calculate the intersection of two curves in $\mathbb{D}$.

Subroutine 3

Goal: to calculate the sequence of copies of P that a curve in $\mathbb{D}$ passes through.

Input:

- a curve $\tilde{\gamma}$ in $\mathbb{D}$.
- a transformation h_0 such that the $P_0 = h_0 P$, where P_0 is the copy of P such that $\tilde{\gamma}(0) \in P_0$.

Output:

- the sequence $P_0, \ldots, P_{n-1}$ (with $P_{k+1} \neq P_k$, $k = 0, \ldots, n-2$) of copies of P which $\tilde{\gamma}$ crosses.
- a sequence of transformations $h_0, \ldots, h_{n-1} \in \Gamma$ s.t. $h_k P = P_k$, $k = 0, \ldots, n-1$.
- a sequence of points $\tilde{p}_0, \ldots, \tilde{p}_{n-2}$ in $\mathbb{D}$ such that $\tilde{p}_i$ is the point in $\mathbb{D}$ where $\tilde{\gamma}$ leaves P_i and enters P_{i+1}, for $i = 0, \ldots, n-2$.
- a sequence of segments $\tilde{\gamma}_0, \ldots, \tilde{\gamma}_{n-1}$ such that each $\tilde{\gamma}_k$ is the segment in P_k from the endpoint of $\tilde{\gamma}_{k-1}$ to where $\tilde{\gamma}_k$ leaves P_k, $k = 1, \ldots, n-2$, and $\tilde{\gamma}_0$ is the segment from $\tilde{\gamma}(0)$ to $\tilde{\gamma}_1(0)$, while $\tilde{\gamma}_{n-1}$ is the segment from $\tilde{\gamma}_{n-2}(1)$ to $\tilde{\gamma}(1)$.

We repeat the following steps for $k = 0, \ldots, n-2$:

1. We find the first point $\tilde{p}_k$ of intersection of $\tilde{\gamma}$ with the boundary of P_k where $\tilde{\gamma}$ will leave P_k. We define $\tilde{\gamma}_k$ to be the segment of $\tilde{\gamma}$ from $\tilde{\gamma}_{k-1}(1)$ (the endpoint of the previous segment) to $\tilde{p}_k$. For the case $k = 0$, we define $\tilde{\gamma}_0$ to be the segment of $\tilde{\gamma}$ from $\tilde{\gamma}(0)$ to $\tilde{p_k}$.

2. We determine the side a_k of P that $h_k^{-1}\tilde{p}_k$ lies on.
3. We let P_{k+1} be the copy of P that shares the side a_k with P_k, and determine it by using subroutine 2. This also gives us the transformation h_{k+1} s.t. $h_{k+1}P = P_{k+1}$. ∎

5.2 FINDING A HOMOTOPY

It is sometimes said that the road is more important than the destination. We turn to the question of determining a homotopy between the curve and its homotopic closed geodesic explicitly.

To make the goal of this section more precise, we will first introduce some terminology, analogous and complementary to that which we introduced in Section 5.1. We define a homotopy of a (piecewise-continuous) curve α in P representing a closed curve $\hat{\alpha}$ on M to be a map $H : [0,1] \times [0,1] \to P$ that projects to a homotopy of $\hat{\alpha}$ on M. The goal of this section is to find an algorithm that calculates a homotopy from the representative α of a closed curve $\hat{\alpha}$ to the representative β of the closed geodesic $\hat{\beta}$ homotopic to $\hat{\alpha}$.

We will again make use of the lift $\tilde{\alpha}$ of α to $\mathbb{D}$ [4]. We define a homotopy of $\tilde{\alpha}$ to be a free homotopy of $\tilde{\alpha}$ that projects to a homotopy (of closed curves) of $\hat{\alpha}$ in M, where $\hat{\alpha}$ is the projection of α to M. A homotopy $\hat{\phi}$ of $\hat{\alpha}$ on M will lift to a unique homotopy $\tilde{\phi}$ of $\tilde{\alpha}$ in $\mathbb{D}$. Let ϕ be a homotopy of α that projects to $\hat{\phi}$ on M. Then we say ϕ is the *representative on P of the homotopy* $\tilde{\phi}$, while we call $\tilde{\phi}$ the *lift of ϕ (to $\mathbb{D}$)*".

In what follows, we will also occasionally use the reverse of a curve in $\mathbb{D}$, defined as follows: let $\tilde{\gamma}$ be a curve in $\mathbb{D}$ with parametrization $\tilde{\gamma}(t) : [0,1] \to \mathbb{D}$. Then its reverse is the curve $\tilde{\gamma}^R$ parametrized as $\tilde{\gamma}^R(t) = \tilde{\gamma}(1-t)$. Similarly, we define the reverse of an homotopy $H(s,t) : [0,1] \times [0,1] \to \mathbb{D}$ by $H^R(s,t) := H(1-s,t)$.

If we have two curves $\tilde{\gamma}_0, \tilde{\gamma}_1$ such that the endpoint of $\tilde{\gamma}_0$ co-

4 We again abuse terminology here, see Subsection 5.1.2 of the previous section.

incides with the begin point of $\tilde{\gamma}_1$, $\tilde{\gamma}_1\tilde{\gamma}_0$ denotes the curve that appends $\tilde{\gamma}_1$ to $\tilde{\gamma}_0$. That is,

$$\gamma_1\gamma_0 = \begin{cases} \gamma_0(2t) & \text{if } 0 \leqslant t \leqslant \frac{1}{2} \\ \gamma_1(2t-1) & \text{if } \frac{1}{2} < t \leqslant 1 \end{cases}.$$

Also, if $\gamma_0, \gamma_1, \gamma_2$ are three curves, H_0 is a homotopy from γ_0 to γ_1 and H_1 a homotopy from γ_1 to γ_2, then H_1H_0 is defined to be the homotopy from γ_0 to γ_2 defined by

$$(H_1H_0)(s,t) = \begin{cases} H_0(2s,t) & \text{if } 0 \leqslant s \leqslant \frac{1}{2} \\ H_1(2s-1,t) & \text{if } \frac{1}{2} < s \leqslant 1 \end{cases}.$$

Our goal is to design an algorithm that, given a representative α in P of a closed curve $\hat{\alpha}$ in M and a representative β of the closed geodesic homotopic to $\hat{\alpha}$, produces a homotopy from α to β. We will do this by building a homotopy from $\tilde{\alpha}$ to $\tilde{\beta}$ in steps, in such a way that each step we add to the homotopy has a known representative in P. The following is a rough overview of the algorithm, that will be made clearer in the rest of this section.

Overview of algorithm 2

1. We first find a homotopy on $\tilde{\alpha}$ in such a way that, after the homotopy, its endpoints coincide with a lift $\tilde{\beta}$ of the closed geodesic homotopic to $\hat{\beta}$. This homotopy on $\tilde{\alpha}$ will be called $\tilde{H}_{\text{end}}$. We apply it to $\tilde{\alpha}$ [5].

2. The curves $\tilde{\alpha}$ and $\tilde{\beta}$ are homotoped to a simplified form, which we will call *standard form*. It will reduce $\tilde{\alpha}$ and $\tilde{\beta}$ to a sequence of geodesic segments between centers of copies of P. This will facilitate the exposition of the rest of the algorithm. The homotopy on $\tilde{\alpha}$ that brings $\tilde{\alpha}$ to standard form will be called $\tilde{H}_{\text{std},\tilde{\alpha}}$, while the homotopy on $\tilde{\beta}$ is called $\tilde{H}_{\text{std},\tilde{\beta}}$.

3. Next, we calculate a homotopy from $\tilde{\alpha}$ to $\tilde{\beta}$ (which are now both in standard form). This will be done in two parts: in part a), a homotopy is calculated which homotopes the first part of $\tilde{\alpha}$ to $\tilde{\beta}$, in the sense that, after applying the homotopy, $\tilde{\alpha} = \tilde{\delta}\tilde{\beta}$, where $\tilde{\delta}$ is a closed loop (in

5 By apply, we mean that we now let $\tilde{\alpha}(t) = \tilde{H}_{\text{end}}(1,t)$ for $t \in [0,1]$. That is, $\tilde{\alpha}$ is now defined to be the curve to which $\tilde{H}_{\text{end}}$ homotopes.

standard form) at the endpoint of $\tilde{\beta}$. We call this homotopy $\tilde{H}_{part}$, and apply it to $\tilde{\alpha}$.
In part b), we calculate a homotopy that homotopes $\tilde{\alpha} = \tilde{\delta}\tilde{\beta}$ to $\tilde{\beta}$, by contracting the loop $\tilde{\delta}$ to its base point. This homotopy is denoted $\tilde{H}_{contr}$, and applied to $\tilde{\alpha}$.

4. We homotope $\tilde{\alpha}$ back to the original $\tilde{\beta}$ (since $\tilde{\beta}$ was homotoped to standard form). This boils down to applying the homotopy $\tilde{H}^R_{std,\tilde{\beta}}$ to $\tilde{\alpha}$.

5. The output of the algorithm will be the homotopy $H = H^R_{std,\beta} H_{contr} H_{part} H_{std,\alpha} H_{end}$, where $H_{std,\beta}$, H_{contr}, H_{part}, $H_{std,\alpha}$ and H_{end} are the homotopies on P representing $\tilde{H}_{std,\tilde{\beta}}$, $\tilde{H}_{contr}$, $\tilde{H}_{part}$, $\tilde{H}_{std,\tilde{\alpha}}$ and $\tilde{H}_{end}$, respectively. H then homotopes α to β, as desired.

We will now present the steps of the algorithm in detail.

5.2.1 *Step 1: Closing the loop*

The first step we take is to find a homotopy of α that brings the endpoints of $\tilde{\alpha}$ to the axis A_g of $g \in \Gamma$ corresponding to $\tilde{\alpha}$ (in the same manner as in section 5.1). To this end, we apply the following algorithm:

Algorithm 2.1

Input:

- The lift $\tilde{\alpha}$ of a closed curve.
- The transformation $g \in \Gamma$ such that $\tilde{\alpha}(1) = g\tilde{\alpha}(0)$.

Output:

- A homotopy H_{end} that moves the endpoints of $\tilde{\alpha}$ to the axis A_g of g.

Steps:

1. Let A_g be the axis of g. We calculate the geodesic through $\tilde{\alpha}(0)$ that meets A_g perpendicularly. Let $\tilde{\zeta}$ be its segment from $\tilde{\alpha}(0)$ to the point where it meets the axis. As we did in section 5.1, we calculate the copies of P that $\tilde{\zeta}$ meets along the way, and calculate $\tilde{\zeta}$'s representative ζ in P.

2. We now homotope $\tilde{\alpha}$ such as to add $\tilde{\zeta}^R$ before the original starting point of $\tilde{\alpha}$ as well as $g\tilde{\zeta}$ after the endpoint of $\tilde{\alpha}$. We let

$$\tilde{H}_{\text{end}}(s,t) = \begin{cases} \tilde{\zeta}(s-4t) & \text{if } 0 \leqslant t < s/4 \\ \tilde{\alpha}\left(\dfrac{t-s/4}{1-s/2}\right) & \text{if } s/4 \leqslant t \leqslant 1-s/4 \\ g\tilde{\zeta}(s+4t-4) & \text{if } 1-s/4 < t \leqslant 1 \end{cases}.$$

Note that $\tilde{H}(s,0) = g\tilde{H}(s,1)$, as it should. Also this homotopy has a clear representative in P, given by

$$H_{\text{end}}(s,t) = \begin{cases} \zeta(s-4t) & \text{if } 0 \leqslant t < s/4 \\ \alpha\left(\dfrac{t-s/4}{1-s/2}\right) & \text{if } s/4 \leqslant t \leqslant 1-s/4 \\ \zeta(s+4t-4) & \text{if } 1-s/4 < t \leqslant 1 \end{cases}.$$

■

We apply $\tilde{H}_{\text{end}}$ to $\tilde{\alpha}$, and will continue with the resulting $\tilde{\alpha}$ from now on. After applying the homotopy $\tilde{H}_{\text{end}}$ given by algorithm 2, $\tilde{\alpha}(0)$ clearly lies on the axis. But $\tilde{\alpha}(1) = g\tilde{\alpha}(0)$, so $\tilde{\alpha}(1)$ now does as well. We define the segment of the axis from $\tilde{\alpha}(0)$ to $\tilde{\alpha}(1)$ to be $\tilde{\beta}$, and $\tilde{\delta}$ to be the closed loop $\tilde{\beta}^R\tilde{\alpha}$.

5.2.2 *Step 2: A homotopy to standard form*

To simplify our algorithm in the steps that follow, we will homotope $\tilde{\alpha}$ and $\tilde{\beta}$ such that they take a grid-like form, which we will call *standard form*. Specifically, we homotope $\tilde{\alpha}$ and $\tilde{\beta}$ such that they consist of geodesic segments between the centers of neighboring copies of P.

Definition 16 *Let $\tilde{\gamma}$ be a curve in $\mathbb{D}$ that passes through the sequence $P_0, \ldots, P_{n-1}$ of copies of P, where P_k shares a side with P_{k+1} for each $k = 0, \ldots, n-2$. Then $\tilde{\gamma}$ is said to be in standard form if it is the sequence of geodesic segments $\tilde{\gamma}_0, \ldots, \tilde{\gamma}_{n-2}$, where each $\tilde{\gamma}_k$ is the geodesic segment from the center of P_k to the center of P_{k+1}.*

See figure 11 for an example of a curve $\tilde{\alpha}$ in standard form. Let $\tilde{\gamma}$ be as in definition 15. Because of the symmetry of admissible polygons, any segment $\tilde{\gamma}_k$ crosses the polygon P_k on the midpoint of one of its sides. This is shown in the following proposition:

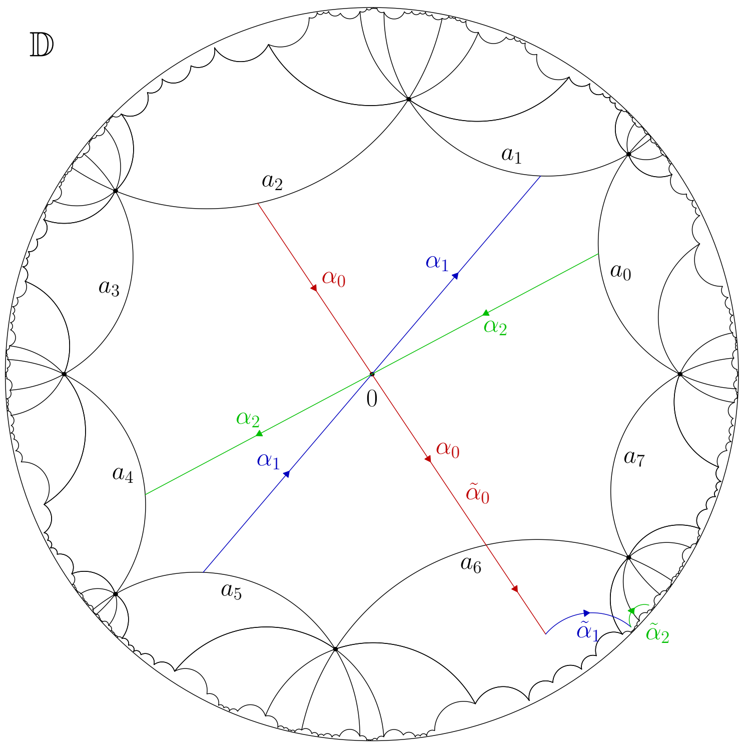

Figure 11: Here we see the curve $\tilde{\alpha}$ from figure 10 brought to "standard form", as well as its representation α in P. Note that we now divide $\tilde{\alpha}$ (and α) in 3 parts, rather than 4. Each $\tilde{\alpha}_k$, $k = 0, 1, 2$ is a geodesic segment between the centers of two adjacent copies of P.

Proposition 17 *Let P be an admissible polygon with sides $a_0, \ldots, a_{4g-1}$ and P' the copy of P that shares the side a_k with P. Then the geodesic between the center o of P and the center c' of P' crosses the side a_k of P at its midpoint m_k.*

Proof. Let b_k be the hyperbolic transformation that sends P to P'. This transformation sends the side a_{k+2g} of P to the side a_k of P. Since P is symmetric, $b_k(-z) = -b_k^{-1}(z)$ (since b_k is fully determined by the two sides a_k and a_{k+2g}, which are symmetric under $z \mapsto -z$. Therefore the endpoints ξ_+, ξ_- of the axis of b_k satisfy

$$\xi_+ = \lim_{n\to\infty} b_k^n(0) = \lim_{n\to\infty} -b_k^{-n}(0) = -\lim_{n\to\infty} b_k^{-n}(0) = -\xi_-.$$

Therefore the axis A of b_k is a (Euclidean) straight line through the origin. There are two possibilities: either it crosses both a_k and a_{k+2g}, or neither. Neither is impossible, since then a_{k+2g} lies on a different side of the axis A as a_k, and $b_k a_{k+2g} = a_k$ must lie on the other side of the axis as well. If A crosses a_k and a_{k+2g}, it is easy to see that the midpoints m_{k+2g}, m_k of a_{k+2g}

and a_k respectively are the only points paired by b_k that are also symmetric under $z \mapsto -z$. Hence, the axis A is a straight line through the origin and the midpoints m_{k+2g}, m_k. In particular, b_k maps the origin to another point on the axis. Hence the center c' of P' lies on the axis, and the geodesic segment between 0 and c' contains m_k. ■

We will not give an explicit form for a homotopy on α and β that brings $\tilde{\alpha}$ and $\tilde{\beta}$ to standard form, but this is not difficult to do. Each geodesic segment of $\tilde{\alpha}$ and $\tilde{\beta}$ is represented in P by two segments, the first from the midpoint of a side to the center of P, and the second from the center of P to the midpoint of a (possibly different) side.
We assume we have a homotopy $H_{\text{std},\alpha}$ and $H_{\text{std},\beta}$ that put $\tilde{\alpha}$ and $\tilde{\beta}$ in standard form, respectively. We make a few remarks regarding these homotopies:

- Note that after $H_{\text{std},\beta}$, in general β does not represent a closed geodesic anymore. However, when we have found a homotopy from α to β, we may simply homotope β and α back to the original β by $H^R_{\text{std},\beta}$ to obtain a homotopy from α to (the representative in P of) its homotopic closed geodesic.

- We will denote the geodesic segments that make up $\tilde{\alpha}$ and $\tilde{\beta}$ by $\tilde{\alpha}_0, \dots, \tilde{\alpha}_{m-2}$ and $\tilde{\beta}_0, \dots, \tilde{\beta}_{r-2}$, respectively. Note that these segments do not correspond to those of section 5.1; here we are talking about segments between centers, each of which lies in two neighboring copies of P, while in section 5.1 we used the same notation for a segment contained in a single copy of P. This also implies there is one fewer segment (which is why we use $m-1$ and $r-1$ segments).

5.2.3 *Step 3 : A homotopy from $\tilde{\alpha}$ to $\tilde{\beta}$*

After step 2, the curves $\tilde{\alpha}$ and $\tilde{\beta}$ are in standard form, given by the geodesic segments $\tilde{\alpha}_0, \dots, \tilde{\alpha}_{m-1}$ and $\tilde{\beta}_0, \dots, \tilde{\beta}_{r-1}$ respectively. Moreover, their endpoints are the same. It is now time to calculate a homotopy from $\tilde{\alpha}$ to $\tilde{\beta}$. This will be done in two steps:

a) First, we find a homotopy of $\tilde{\alpha}$ such that the first $r-1$ segments of $\tilde{\alpha}$ become those of $\tilde{\beta}$. This does not mean we are done: after its first $r-1$ segments $\tilde{\alpha}$ may still continue. As $\tilde{\alpha}$ will eventually end at the endpoint of $\tilde{\beta}$, this implies that, after applying the homotopy found in this step, $\tilde{\alpha} = \tilde{\delta}\tilde{\beta}$, where $\tilde{\delta}$ is a closed loop.

b) We find a homotopy that contracts the closed loop $\tilde{\delta}$ to its base point (the endpoint of $\tilde{\beta}$). Extending this homotopy to $\tilde{\alpha} = \tilde{\delta}\tilde{\beta}$ we find a homotopy from $\tilde{\alpha}$ to $\tilde{\beta}$.

See figure 12 for a schematic example of step 3a.

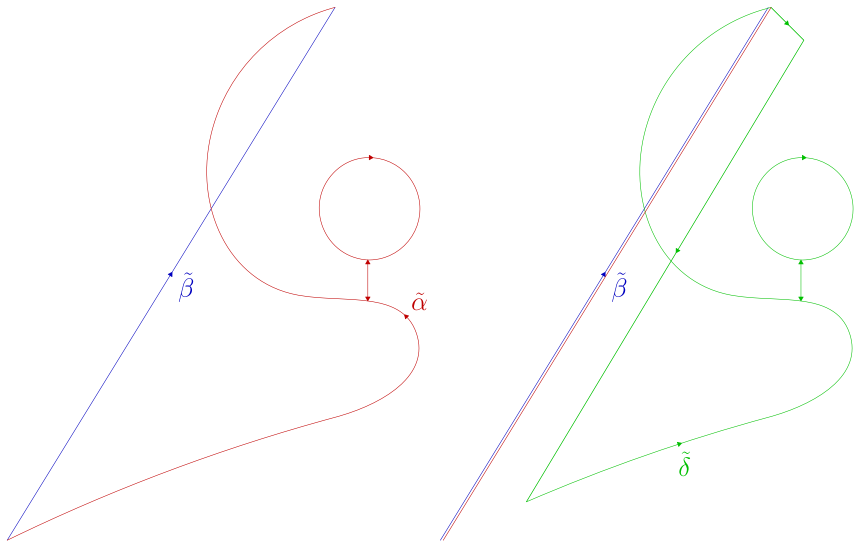

Figure 12: In the left figure we see two curves $\tilde{\alpha}$ and $\tilde{\beta}$ before applying the homotopy of algorithm 3a. After the homotopy we obtain the figure on the right, where the curve $\tilde{\alpha}$ is homotoped to $\tilde{\delta}\tilde{\beta}$ for a closed loop $\tilde{\delta}$ at the endpoint of $\tilde{\beta}$

.

Step 3a: A homotopy for the first $r-1$ *segments of* $\tilde{\alpha}$

The homotopy in this section and the next will be constructed using small building blocks, which we will call *elementary moves*. Each such elementary move is a small algorithm that produces a homotopy $\tilde{\mu}$. The result of algorithm 2.3a and 2.3b will be compositions $\tilde{\mu}_{n-1} \ldots \tilde{\mu}_0$ of homotopies resulting from elementary moves $\tilde{\mu}_0, \ldots, \tilde{\mu}_{n-1}$, that together form the homotopies desired.

For algorithm 2.3a, we will need only one elementary move, which we call $\mathtt{Rotate}$. Suppose that the first k segments of $\tilde{\alpha}$ are already those of $\tilde{\beta}$ for some $0 \leqslant k < r-1$, i.e. $\tilde{\alpha}_i = \tilde{\beta}_i$ for

$i = 0, \ldots, k-1$, but the $(k+1)$-th segment $\tilde{\alpha}_k$ is not $\tilde{\beta}_k$. Since $\tilde{\alpha}$ is in standard form, there are only $4g$ options for the segment $\tilde{\alpha}_k$, given that we know its starting point; one through each side of the copy of P that its starting point lies in. Suppose $\tilde{a}'_j$ is the side of the copy P' that $\tilde{\alpha}_k$ crosses. We will present an elementary move that, given a direction, clockwise or counter-clockwise, replaces the segment $\tilde{\alpha}_k$ by $4g-1$ segments, the first of which is $\tilde{\alpha}_k$ "rotated" in the direction given. That is, it is the segment going through the side $\tilde{a}'_{j+1}$ if the direction was clockwise, and $\tilde{a}'_{j-1}$ if the direction was counter-clockwise. See figure 13. We now present the elementary move in detail.

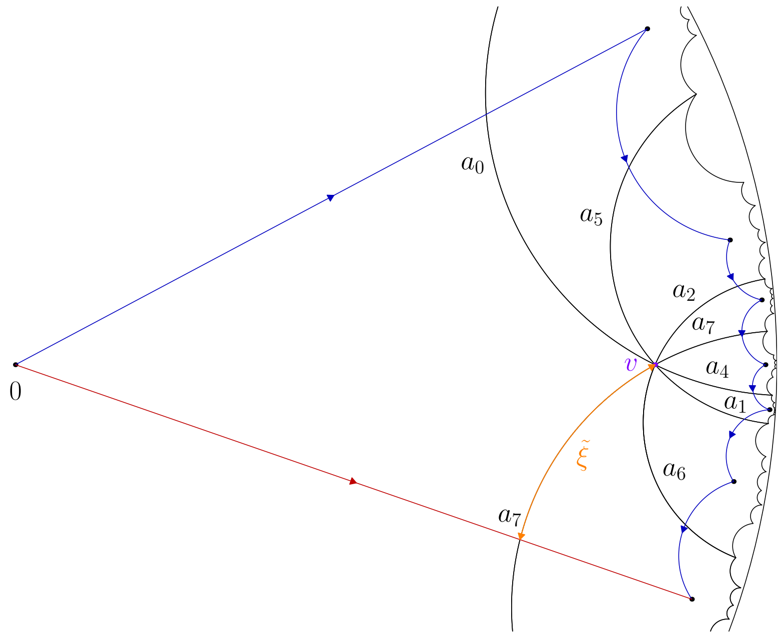

Figure 13: The homotopy $\mathtt{Rotate}(\tilde{\gamma}', \text{counter-clockwise})$ will homotope the red segment $\tilde{\gamma}'$ in the figure to the curve consisting of the sequence of blue segments. Note that $g = 2$ in this example. We first add the segment $\tilde{\xi}$ (in orange), that connects the midpoint of side a_7 with vertex v on the left side of a_7 (in purple). We then bring the curve to standard form, in such a way that the resulting blue curve is such that it takes the other way around the vertex v (when compared to the red segment we started with). The first segment of the blue curve now crosses the side a_0 next to a_7 in the counter-clockwise direction.

Elementary move: Rotate

Input:

- A curve $\tilde{\gamma}$ in $\mathbb{D}$ in standard form, given by the segments $\tilde{\gamma}_0, \ldots, \tilde{\gamma}_n$.
- A specific segment $\tilde{\gamma}'$ of $\tilde{\gamma}$.
- The direction in which we will rotate, denoted by $\mathfrak{o}$, which is clockwise or counter-clockwise.

Output:

- A homotopy of $\tilde{\gamma}$ that replaces the segment $\tilde{\gamma}'$ by $4g - 1$ segments, of which the first is "rotated" in the direction $\mathfrak{o}$ with respect to the original segment $\tilde{\gamma}'$.

Steps:

1. We first homotope $\tilde{\gamma}'$ as to add a segment along the side of P it crosses, to the vertex on the right if the direction $\mathfrak{o}$ given as input is clockwise, and to the vertex on the left if it is anti-clockwise. See figure 13
2. $\tilde{\gamma}$ now includes a vertex which lies in $4g$ copies of P. We homotope $\tilde{\gamma}$ such that $\tilde{\gamma}'$ is deformed into a sequence of segments $\tilde{\gamma}_0, \ldots, \tilde{\gamma}_{4g-2}$ that are each segments between neighboring copies of P, as in figure 13. ∎

Given this elementary move, homotoping the first $r - 1$ segments of $\tilde{\alpha}$ to those of $\tilde{\beta}$ is fairly simple: starting from the first segment, we rotate every segment $\tilde{\alpha}_k$ either clockwise or counter-clockwise until it is the same as the corresponding segment of $\tilde{\beta}$. This gives us the following algorithm:

Algorithm 2.3a

Input:

- A curve $\tilde{\alpha}$ in standard form.
- A curve $\tilde{\beta}$ in standard form, whose endpoints are the same as those of $\tilde{\alpha}$.

Output:

- A homotopy $\tilde{H}_{part}$ that homotopes $\tilde{\alpha}$ to $\tilde{\delta}\tilde{\beta}$, where $\tilde{\delta}$ is a closed loop in standard form at the endpoint of $\tilde{\beta}$.

Steps:

1. We define the variable s to count the number of elementary moves we apply in this algorithm. We set $s = 0$.

2. For $k = 0, \dots, r-2$, we perform the following step:
   - While $\tilde{\alpha}_k \neq \tilde{\beta}_k$, we let $\tilde{\mu}_s = \mathtt{Rotate}(\tilde{\alpha}, \tilde{\alpha}_k, \text{clockwise})$, apply $\tilde{\mu}_s$ to $\tilde{\alpha}$ and increase s by 1.

3. We output the homotopy $\tilde{H}_{\text{part}} = \tilde{\mu}_s \dots, \tilde{\mu}_0$. ∎

Step 3b: Contracting the extraneous part of $\tilde{\alpha}$

We now proceed with the second part of step 3. After applying the homotopy of algorithm 3a, $\tilde{\alpha}$ is given by $\tilde{\delta}\tilde{\beta}$ for some closed loop $\tilde{\delta}$ at the endpoint of $\tilde{\beta}$, in standard form. We will use a combination of two elementary moves to contract the loop $\tilde{\delta}$, RemoveDeadEnd, and RemoveVertex. We will first introduce RemoveDeadEnd, which removes what we call a *dead-end* of $\tilde{\delta}$. We define a dead-end as follows:

Definition 18 *Let $\tilde{\gamma}$ be a curve in $\mathbb{D}$ in standard form, given by the geodesic segments $\tilde{\gamma}_0, \dots, \tilde{\gamma}_{n-1}$. Let P_k be the copy of P whose center is $\tilde{\gamma}_k(0)$, for some $k \in \{1, \dots, n-1\}$. Then if $\tilde{\gamma}_k = \tilde{\gamma}_{k-1}^R$, we call the ordered pair $(\tilde{\gamma}_{k-1}, \tilde{\gamma}_k)$ a dead-end.*

See figure 14 for an example. During the algorithm that finds a homotopy to contract $\tilde{\delta}$ that we will present later, we often require a homotopy to remove a dead-end [6]. During the algorithm to find a homotopy to contract $\tilde{\delta}$ that we will present, we often require a homotopy to remove a dead-end. To this end, we define the elementary move RemoveDeadEnd:

Elementary move: RemoveDeadEnd

Input:

- A curve $\tilde{\gamma}$ in standard form, given by the segments $\tilde{\gamma}_0, \dots, \tilde{\gamma}_{n-1}$.

- A dead-end $(\tilde{\gamma}_{k-1}, \tilde{\gamma}_k)$ of $\tilde{\gamma}$.

6 Note that we do not include the pair $(\tilde{\gamma}_{n-1}, \tilde{\gamma}_0)$ in the definition; in what follows we wish to contract a curve to its base point, so we do not want to remove such a pair.

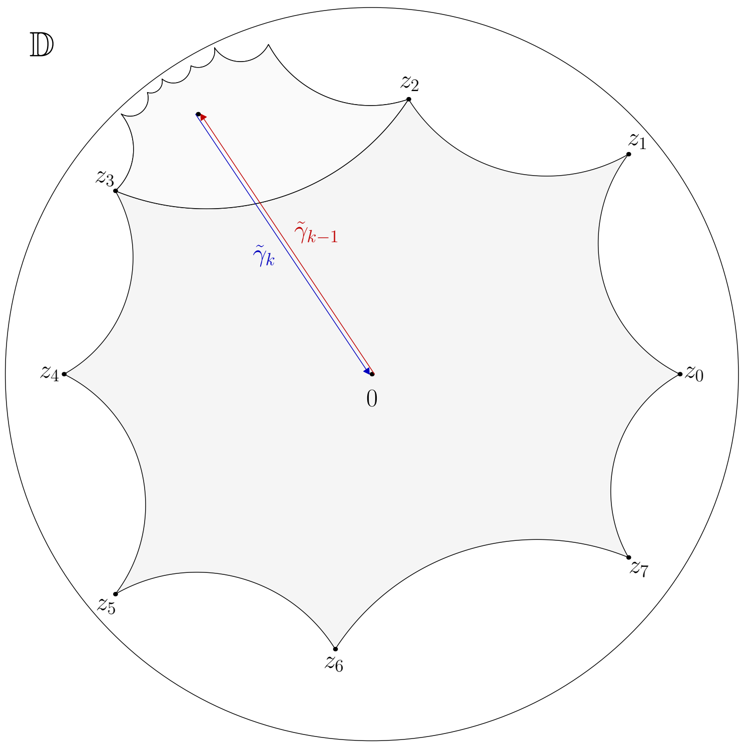

Figure 14: A dead-end $(\tilde{\gamma}_{k-1}, \tilde{\gamma}_k)$ of a curve $\tilde{\gamma}$ (the rest of the curve is omitted). Though in reality the segments overlay one another, we moved them apart a little for clarity of illustration.

Output:

- A homotopy of $\tilde{\gamma}$ that removes the dead-end from $\tilde{\gamma}$.

Steps:

1. We will first give the representative in P of the homotopy that will remove the dead-end. In P, $\tilde{\gamma}_{k-1}$ is represented by two segments, γ_{k-1} from 0 to some midpoint m of a side of P and γ'_{k-1} from the opposite midpoint m' back to 0. Since $\tilde{\gamma}_k$ is $\tilde{\gamma}^R_{k-1}$, it is represented by the segments γ'^R_{k-1} and γ^R_{k-1}. Let γ' be piecewise-continuous curve in P made up from the segments $\gamma_{k-1}, \gamma'_{k-1}$. Then $\gamma'^R\gamma'$ represents the segments $\tilde{\gamma}_{k-1}$ and $\tilde{\gamma}_k$ in P. We apply the following homotopy to $\gamma'^R\gamma'$:

$$H(s,t) = \begin{cases} \gamma'((1-s)2t) & \text{if } 0 \leqslant t \leqslant \frac{1}{2} \\ \gamma'^R((1-s)(2t-1)) & \text{if } \frac{1}{2} < t \leqslant 1 \end{cases}.$$

   This homotopy keeps the endpoints of $\gamma'^R\gamma$ fixed, so we may easily extend this homotopy to a homotopy on γ that leaves the rest of γ untouched. This homotopy reduces $\gamma'^R\gamma'$ to the point $\{0\}$.

2. The homotopy H given above lifts to the following homotopy on $\mathbb{D}$:

$$\tilde{H}(s,t) = \begin{cases} \tilde{\gamma}'((1-s)2t) & \text{if } 0 \leqslant t \leqslant \frac{1}{2} \\ \tilde{\gamma}'^R((1-s)(2t-1)) & \text{if } \frac{1}{2} < t \leqslant 1 \end{cases},$$

and (again, extended to the whole of $\tilde{\gamma}$) clearly removes the dead-end $(\tilde{\gamma}_{k-1}, \tilde{\gamma}_k)$ from $\tilde{\gamma}$. ∎

Apart from removing dead-ends from $\tilde{\alpha}$, our strategy to contract $\tilde{\delta}$ is based on finding and removing what we call "pockets" of $\tilde{\alpha}$. We make the following definitions:

Definition 19 *Let $\tilde{\gamma}$ be a curve in $\mathbb{D}$ in standard form, given by the segments $\tilde{\gamma}_0, \ldots, \tilde{\gamma}_{n-1}$. Then a subcurve of $\tilde{\gamma}$ is a curve in standard form given by the segments $\tilde{\gamma}_i, \ldots, \tilde{\gamma}_j$, $0 \leqslant i < j \leqslant n-1$. We will use the notation $\tilde{\gamma}[i,j]$ for this subcurve.*

Definition 20 *Let $\tilde{\gamma}$ be a closed curve in $\mathbb{D}$ in standard form. Then a pocket of $\tilde{\gamma}$ is a simple closed subcurve of $\tilde{\gamma}$.*

See figure 15 for a schematic example of a pocket. As we will show below, any non-trivial closed curve in standard form which does not have a dead-end will have at least one pocket. The strategy of our algorithm will be to contract these pockets to their begin point. Each pocket will enclose one or more vertices of copies of P. We will introduce an elementary move that, given a pocket of $\tilde{\delta}$, will give us a homotopy of $\tilde{\alpha}$ that removes a vertex of a copy of P from the interior of the pocket (and does not add another).

We have seen that the elementary move $\mathtt{Rotate}$ defined for algorithm 2.3a results in a homotopy which homotopes a curve to go "the other way around" a vertex of a copy of P in $\mathbb{D}$. See again figure 13. Suppose we are given a pocket $\tilde{\gamma}$ of $\tilde{\delta}$. Then its first segment $\tilde{\gamma}_0$ crosses the midpoint of a side of a copy of P, and one of the vertices at the endpoints of this side must lie in the interior of $\tilde{\gamma}$. Which of the two is determined by the orientation of $\tilde{\gamma}$. Thus we may remove a vertex of a copy of P from the interior of $\tilde{\gamma}$ by first determining its orientation $\mathfrak{o}$ and then calculating and applying $\mathtt{Rotate}(\tilde{\gamma}, \tilde{\gamma}_0, \mathfrak{o})$. This is how we will define the elementary move called $\mathtt{RemoveVertex}$.

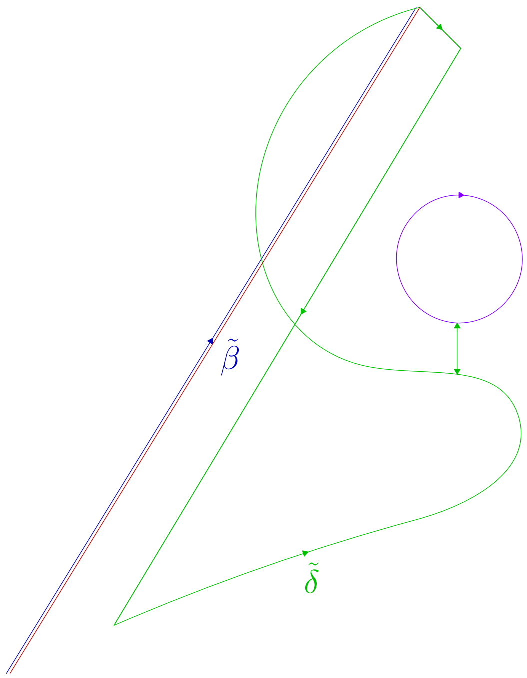

Figure 15: The curve $\tilde{\delta}$ as given in figure 12 has one pocket, indicated here in purple.

Elementary move: $\mathsf{RemoveVertex}$

Input:

- A curve $\tilde{\gamma}$ in standard form, given by the segments $\tilde{\gamma}_0, \dots, \tilde{\gamma}_{n-1}$.
- A pocket $\tilde{\gamma}'$ of $\tilde{\gamma}$ in $\mathbb{D}$, given by the segments $\tilde{\gamma}'_0, \dots, \tilde{\gamma}'_{n-1}$.

Output:

- A homotopy on $\tilde{\gamma}$ which replaces $\tilde{\gamma}'_0$ by a sequence of segments $\tilde{\gamma}'_0, \dots, \tilde{\gamma}'_{4g-2}$ such that, after applying the homotopy to $\tilde{\gamma}$, the vertex v lying in the interior of $\tilde{\gamma}'$ and on the side of the copy of P crossed by $\tilde{\gamma}'_0$ is removed from the interior of $\tilde{\gamma}'$.

Steps:

1. The orientation of $\tilde{\gamma}'$ is determined as follows: we take a point $\tilde{p}$ on $\tilde{\gamma}'$. We then draw a geodesic $\tilde{\chi}$ through the point $\tilde{p}$, transverse to $\tilde{\gamma}'$ at $\tilde{p}$. We include the endpoints ξ_-, ξ_+ on the boundary of $\mathbb{D}$ in $\tilde{\chi}$, and parametrize χ such that $\tilde{\chi}(0) = \xi_-$, $\tilde{\chi}(1) = \xi_+$. Let $\tilde{x}$ be the first intersection of

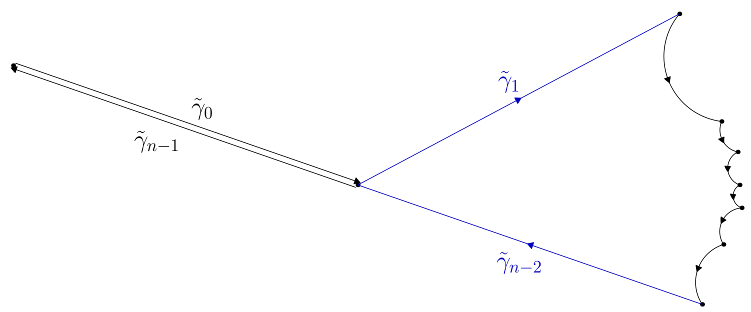

Figure 16: A closed curve $\tilde{\gamma}$ with split $(\tilde{\gamma}_1, \tilde{\gamma}_{n-2})$ (so $k = 1$ in the definition of a split).

$\tilde{\chi}(t)$ with $\tilde{\gamma}'$. The orientation can now be determined by the direction of $\tilde{\gamma}'$ at $\tilde{x}$: if it crosses $\tilde{\chi}$ from the left of $\tilde{\chi}(t)$ to the right of $\tilde{\chi}$, it is anti-clockwise, and if from right to left, clockwise.

2. We calculate and output $\mathtt{Rotate}(\tilde{\gamma}, \tilde{\gamma}_0', \mathfrak{o})$. ■

Our goal is to contract the curve $\tilde{\delta}$ by removing pockets and dead-ends. Before we present an algorithm to remove a pocket by using our elementary moves, we first turn to the matter of finding a pocket of $\tilde{\delta}$ in the first place, and indeed, showing that $\tilde{\delta}$ *has* a pocket. To this end we define a subroutine called `FindPocket`. Before presenting this subroutine we make two more definitions. The first clarifies the notion of the interior of a closed curve in $\mathbb{D}$ that may not be simple.

Definition 21 *Let $\tilde{\gamma}$ be a closed curve in $\mathbb{D}$. Then the interior of $\tilde{\gamma}$ is the set of all points $\tilde{z} \in \mathbb{D}$ such that any continuous curve $\tilde{\zeta} : [0, \mathbb{R}) \to \mathbb{D}$ from $\tilde{z}$ to the boundary of $\mathbb{D}$ (i.e. $\lim_{t\to\infty} \tilde{\zeta}(t) \in \partial\mathbb{D}$) intersects $\tilde{\gamma}$.*

The second definition concerns what we call the "split" of a closed curve in standard form.

Definition 22 *Let $\tilde{\gamma}$ be a closed curve with non-empty interior in $\mathbb{D}$ in standard form, given by the segments $\tilde{\gamma}_0, \dots, \tilde{\gamma}_{n-1}$. Then the split of $\tilde{\gamma}$ is the pair of segments $(\tilde{\gamma}_k, \tilde{\gamma}_{n-1-k})$ such that $\tilde{\gamma}_k$ is the first segment of $\tilde{\gamma}$ satisfying $\tilde{\gamma}_k \neq \tilde{\gamma}^R_{n-1-k}$.*

See Figure 16 for an example. It is easy to see that any closed curve with non-empty interior in standard form must have a split. We are now able to define the subroutine `FindPocket`, as follows:

Subroutine: FindPocket

Input:

- A closed curve $\tilde{\gamma}$ in $\mathbb{D}$ in standard form, given by the segments $\tilde{\gamma}_0, \dots, \tilde{\gamma}_{n-1}$, which has a non-empty interior and no dead-ends.

Output:

- A pocket $\tilde{\gamma}[i, j]$ of $\tilde{\gamma}$.

Steps:

1. We let $\tilde{\gamma}' = \tilde{\gamma}$. This will be our working copy of $\tilde{\gamma}$.
2. We find the split $(\tilde{\gamma}'_{k_0}, \tilde{\gamma}'_{l_0})$ of $\tilde{\gamma}'$. Since $\tilde{\gamma}'$ has non-empty interior, such a split must exist.
3. Now $\tilde{\gamma}'[k_0, l_0]$ must be a closed curve. We calculate its self-intersections. If it has none, $\tilde{\gamma}'[k_0, l_0]$ is a pocket of $\tilde{\gamma}$, and we output $\tilde{\gamma}'[k_0, l_0]$.
4. If $\tilde{\gamma'}$ does have a self-intersection, we take the first self-intersection $\tilde{p}$, and find the first segment $\tilde{\gamma}'_{k_1}$ that leaves the point $\tilde{p}$ (so $\tilde{\gamma}'_{k_1}(0) = \tilde{p}$), as well as the first segment $\tilde{\gamma}'_{l_1}$, $k_1 < l_1$, that gets back to p again (so $\tilde{\gamma}'_{l_1}(1) = \tilde{p}$). Note that $k_1 > k_0$ and $l_1 < l_0$.
5. Since $\tilde{\gamma}$ had no dead-ends, $\tilde{\gamma}'[k_1, l_1]$ must have non-empty interior (otherwise it would contain a dead-end). We let $\tilde{\gamma}' = \tilde{\gamma}'[k_1, l_1]$ and repeat this algorithm from step 2 onward. ∎

To see that the algorithm terminates, note that the number of segments of $\tilde{\gamma}'$ is finite, and strictly decreases every iteration of the algorithm. In particular, this also shows the fact that any closed curve in standard form with no dead-ends has a pocket. Therefore, if we first remove $\tilde{\delta}$'s dead-ends using RemoveDeadEnd, we can be sure it has a pocket, and we may find one using FindPocket.

As said, $\tilde{\delta}$ will be contracted by removing pockets. To remove a pocket $\tilde{\gamma}$, we define a subroutine RemovePocket. It will use RemoveVertex to remove the vertices of copies of P inside $\tilde{\gamma}$. After removing a vertex, dead-ends may have been created, which we can remove by RemoveDeadEnd. It may also be however that $\tilde{\gamma}$ is no longer simple, and hence no longer a pocket. In this case we use $\mathtt{FindPocket}$ to find a pocket of $\tilde{\gamma}$, and apply $\mathtt{RemovePocket}$ to the found pocket [7]. It is easy to see that every iteration of the subroutine removes a vertex of a copy of P from the interior of $\tilde{\gamma}$, while not adding any. As a pocket contains only finitely many vertices of copies of P, the algorithm will terminate.

Subroutine: $\mathtt{RemovePocket}$

Input:

- A curve $\tilde{\gamma}$ in standard form.
- A pocket $\tilde{\gamma}'$ of $\tilde{\gamma}$.

Output:

- A homotopy that homotopes $\tilde{\gamma}$ so that $\tilde{\gamma}'$ is removed.

Steps:

1. We let s be a counter for the elementary moves that are found in this algorithm, and set $s = 0$.
2. We apply and list $\tilde{\mu}_s = \mathtt{RemoveVertex}(\tilde{\gamma}, \tilde{\gamma}')$, and increase s by 1.
3. Applying $\mathtt{RemoveVertex}$ may have created dead-ends and/or self-intersections in $\tilde{\gamma}'$. While there are dead-ends $(\tilde{\gamma}_{k-1}, \tilde{\gamma}_k)$ in $\tilde{\gamma}'$ we calculate $\tilde{\mu}_s = \mathtt{RemoveDeadEnd}(\tilde{\gamma}, (\tilde{\gamma}_{k-1}, \tilde{\gamma}_k))$, apply the result and increase s by 1.
4. If $\tilde{\gamma}'$ has not been contracted to its starting point, we calculate $\tilde{\gamma}'' := \mathtt{FindPocket}(\tilde{\gamma}')$, and then calculate $\mathtt{RemovePocket}(\tilde{\gamma}, \tilde{\gamma}'')$.
5. We output the homotopy $\tilde{\mu}_s \cdots \tilde{\mu}_0$. ■

We are now ready to present algorithm 2.3b, which finds a homotopy of $\tilde{\alpha} = \tilde{\delta}\tilde{\beta}$ that contracts $\tilde{\delta}$ to its starting point. In figure 17 we see a schematic example of the two main steps of the algorithm, removing pockets and removing dead-ends.

7 Note that we define $\mathtt{RemovePocket}$ recursively in this way.

Algorithm 2.3b

Input:

- A curve $\tilde{\alpha} = \tilde{\delta}\tilde{\beta}$ in standard form, where $\tilde{\delta}$ is a closed loop at the endpoint of $\tilde{\beta}$.

Output:

- A homotopy $\tilde{H}_{\text{contr}}$ from $\tilde{\alpha}$ to $\tilde{\beta}$.

Steps:

1. We define the variable s as a counter for the elementary moves, and set $s = 0$.
2. While $\tilde{\alpha} \neq \tilde{\beta}$, we repeat the following steps:
   a) While $\tilde{\delta}$ has a dead-end $(\tilde{\delta}_k, \tilde{\delta}_{k+1})$, we let $\tilde{\mu}_s = \mathtt{RemoveDeadEnd}(\tilde{\alpha}, (\tilde{\delta}_k, \tilde{\delta}_{k+1}))$. We apply $\tilde{\mu}_s$ to $\tilde{\alpha}$ and increase s by 1.
   b) If $\tilde{\alpha}$ is now $\tilde{\beta}$, we output $\tilde{\mu}_s \cdots \tilde{\mu}_0$ and terminate the algorithm.
   c) We calculate $\chi = \mathtt{FindPocket}(\tilde{\delta})$.
   d) We calculate and apply $\tilde{\mu}_s = \mathtt{RemovePocket}(\tilde{\alpha}, \chi)$, and increase s by 1.
3. We output $\tilde{H}_{\text{contr}} = \tilde{\mu}_s \cdots \tilde{\mu}_0$ and terminate the algorithm. ■

To see that this algorithm terminates, note that $\tilde{\delta}$ has a finite number of closed subcurves. Each time we remove a pocket, we remove a closed subcurve and do not create a new closed subcurve. Hence the algorithm will terminate.

5.2.4 *Step 4: Back to the original* $\tilde{\beta}$

Now all that is left is to homotope $\tilde{\alpha}$ back to the original version of $\tilde{\beta}$, i.e. $\tilde{\beta}$ before we put it in standard form. This comes down to simply applying $\tilde{H}^R_{\text{std},\tilde{\beta}}$, the reverse of the homotopy $\tilde{H}_{\text{std},\tilde{\beta}}$ that we applied to $\tilde{\beta}$ in section 5.2.2.

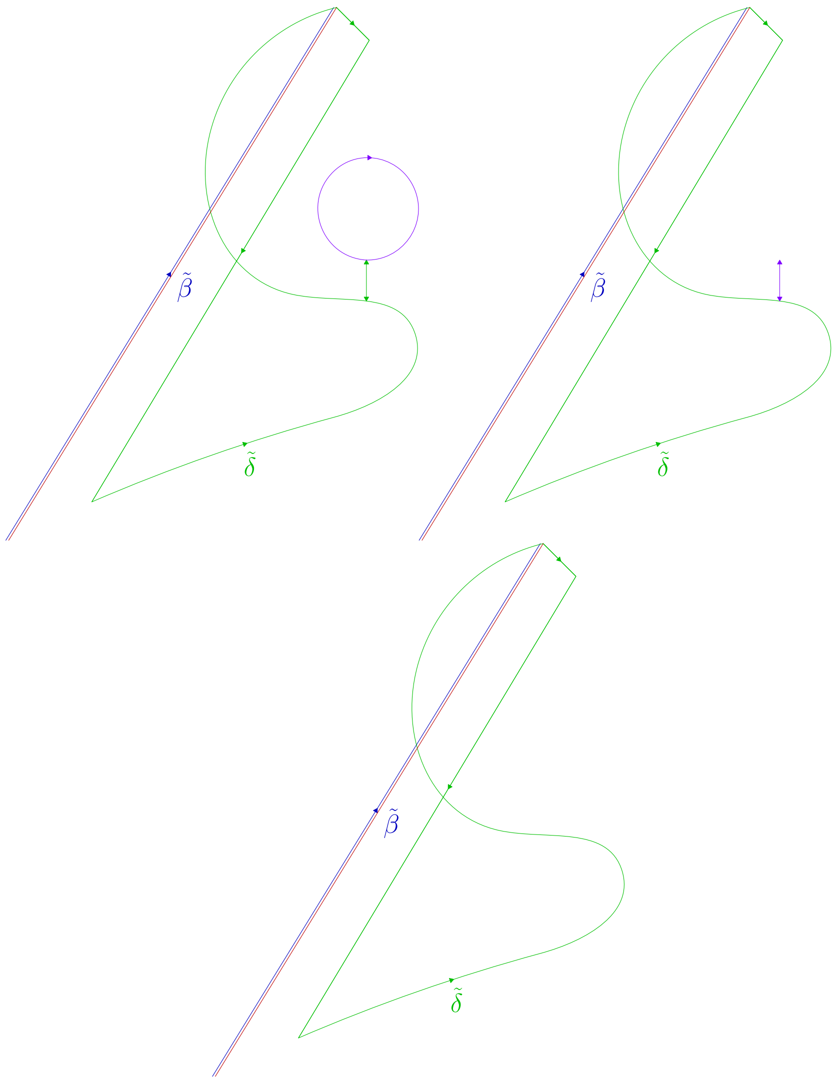

Figure 17: In first figure we have found a pocket of $\tilde{\delta}$, indicated in purple. In the second figure we remove this pocket by RemovePocket. A dead-end remains, which is removed in the third figure by RemoveDeadEnd.

5.2.5 *Step 5: Output*

The end result of algorithm 2 is the homotopy $H = H^R_{std,\beta} H_{contr} H_{part} H_{end} H_{std,\alpha}$, where $H_{std,\beta}$, H_{contr}, H_{part}, $H_{std,\alpha}$ and H_{end} are the homotopies on P representing $\tilde{H}_{std,\tilde{\beta}}$, $\tilde{H}_{contr}$, $\tilde{H}_{part}$, $\tilde{H}_{std,\tilde{\alpha}}$ and $\tilde{H}_{end}$, respectively. It homotopes α to β, as desired.

6 DISCUSSION

At the outset, the goal of this master project was the following: to generalise the results of Aigon-Dupuy's article on the description of Teichmüller space of genus 2 by admissible polygons to higher genus. The hope was that finding such a generalisation would allow us to use the tools provided in Aigon-Dupuy's description to facilitate computation on closed Riemann surfaces of higher genus.

We realised that just taking symmetric $4g$-gons would not work, as we do not get the required $6g-6$ real parameters to parametrize Teichmüller space. Our initial attempt, not described in thesis, was to take three of its vertices, mirror these and leave the rest to vary. This seemed to give us the required $6g-6$ parameters. We were not able, however, to show that this construction would actually parametrize Teichmüller space. We tried to relate the construction to that of Zieschang-Vogt-Coldewey, which also provides coordinates to Teichmüller space using hyperbolic polygons, but could not make progress.

We decided to switch focus to the symmetric polygons of higher genus. These we found, are known to correspond to the closed hyperelliptic Riemann surfaces. We made a study of this result, which was first proven by Schaller.

Now, given this relation between admissible polygons and closed hyperelliptic Riemann surfaces, we wished to see whether the results of Aigon-Dupuy would carry over to this more general case. We went through the arguments of Aigon-Dupuy proving criteria on how we may choose vertices of admissible octagons, and showed that, apart from some small adaptations, the same results hold for admissible polygons of higher genus.

In looking ahead through the rest of the article of Aigon-Dupuy, it seemed that its results, apart from section (viii) on the generators of the modular group, would generalise almost immediately. We therefore thought it to be more interesting to spend the remaining time of the project on a computational problem

in which we might apply the studied parametrization of hyperelliptic surfaces by admissible polygons. The problem we settled on was determining a homotopic closed geodesic to a curve on a closed hyperelliptic Riemann surface, and if possible, find a homotopy explicitly.
There are some comments to be made regarding the algorithms we presented. The algorithms have not been implemented and tested, the complexity of the algorithms has not been analysed and they are probably not as efficient as they could be. They are still, very much, theoretical algorithms.

Furthermore, there are different approaches one might take in the construction of the algorithms. A different idea for the algorithm to find vertices inside $\tilde{\delta}$ is a so-called sweep-line algorithm. With a $\tilde{\delta}$ in the Euclidean plane and the case of the torus, this would work as follows: we start with the vertex of $\tilde{\delta}$ that is leftmost, i.e. has the lowest x coordinate. We then work rightwards. We know that each gridpoint between the edges that leave the leftmost vertex will be inside $\tilde{\delta}$. Moving rightwards, and taking care to note where new vertices appear and where edges come together, and which gridpoints we encounter inside $\tilde{\delta}$ along the way, we can find all vertices inside $\tilde{\delta}$.
The problem in the hyperbolic case, however, is that we do not know beforehand where our gridpoints, the vertices of copies of P, lie. We could try to calculate them along the way, and probably this could work as well, though you might have to calculate and check points that lie left of your sweepline, making the process messier.

In this thesis we used the approach as described in section 5.2. Follow up studies are recommended to explore such algorithms further and provide actual implementation.

BIBLIOGRAPHY

[1] Aline Aigon-Dupuy, Peter Buser, Michel Cibils, Alfred F. Künzle, and Frank Steiner. "Hyperbolic octagons and Teichmüller space in genus 2." In: *Journal of Mathematical Physics* 46.3 (2005). DOI: 10.1063/1.1850177.

[2] Alan F. Beardon. *The geometry of discrete groups*. Springer-Verlag, 1983, p. 337. ISBN: 0387907882.

[3] Peter Buser. *Geometry and Spectra of Compact Riemann Surfaces (Reprint of the 1992 Edition)*. Tech. rep.

[4] Yoichi Imayoshi and Masahiko Taniguchi. *An introduction to Teichmüller spaces*. Springer-Verlag, 1992, p. 279. ISBN: 4431700889.

[5] John H. Hubbard. *Teichmüller Theory and Applications to Geometry, Topology, and Dynamics. Volume 1 Teichmüller Theory*. Matrix Editions, 2006.

[6] Henri Poincaré. *Oeuvres, Tome II*. Les Grands Classiques Gauthier-Villars, 1916.

[7] Paul Schmutz Schaller. "Teichmüller Space and Fundamental Domains of Fuchsian Groups." In: *L'Enseignement mathématique* 45 (1999).