2015-03-01

The Steiner Problem on Closed Surfaces of Constant Curvature

Andrew Logan

Brigham Young University - Provo

The Steiner Problem on Closed Surfaces of Constant Curvature

Andrew E. Logan

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Denise M. Halverson, Chair
Gary R. Lawlor
Michael J. Dorff

Department of Mathematics

Brigham Young University

March 2015

ABSTRACT

The Steiner Problem on Closed Surfaces of Constant Curvature

Andrew E. Logan
Department of Mathematics, BYU
Master of Science

The n-point Steiner problem in the Euclidean plane is to find a least length path network connecting n points. In this thesis we will demonstrate how to find a least length path network T connecting n points on a closed 2-dimensional Riemannian surface of constant curvature by determining a region in the covering space that is guaranteed to contain T. We will then provide an algorithm for solving the n-point Steiner problem on such a surface.

ACKNOWLEDGMENTS

I would like to primarily thank my God for the people He has placed in my life and the gifts that He has given me that have helped me with this thesis. I know that I am nothing without Him.

He has blessed me with an incredible spouse. I am grateful for her support and understanding throughout the process of writing this thesis. Eternity is going to be wonderful.

He has blessed me with an amazing adviser. I know I would not have made it this far if it were not for Dr. Halverson's patience. Having her as an adviser for the past four years has been an immense blessing. Her training has helped me develop the critical thinking skills I needed to succeed as a mathematician.

He has inspired wonderful research students to join Dr. Halverson's Origami Math Research Group. They have given me wonderful advice on my research interests. They are all great examples to me.

He has blessed me with a supportive family and many encouraging friends. I am grateful for their prayers for me. One such friend who is especially worthy of mention is Nick Callor. He helped me through several ideas and edits pertaining to this thesis. I am indeed grateful for his friendship.

CONTENTS

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

The n-point Steiner problem (named after Jakob Steiner (1796-1863)) on a given surface X is to find the least length path network connecting n points on X. Although the Steiner problem in the plane for the 3-point case was purposed by Fermat and solved by Torricelli in the 17th century [18], it became popularized in 1979 as the Steiner problem by Courant and Robbins who attributed it to Jakob Steiner[7]. Many results for the Steiner problem on the plane soon developed [5, 9, 17, 18, 19, 25]. Ever since its popularization, the underlying goal has been to solve the Steiner problem on all surfaces. In this thesis, we look at solving the Steiner problem on closed surfaces of constant curvature.

Recent studies address the Steiner problem on some closed surfaces of constant curvature. Such closed surfaces include, but are not limited to, the 2-sphere, the flat torus, the Klein bottle, and the g-holed torus for some $g \in \mathbb{N}$. Ivanov and Tuzhilin provide classifications for closed minimal path networks (networks on closed graphs) on closed surfaces of constant nonnegative curvature in chapter 8 of [18], but do not provide an algorithm that solves the Steiner problem on these surfaces. Many results regarding the Steiner problem on the 2-sphere have been studied by [2, 6, 8]. Results on the Steiner problem on a flat torus were obtained in [14, 22]; and Halverson and Logan solve the n-point Steiner problem on a flat torus $\mathcal{T}$ and provide a complete algorithm for solving the 3-point Steiner problem on $\mathcal{T}$ in [11]. Results on surfaces in general, which we use in this thesis, can be found in [4, 26, 27].

The Steiner problem in the plane is $\mathcal{NP}$-complete [18]. The Steiner problem is even more difficult on non-simply connected surfaces for which there is a "wrapping" issue. In non-simply connected surfaces geodesics (see Subsection 2.5.1) connecting two points may not be unique. In fact there may be infinitely many such geodesics on some surfaces, i.e. surfaces of constant nonpositive curvature. Many authors address this wrapping issue while solving the Steiner problem on surfaces such as cones [20], cylinders [10], tetrahedron [3, 13], and more.

The main results of this thesis provide tools to reduce to a finite number the geodesics considered to find a least length geodesic. With these tools, we also provide an algorithm for solving the Steiner problem on closed surfaces of constant curvature that have a certain property discussed in Section 2.4.

Chapter 2. Preliminaries

The following preliminaries are standard results from algebraic topology, graph theory, and Riemannian geometry. Readers familiar with these subjects are invited to skim through this chapter, but pay particular attention to the section on fundamental domains.

2.1 Covering Spaces

The results in this section can be found in [21]. We assume that all spaces in this section are arcwise connected and locally arcwise connected. Let X be a topological space. A *covering space* of X is a pair consisting of a space $\tilde{X}$ and a continuous map $p : \tilde{X} \to X$ called a *covering map* such that the following condition holds: Each point $x \in X$ has an arcwise-connected open neighborhood U such that each arc component of $p^{-1}(U)$ is mapped topologically onto U by p.

A continuous map $f : X \to Y$ is a *local homeomorphism* if each point $x \in X$ has an open neighborhood V such that $f(V)$ is open and f maps V topologically onto $f(V)$. In a locally arcwise connected space, the arc components of an open set are open. Hence if $(\tilde{X}, p)$ is a covering space of X, then p is a local homeomorphism.

Lemma 2.1.1. *Let $(\tilde{X}, p)$ be a covering space of X, let A be a subspace of X which is arcwise connected and locally arcwise connected, and let $\tilde{A}$ be an arc component of $p^{-1}(A)$. Then $(\tilde{A}, p|\tilde{A})$ is a covering space of A.*

Definition 2.1.2. Let $p : X \to Y$ be a covering map and let Z be a subset of Y. A

continuous map $f : Z \to X$ is said to be a *lift of the inclusion map* $Z \hookrightarrow Y$ provided, for all $z \in Z$, $z = p \circ f(z)$. We also say the set $f(Z)$ *is a lift of Z.*

Lemma 2.1.3. *Let* $(\tilde{X}, p)$ *be a covering space of* X, $\tilde{x}_0 \in \tilde{X}$, *and* $x_0 = p(\tilde{x}_0)$. *Then for any path* $f : I \to X$ *with initial point* x_0, *there exists a unique path* $g : I \to \tilde{X}$ *with initial point* $\tilde{x}_0$ *such that* $pg = f$.

Theorem 2.1.4. *Let* X *be a topological space which is connected, locally arcwise connected, and semilocally simply connected. Then there exists a covering space* $(\tilde{X}, p)$ *of* X.

Definition 2.1.5. Any compact surface M is a quotient space of a polygonal disc D under a map

$$f : D \to M$$

which identifies certain edges of D in pairs. If $(\tilde{M}, p)$ is any covering space of M, then there exist lifts

$$f_i : D \to \tilde{M}, \qquad pf_i = f,$$

of f. Then, $\tilde{M}$ has the largest topology which makes all the f_i's continuous. The images $f_i(D)$ cover $\tilde{M}$. Each such image is called a *fundamental domain* in $\tilde{M}$ [21].

2.2 Graph Theory

The definitions and results in this section can be found in [21, 24].

A *graph* is a topological space which consists of points, called *vertices*, and a collection of *edges*. Each edge is either homeomorphic to an interval of the real line and joins two distinct vertices, or it is homeomorphic to a circle and joins a given vertex to itself. An edge e connecting vertices v_1 and v_2 may be assigned one of two different *orientations*; either

   (i) e begins at initial point v_1 and moves towards terminal point v_2, or

   (ii) e begins at initial point v_2 and moves towards terminal point v_1.

3

A *tree* is a connected graph that contains no closed edge paths (cycles). Trees have the following properties:

(i) Any connected subgraph of a tree is also a tree.

(ii) Any two distinct vertices can be joined by a unique reduced edge path.

Theorem 2.2.1. *Let X be a connected graph with vertex set X^0, let (Y, p) be a covering space of X, and let $Y^0 = p^{-1}(X^0)$. Then Y is a graph with vertex set Y^0.*

2.3   SIMPLICIAL COMPLEXES

We introduce definitions in this section to help us clearly define a compact region in a covering space of a closed surface X of constant curvature that is guaranteed to contain the lift of a minimal path network connecting n points in X (see Section 4.1). These definitions can be found in [23].

Let $\mathcal{G}$ be a geometric space (such as the 2-sphere, Euclidean space, or hyperbolic space). A set $\{a_0, \ldots a_n\}$ of points in $\mathcal{G}$ is said to be *geometrically independent* if for any scalars t_i, the equations

$$\sum_{i=0}^{n} t_i = 0 \quad \text{and} \quad \sum_{i=0}^{n} t_i a_i = \mathbf{0}$$

imply that $t_0 = t_1 = \cdots = t_n = 0$.

Let $\{a_0, \ldots a_n\}$ be a geometrically independent set. The *n-simplex σ spanned by $a_0, \ldots a_n$* is the set of all points x in $\mathcal{G}$ such that

$$x = \sum_{i=0}^{n} t_i a_i \quad \text{where} \quad \sum_{i=0}^{n} t_i = 1$$

and $t_i \geq 0$ for all i.

Let σ be an n-simplex spanned by $a_0, \ldots a_n$. A topological space X is called a *topological n-dimensional simplex* if there is a homeomorphism $\phi : \sigma \to X$.

4

The points $a_0 \ldots a_n$ that span the n-simplex σ are called the *vertices* of σ, and n is called the *dimension* of σ. Any simplex spanned by a subset of $\{a_0, \ldots a_n\}$ is called a *face* of σ. In particular, the face of σ spanned by $a_1, \ldots a_n$ is called the face *opposite* a_0. The faces of σ different from σ itself are called the *proper faces* of σ; their union is called the *boundary* of σ and is denoted Bd σ. The *interior* of σ is defined by the equation $\sigma^o = \sigma - \text{Bd } \sigma$.

A *simplicial complex* K in $\mathcal{G}$ is a collection of simplices in $\mathcal{G}$ such that:

(i) Every face of a simplex of K is in K.

(ii) The intersection of any two simplices of K is a face of each of them.

Definition 2.3.1. If v is a vertex of K, the *star* of v in K, denoted by $\text{St}(v)$, is the union of the interiors of those simplices of K that have v as a vertex. Its closure, denoted $\overline{\text{St}}(v)$, is called the *closed star* of v in K. It is the union of all simplices of K having v as a vertex. If $A \subset \mathcal{G}$, then $\text{St}(A)$ is the union of the interiors of those simplices of K whose intersection with all the vertices in $\overline{A}$ is nonempty. We can thus define $\text{St}^k(v) = \text{St}^{k-1}(\text{St}(v))$ for any $k \in \mathbb{N}$.

2.4 Fundamental Domains

In this section we take several ideas as presented in [1]. These ideas require a metric structure so we adopt the typical definitions and theories of metric spaces. Results in this thesis depend heavily on the concept of isometry of metric spaces.

Definition 2.4.1. Let (X, d_X) and (Y, d_Y) be metric spaces. A map $p : X \to Y$ is an *isometry*, if $d_X(x, y) = d_Y(p(x), p(y))$ for all $x, y \in X$. If such a map p exists we say X is *isometric* to Y and likewise say that p is an *isometric map* from X to Y. Furthermore we say p is a *local isometry* if for any $x \in X$ there exists an open subset U of X containing x such that $p|_U : U \to Y$ is an isometry.

A collection of subsets in a topological space X is *locally finite* if each point in X has a neighborhood that intersects only finitely many sets in the collection. A *polygonal tiling*

is a locally finite collection of polygonal regions $\mathcal{P}$ with pairwise disjoint interiors so that $\bigcup_{P \in \mathcal{P}} P$ is the entire space (typically Euclidean or hyperbolic). A tiling is said to be *regular* if each element of $\mathcal{P}$ is a regular n-gon.

Suppose the polygon F is the only polygon in a polygonal tiling $\mathcal{P}$. Then we say F is a *fundamental tile of $\mathcal{P}$*.

Let $(\tilde{X}, p)$ be a covering space for a surface X. Note that each fundamental domain F_0 from Section 2.1 is a fundamental tile of $\tilde{\mathcal{P}}_{F_0}$ where $\tilde{\mathcal{P}}_{F_0}$ covers $\tilde{X}$. To emphasize that these fundamental domains are polygons we refer to such a fundamental domain as a *fundamental polygon.*

As a remark it is not necessary for a fundamental polygon to be a regular polygon (in hyperbolic space a polygon is said to be regular if all edges have the same length and every angle has the same measure).

Definition 2.4.2. A fundamental polygon F is said to have a *center point* q if the perpendicular bisectors of the sides of F meet at $q \in F$. We denote such a fundamental polygon by $F[q]$.

Note that $F[q]$ is unique if the tiling of $\tilde{X}$ has been established.

2.5   RIEMANNIAN MANIFOLDS

Let n be a positive integer. An *n-dimensional manifold* is a Hausdorff space such that each point has an open neighborhood homeomorphic to the open n-dimensional disc [21]

$$U^n = \{x \in \mathbb{R}^n : \|x\| < 1\}.$$

In this section we will mostly refer to results in [18].

We say two curves passing through the point x_0 are in the same *equivalence class* if they have the same velocity vector through x_0.

Definition 2.5.1. An equivalence class of curves passing through a point $x_0 \in M$ is called a *tangent vector to manifold M at x_0*. The set of all vectors tangent to M at a point x_0 is called the *tangent space to M at x_0* and is denoted by $T_{x_0}M$.

Let $x \in M$ be a point of a smooth n-dimensional manifold M. Let us consider the set of all linear mappings from $T_x M$ to $\mathbb{R}$. This set forms an n-dimensional linear space called the *cotangent space* (also called the dual space of $T_x M$) and is denoted $T_x^* M$. Mappings that are linear on any Cartesian factor of the Cartesian products of vector spaces are called *tensors*. In the case of smooth manifolds these mappings act on the Cartesian products of p cotangent spaces and q tangent spaces and are called *tensors of type (p,q)* [18].

Definition 2.5.2. A tensor of type (0,2) giving an inner product on each tangent space is called a *Riemannian metric*. A manifold endowed with a Riemannian metric is called *Riemannian*.

2.5.1 Geodesics. The Riemannian metric enables us to measure the lengths of curves on smooth Riemannian manifolds. We define the distance between two points x and y in a space X denoted $d_X(x,y)$ to be the greatest lower bound of the lengths of the curves joining these points. Suppose for a given pair of points the infimum is obtained by γ. We call γ and all other locally length minimizing curves *geodesics*. The following are properties of geodesics.

Proposition 2.5.3.

(i) *For any point of a Riemannian manifold, and for any direction (nonzero tangent vector) at that point there exists a geodesic starting at that point and going in that direction.*

(ii) *If two geodesics go out from the same point in the same direction then one of them is included in the other one.*

(iii) *The length of the arc of a geodesic between its points at a sufficiently short distance equals the distance between the points.*

(iv) *For any point of a Riemannian manifold there exists a neighborhood U such that any two points from U can be joined by a shortest geodesic included in U.*

(v) *On a complete (as a metric space) Riemannian manifold any two points can be joined by a geodesic.*

(vi) *If the length of a curve connecting two given points of a complete Riemannian manifold equals the distance between the points, then the curve is a geodesic (of the shortest possible length).*

For more results on geodesics on two-dimensional Riemannian manifolds see [15].

2.5.2 Classification of closed surfaces of constant curvature. An important characteristic of a Riemannian manifold is its curvature, or more specifically Gaussian curvature. In this thesis we will only refer to curvature of 2-dimensional Riemannian manifolds. For a more formal definition of curvature see pages 48 through 51 of [18]. We also introduce *closed* manifolds which are compact manifolds without boundary (see [18] pages 12 through 15). The propositions discussed in this subsection are found in chapter 8 of [18].

Consider the standard sphere $S^2(r)$ of radius r embedded in $\mathbb{R}^3$. We can calculate the curvature of $S^2(r)$ to be $1/r^2$.

Furthermore we can look at the projective plane $\mathbb{R}P^2$ which can be realized as a quotient space $S^2(r)/\mathbb{Z}_2$. The curvature of $\mathbb{R}P^2$ then equals the curvature of the sphere $S^2(r)$. Such a projective plane we denote by $\mathbb{R}P^2(r)$

Proposition 2.5.4. *Let W be a closed two-dimensional Riemannian manifold of constant positive curvature $1/r^2$. Then W is isometric either to the sphere $S^2(r)$, or to the projective plane $\mathbb{R}P^2(r)$.*

Consider a parallelogram that is the span of a pair of vectors e and f, and glue from it a closed two-dimensional surface by the word $aba^{-1}b^{-1}$ where consecutive letters correspond to consecutive sides of the parallelogram, and same letters correspond to glued sides. The letter

8

is raised to a power of -1 if the orientation of the corresponding side is opposite to the given orientation of the contour of the parallelogram. The obtained surface is a two-dimensional flat torus denoted by $T^2(e, f)$.

If the parallelogram is glued by the work $aba^{-1}b$, then the resulting surface is a flat Klein bottle denoted by $K^2(e, f)$.

Proposition 2.5.5. *Let W be a closed two-dimensional Riemannian manifold of zero curvature. Then W is isometric either to the two-dimensional torus $T^2(e, f)$, or the the Klein bottle $K^2(e, f)$, for some vectors e and f.*

Let P_g be some $4g$-gon, where $g > 1$. Write the word

$$w = a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_g b_g a_g^{-1} b_g^{-1},$$

and suppose that $l(c_i) = l(c_i^{-1})$ where $c_i \in \{a_i, b_i\}$ and c_i^{-1} indicates the opposite orientation of c_i. Glue the two-dimensional surface by the word w. We get the surface of genus g which is a smooth Riemannian manifold of constant negative curvature k. Such surfaces are denoted by $M_g(k)$. Note that topologically $M_g(k)$ is a g holed torus of constant curvature k.

Proposition 2.5.6. *Let W be a closed two-dimensional oriented Riemannian manifold of constant negative curvature k. Then W is isometric to some surface $M_g(k)$.*

We close this section with some fundamental properties of Riemannian surfaces.

Proposition 2.5.7. *Any closed surface of constant curvature will have one of $\mathbb{R}^2$, $\mathbb{H}^2$, or S^2 as a universal covering space.*

A family of open subsets $\{U_\alpha, V_\beta\}$ covering a locally Euclidean space X (i.e. a smooth manifold) such that each U_α is homeomorphic to $\mathbb{R}^{n_\alpha}$, but each V_β is homeomorphic to $\mathbb{R}_+^{n_\beta}$ (where $\mathbb{R}_+^n := \{(x_1, \ldots, x_n) : x_1 \geq 0\}$), together with the corresponding homeomorphisms $\phi_\alpha : U_\alpha \to \mathbb{R}^{n_\alpha}$ and $\psi_\beta : V_\beta \to \mathbb{R}_+^{n_\beta}$, is called an *atlas* of X consisting of the *charts* $\{(U_\alpha, \phi_\alpha)\}$

and $\{(V_\beta, \psi_\beta)\}$. If (U_1, ϕ_1) and (U_2, ϕ_2) are two charts such that the intersection $U_1 \cap U_2$ is nonempty, then two coordinate systems appear in $U_1 \cap U_2$. The homeomorphism $\phi_2 \circ \phi_1^{-1}$ is called the *transition function* that sets the rule of transition from one coordinate system to the other.

A smooth manifold M is called *orientable* if there exists an atlas such that all its transition functions have positive Jacobian. If no such atlas exists, then M is *non-orientable*. Examples of orientable surfaces are the 2-sphere, the flat torus, and the g-holed torus. Examples of non-orientable surfaces are the projective plane and the Klein bottle.

Proposition 2.5.8. *Every non-orientable surface has an orientable double cover.*

Chapter 3. The Steiner Problem

In this section we will discuss general results regarding the n-point Steiner problem on a smooth surface and algorithms for solving an n-point Steiner problem on the 2-dimensional sphere S^2, Euclidean plane $\mathbb{R}^2$, and hyperbolic plane $\mathbb{H}^2$.

3.1 General notations, definitions, and results

The following definition makes sense in the context of the path networks we are considering in this paper, which can be viewed as a finite collection of distinct rectifiable paths.

Definition 3.1.1. Let T be the union of a finite collection of rectifiable paths in $\mathbb{R}^2$. Then the *length* of T, denoted $l(T)$, is the sum of the lengths of the paths in the collection.

Definition 3.1.2. A *Steiner minimal tree connecting* A_1, A_2, $\ldots$, A_n in a path-connected space $\mathcal{X}$ is a minimal length path network connecting A_1, A_2, $\ldots A_n$ in $\mathcal{X}$. We denote the set of all such Steiner minimal trees as $SMT(A_1, A_2, \ldots, A_n)$.

Remark 3.1.3. In the plane, A Steiner minimal tree connecting points A_1, $A_2, \ldots, A_n$ is known to be unique [18] for $n \leq 3$, and hence the set $SMT(A_1, A_2, \ldots, A_n)$ contains only

one element. In the case $n = 3$ the unique element of $SMT(A, B, C)$ is contained in the convex hull of $\triangle ABC$ [10]. If $n > 3$, then $SMT(A_1, A_2, \ldots, A_n)$ may contain more than one element.

In an arbitrary surface, $SMT(A_1, A_2, \ldots, A_n)$ may not be unique for any $n \geq 2$. For example, and relevant to this paper, in the flat torus, given any two specified points, there are infinitely many geodesic segments connecting them.

In surfaces $S^2(r)$, $\mathbb{R}^2$, and $\mathbb{H}^2$ a solution to the 3-point Steiner problem has several characteristics. If an interior angle of $\triangle ABC$ (the inner triangle in $S^2(r)$) is greater than or equal to $120°$, then an element of $SMT(A, B, C)$ is the union of the shortest two sides of $\triangle ABC$ [18]. In this case the Steiner minimal tree is said to be *degenerate*. In the case that no interior angle of $\triangle ABC$ is greater than or equal to $120°$, then the solution to the 3-point Steiner problem requires a forth point S, called the *Steiner point*, such that an element of $SMT(A, B, C)$ is $\overline{AS} \cup \overline{BS} \cup \overline{CS}$. In this case the Steiner minimal tree is said to be *full*. It is a classical result, observed by Cavalieri in the plane [18] and then extended to a regular, or smooth, surface by Xin-yao [27], that in the full case

$$m\angle ASB = m\angle ASC = m\angle BSC = 120°.$$

We say a point is *terminal* in a Steiner minimal tree if it is a vertex that is not a Steiner point.

Note in the degenerate case above if we let $S = B$ we can also represent an element of $SMT(A, B, C)$ as $\overline{AS} \cup \overline{BS} \cup \overline{CS}$. We thus define the *generalized Steiner point* to be the point S such that an element of $SMT(A, B, C)$ is $\overline{AS} \cup \overline{BS} \cup \overline{CS}$.

The following results are proven in [26].

Theorem 3.1.4. *If T is a Steiner minimal tree on a smooth surface, then*

   (i) any angle in T is no less than $120°$, in particular, all angles at a Steiner point equal $120°$

(ii) any vertex in T is of degree no more than 3, in particular, any Steiner point is exactly of degree 3.

We now prove some facts about the n-point Steiner problem on such a surface. We note that Steiner points have degree 3, generalized Steiner points may have degree 2 or 3, and all other vertices in $SMT(a_1, \ldots, a_n)$ have degree 1, where degree of a vertex is the number of edges assigned to the vertex.

Proposition 3.1.5. *Let $a_1, \ldots a_n$ be points on a smooth surface and $\tau \in SMT(a_1, \ldots, a_n)$. Let*

- *s be the number of Steiner points of τ;*

- *g be the number of generalized Steiner points of degree 2 of τ; and*

- *h be the number of generalized Steiner points of degree 3 of τ that are not Steiner points.*

Then $n - 2 = s + g + 2h$.

Proof. The total number of vertices of any element in $SMT(a_1, \ldots, a_n)$ is $n+s$ and the total number of edges is $1/2(3h + 3s + 2g + n - (g + h))$ (obtained by considering the degree of each vertex and dividing by 2 since each edge will be counted twice in this procedure) [19]. By the Euler characteristic (see [24])

$$n + s - \frac{3h + 3s + 2g + n - g - h}{2} = 1.$$

Simplifying this expression yields $n - 2 = s + g + 2h$.

$\square$

Corollary 3.1.6. *Any Steiner minimal tree connecting n points on a smooth surface has at most $n - 2$ generalized Steiner points.*

Proof. The number of generalized Steiner points is $s + g + h \leq s + g + 2h = n - 2$.

$\square$

Proposition 3.1.7. *Any Steiner minimal tree connecting n-points contains a minimal sub-tree that only connects the generalized Steiner points.*

Proof. Let T be a Steiner minimal tree. By removing all vertices of degree 1 and the interior of the edges connecting them, we will obtain a subtree of T connecting the generalized Steiner points. Furthermore this subtree is a minimal tree else T would not be minimal.

$\square$

3.2   SOLVING THE STEINER PROBLEM IN THE EUCLIDEAN PLANE

For completeness we include a standard algorithm for solving the 3-point Steiner problem in the plane, based off of [18], which extends to the n-point Steiner problem. In particular, this algorithm returns the generalized Steiner point S, the Steiner minimal tree, and the length of the Steiner minimal tree for three points in the plane.

3.2.1   Algorithm.   Let A, B, and C be three points in the plane.

Case 1: An interior angle of $\triangle ABC$ is greater than $120°$. Without loss of generality suppose $m\angle ABC \geq 120°$. Then $S = B$ is the generalized Steiner point, $SMT(A, B, C) = \overline{AB} \cup \overline{BC}$, and $l(SMT(A, B, C)) = AB + BC$.

Case 2: No interior angle of $\triangle ABC$ is greater than $120°$.

(i) Construct the point E which is the point opposite C of $\overleftrightarrow{AB}$ such that $\triangle ABE$ is equilateral.

(ii) Draw a straight line from E to C.

(iii) Construct the unique circle, $\mathcal{C}$, that circumscribes $\triangle ABE$.
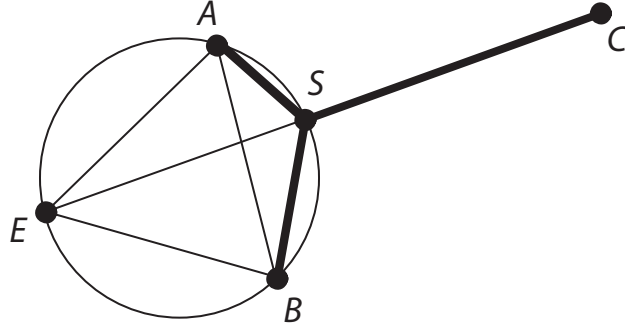
Figure 3.1: Constructing a full Steiner minimal tree in the plane.

Then S is the point distinct from E in $\mathcal{C} \cap \overline{EC}$ (in particular S is on the short arc of $\mathcal{C}$ from A to B), $SMT(A, B, C)$ is $\overline{AS} \cup \overline{BS} \cup \overline{CS}$ (see Figure 3.1) and $l(SMT(A, B, C)) = EC$.

In the $n > 3$ case there are some difficulties that arise in solving the Steiner problem. If $n > 3$ there may be many possible constructions of the Steiner tree with different topology. A topology is a *Steiner topology* if it meets the degree requirement of a Steiner minimal tree (see Figure 3.2). In order to successfully solve the n-point Steiner problem in the plane one must compare the lengths of the path networks corresponding to each Steiner topology.
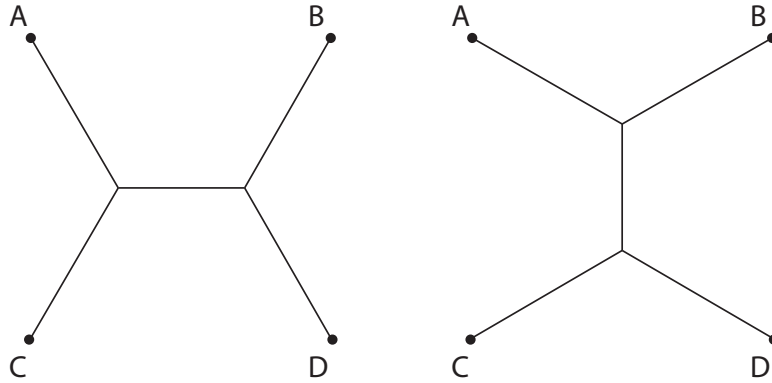

Figure 3.2: Two different Steiner topologies for $n = 4$.

Finding the length of each Steiner topology requires finding each Steiner point. To do this we look at sister terminals (pairs of terminal points that are closer in length to each other than other terminal points) and construct an E point as in the 3-point algorithm. This

E now replaces the previous two terminal points so that the n-point Steiner problem reduces to an $(n-1)$-point Steiner problem. Continuing in this manner we reduce to problem to a 2-point problem whose length represents the length of the Steiner tree under the given Steiner topology. For more results on this problem see [17, 19].

3.3 Solving the Steiner problem in the Hyperbolic plane

In the hyperbolic plane $\mathbb{H}^2$, the algorithm for solving the 3-point Steiner problem requires a different construction than in the plane. This is due to the fact that the point E that is constructed in Subsection 3.2.1 cannot be constructed in $\mathbb{H}^2$.

Halverson and March provide an algorithm for solving the Steiner problem in [12]. The solution requires the construction of a curve

$$L(p_1, p_2) := \{q : m\angle p_1 q p_2 = 120°\},$$

called a *locus*, on the upper-half plane model of $\mathbb{H}^2$ between two points p_1 and p_2.

To solve the Steiner problem connecting the three points p_1, p_2, and p_3, construct the two locus curves $L(p_1, p_2)$ and $L(p_2, p_3)$. Then $L(p_1, p_2) \cap L(p_2, p_3)$ contains the Steiner point of an element in $SMT(p_1, p_2, p_3)$ (see Figure 3.3). We refer the reader to [12] for details on the construction of the locus curves.

Halverson and March also generalize their results to the n-point Steiner problem on $\mathbb{H}^2$. For other results on length minimizing paths on the hyperbolic plane see [16].

3.4 Solving the Steiner problem on the 2-sphere

Solving the Steiner problem on the 2-sphere is similar to how it is solved in the hyperbolic plane. It includes making the locus curves and finding where they intersect. For more results on this topic see [2, 6, 8, 26].
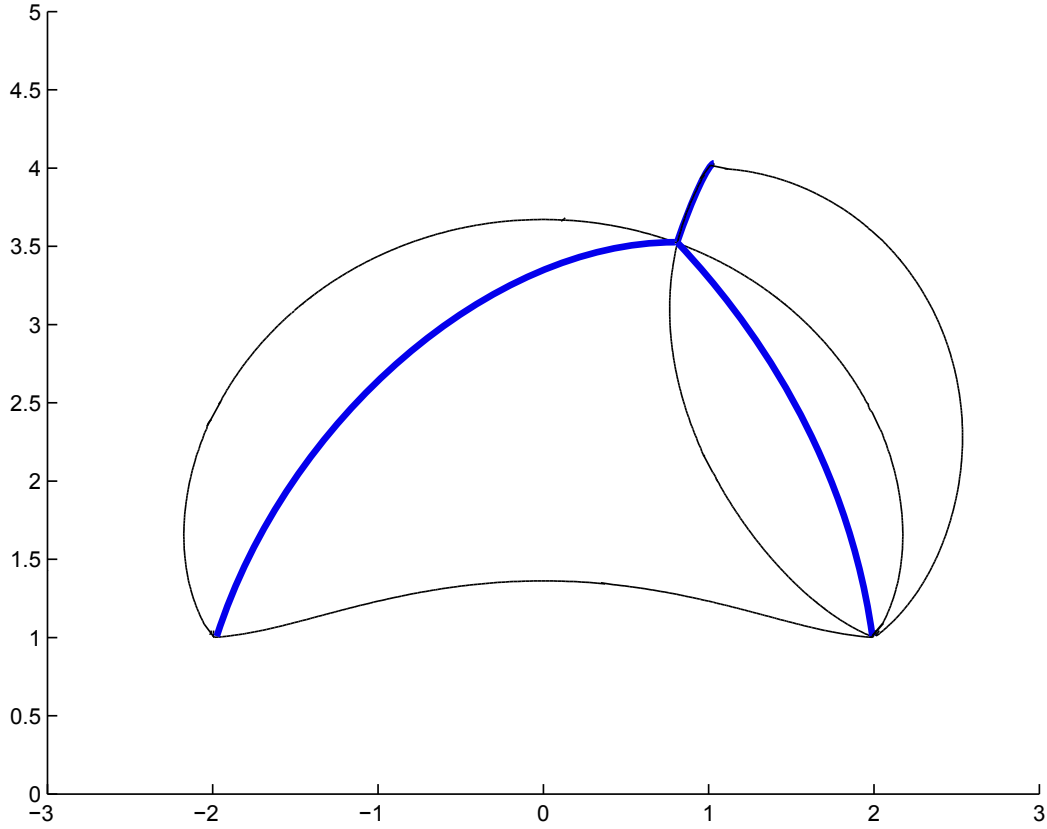
Figure 3.3: Constructing a full Steiner minimal tree in the hyperbolic plane at points $(-2, 1)$, $(2, 1)$, and $(1, 4)$ in the upper half model of the hyperbolic plane. Locus curves are in black and geodesics are in blue.

CHAPTER 4. MAIN RESULTS

In this section let X be a Riemannian two-dimensional closed surface with constant curvature. We will simply refer to X as a surface. Furthermore we will suppose $(\tilde{X}, p)$ is a locally isometric covering space of X with p chosen so that the fundamental domain is a fundamental polygon $F[q]$ with a center point q. Let w be the word that determines the identification of edges of $F[q]$ to obtain the surface X. For instance, suppose $F[q]$ is a rectangle. If $w = aba^{-1}b^{-1}$ then X is a torus; but if $w = aba^{-1}b$ then X is a Klein bottle. Thus for every edge e_i of $F[q]$ there is an edge e_j such that $l(e_i) = l(e_j)$ and e_j identifies

with e_i. Note that if $F[q]$ and $F[q']$ are adjacent in the tiling, then $F[q]$ meets $F[q']$ at an edge corresponding to the same letter of the word w with opposite orientation indicated by a superscript of -1.

It follows by Propositions 2.5.4, 2.5.5, and 2.5.6, that $\tilde{X}$ could be one of $\mathbb{R}^2$, $\mathbb{H}^2$, or S^2. In the following theorems we require reflection across a geodesic, or rather reflection across the arc formed by the geodesic, on $\tilde{X}$ to be an isometry. We know each of these three spaces have this property. It may therefore be helpful to consider $\tilde{X}$ to be one of these spaces.

Proposition 4.0.1. *Let $F[q]$ be a fundamental polygon with center q that tiles a covering space $(\tilde{X}, p)$ of the surface X. Then $qv = qw$ for all vertices v and w of $F[q]$.*

Proof. Let $e = \overline{v_1 v_2}$ be an edge in $F[q]$ with midpoint x. Then consider the triangle $\triangle v_1 v_2 q$. Note that $\triangle xqv_1$ is a right triangle and by SAS (side-angle-side) is congruent to $\triangle xqv_2$. Thus triangle $\triangle v_1 v_2 q$ is isosceles. Let $\overline{u_1 u_2}$ be an edge adjacent to e. Suppose without loss of generality that the point $v_1 = u_1$. By the argument above $qu_1 = qu_2$ and hence $\triangle qu_1 u_2$ is also isosceles. It follows that $qv_2 = qv_1 = qu_1 = qu_2$. Continuing inductively we see that $qu = qv$ for all vertices v, u in $F[q]$.

$\square$

Proposition 4.0.2. *Let $F[q]$ be a fundamental polygon with center q that tiles a covering space $(\tilde{X}, p)$ of the surface X. Let w be the word that determines the identification of edges of $F[q]$ such that e identifies with e' in w. Let x be the midpoint of e and x' be the midpoint of e'. Then $xq = x'q$.*

Proof. Let $e = \overline{v_1 v_2}$ be identified with $e' = \overline{v_3 v_4}$ in $F[q]$. It follows by Proposition 4.0.1 that the triangles $\triangle v_1 v_2 q$ and $\triangle v_3 v_4 q$ are isosceles. Since $l(e) = l(e')$, then $\triangle v_1 v_2 q$ is congruent to $\triangle v_3 v_4 q$ so that $xq = x'q$ as desired.

$\square$

The purpose of this next proposition is to show that if $F[q]$ is reflected across one of its edges, then q maps to the center point of the adjacent fundamental domain meeting at that

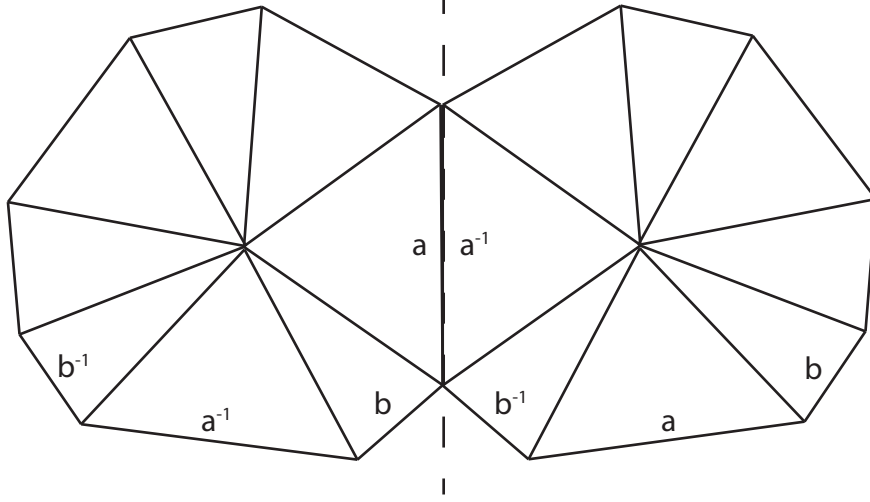edge. This is demonstrated in Figures 4.1 and 4.2.

Figure 4.1: Reflection of a polygon over one of its edges.
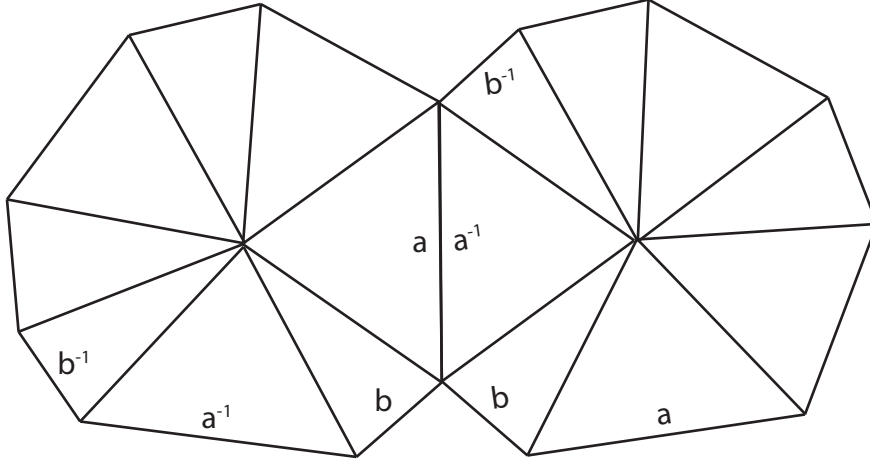
Figure 4.2: Two fundamental polygons in a polygonal tiling.

Proposition 4.0.3. *Let $(\tilde{X}, p)$ be a covering space of the surface X with fundamental polygon $F[q]$. Let $F[q']$ be an adjacent fundamental polygon to $F[q]$ with common edge e. Then the reflection of q across e is q'.*

Proof. Let x be the midpoint of e. Then $\overline{xq}$ lies on the perpendicular bisector of e (by definition of center point). Hence $\overleftrightarrow{xq}$ reflects about e to itself. Since e is also an edge of $F[q']$, it must be that q' is on the line $\overleftrightarrow{xq}$. By Proposition 4.0.2 if x' is the midpoint of the edge e' that identifies to e then $xq = x'q = xq'$. Therefore q' is the image of the reflection of

18

q across e.

$\square$

Observe that q is the unique point of $F^o[q]$ so that $p(q) = p(q')$ in X. It is this fact that we will exploit in our strategy for finding a bounded region in $\tilde{X}$ which must contain a lift of a Steiner minimal tree in X.

The following theorem is the building block of our results. The strategy employed to prove this theorem was developed while trying to solve the n-point Steiner problem on a flat torus. We then realized that this same strategy could be generalized to other surfaces as well. This became the motivation for this paper. Before we introduce and prove this theorem we will need to provide a definition.

Definition 4.0.4. A *radial segment* of a polygon P with center q is any segment $\overline{vq}$ where v is a point of the boundary of P.

Theorem 4.0.5. *Let $\overline{ab}$ be a shortest geodesic segment in X and $\overline{AB}$ a lift of $\overline{ab}$ in $\tilde{X}$. Then $\overline{AB}$ is contained in a radial segment of the fundamental polygon $F[A]$.*

Proof. Suppose that $\overline{AB}$ is not contained in a radial segment of the fundamental polygon $F[A]$ with center A. It suffices to show that $AB > \min\{A'B : A' \in p^{-1}(a)\}$. Since $\overline{AB}$ is not contained in a radial segment of $\overline{F}[A]$, then $\overline{AB}$ meets a side of $F[A]$ at a point Q. Let L be the arc determined by the side of $\overline{F}[A]$ containing Q. By reflecting the part of $\overline{AB}$ opposite B over the line L, we obtain a union of edges $\overline{A'Q} \cup \overline{QB}$ where $A' \in p^{-1}(a)$ (by Proposition 4.0.3) and $A'Q + QB = AB$. By the triangle inequality $A'B \leq A'Q + QB = AB$. Since A' is on the circle of radius AQ centered at Q with $A' \neq A$, then $m\angle A'QB < 180°$ so that $A'B < A'Q + QB = AB$. Hence $AB > \min\{A'B : A' \in p^{-1}(a)\}$.

$\square$

Corollary 4.0.6. *The lift of a minimal path network connecting n points on X having e edges is contained in a collection of e radial segments. Moreover, the radial segments can be chosen to be based at either endpoint of a lifted edge.*

19

Corollary 4.0.7. *Any minimal path network connecting three points on X has a lift that is contained in the closure of the fundamental polygon centered at the point in the preimage of the generalized Steiner point.*

Proposition 4.0.8. *The maximum number of edges on a n-point Steiner minimal tree is $2n - 3$ for $n \geq 2$.*

Proof. The maximum number of Steiner points is known to be $n - 2$ [18, 12]. Hence the Steiner minimal tree connecting the n points will have at most $2n - 2$ vertices. As a standard result from graph theory the number of edges connecting s vertices on a tree is $s - 1$. Therefore the maximum number of edges on a n-point Steiner minimal tree in the plane is $2n - 3$.

$\square$

These results combine to give the following general result.

Theorem 4.0.9. *Any minimal path network τ connecting n points on X has a lift that is*

  (i) *contained in the connected union of e radial segments where $e \leq 2n - 3$ is the number of edges of τ; and*

  (ii) *contained in the connected union of k closed fundamental polygons where $k \leq n - 2$ is the number of generalized Steiner points of τ.*

Moreover, the radial segments in (1) can be chosen to be based at generalized Steiner points and the fundamental polygons in (2) can be chosen to be centered at the generalized Steiner points.

4.1 ALGORITHM

Let $a_1, \ldots, a_n$ be points on a closed surface X of constant curvature with locally isometric covering space $(\tilde{X}, p)$ and p chosen so that $F[q]$ is a fundamental polygon. Consider the simplicial complex K that is the collection of m isosceles triangles (see Proposition 4.0.1) each

one having vertex q and edge e where e is an edge of $F[q]$ together with their faces. We can now form a tiling of $\tilde{X}$ with isosceles triangles which determines a simplicial complex $\tilde{K}$ with underlying space $\tilde{X}$. We consider $\mathrm{St}(q)$ (which in fact is $F^o[q]$) in $\tilde{K}$ (see Definition 2.3.1).

(i) Choose an element $A \in p^{-1}(a_1)$.

(ii) Let

$$\mathcal{A}_i = p^{-1}(a_i) \cap \overline{\mathrm{St}}^{2n-3}(A).$$

(iii) Calculate the lengths of the possible Steiner minimal trees connecting to exactly one point in each of the sets $\mathcal{A}_i$.

(iv) Identify the trees that have minimal length. Their projections to X are the elements of $SMT(a_1, \ldots, a_n)$.

To reduce the number of calculations in Step (iii) substantially, make sure the Steiner minimal tree in question meets the criteria of Theorem 4.0.9.

CHAPTER 5. EXAMPLES

The main results that have been proven so far apply to a subcollection of all closed surfaces of constant curvature. In this chapter we will explore specific spaces for which these results apply. Let Y be a closed surface of constant curvature such that Y has a locally isometric covering space $(\tilde{Y}, p)$ that is tiled by a fundamental polygon with a center point. From previous results $\tilde{Y}$ is one of $\mathbb{R}^2$, $\mathbb{H}^2$, or S^2. Now suppose X is any surface that has an isometry $y : X \to Y$. Then $(\tilde{Y}, py^{-1})$ is a locally isometric covering space for X. In the following examples we will provide solutions to the n-point Steiner problem on some surface Y (such as S^2, $\mathbb{R}P^2$, $\mathcal{T}^2(e,f)$, $\mathcal{K}^2(e,f)$. or $M_g(k)$).

5.1 Surfaces of constant positive curvature

Proposition 2.5.4 states that for any closed surface X of constant positive curvature there exists an isometry from X to the 2-sphere or the projective plane. The algorithm provided in this thesis is trivial for the 2-sphere and trivial for the projective plane if $n > 3$. This is because the real projective plane has a double cover by the 2-sphere (see Proposition 2.5.8).

5.2 Surfaces of constant zero curvature

The results in this paper clearly apply to a flat torus $\mathcal{T}^2(e, f)$ with e perpendicular to f (a complete algorithm solving the 3-point Steiner problem on such a torus was done by Halverson and Logan in [11]) and the flat Klein bottle $\mathcal{K}^2(e, f)$ with e perpendicular to f. This is because there exist locally isometric covering maps from the plane to the flat torus or the Klein bottle with fundamental polygons (rectangles) that have center points. Ivanov and Tuzhilin prove in [18, pages 321-323, 353-355] that f can be chosen to be perpendicular to e on $\mathcal{K}^2(e, f)$, but not necessarily on $\mathcal{T}^2(e, f)$. Hence the results in this thesis apply to any closed surface of constant zero curvature except flat tori where e and f are not perpendicular.

5.3 Surfaces of constant negative curvature

There exist locally isometric covering maps from any g-holed torus of constant negative curvature k, $M_g(k)$, to the hyperbolic space. Any $M_g(k)$ is formed by the word

$$w = a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_g b_g a_g^{-1} b_g^{-1}.$$

Replacing any countable collection of $a_i b_i a_i^{-1} b_i^{-1}$, $1 \leq i \leq g$, with $a_i b_i a_i^{-1} b_i$ in w, yields non-orientable closed surfaces of constant negative curvature k. We will denote the set of such surfaces by $N_g(k)$. Some of these surfaces have fundamental domains that have the properties needed to apply the results of this paper. For instance any $M_g(k)$ or any element

of $N_g(k)$ with a fundamental domain that is a regular polygon will definitely have a center point. In these and other cases the algorithm above applies.

5.4   OTHER SURFACES

The results in this thesis may also be applied, with some modifications, to other surfaces of constant curvature that are non-simply connected and not closed such as the cylinder. Halverson and Logan provide an algorithm for solving the 3-point Steiner problem on the cylinder [10]. Their exists a locally isometric covering map from the plane to the cylinder of constant zero curvature. The fundamental domain of this covering space is an infinite *strip* $\mathcal{S} := \{(x, y) \in \mathbb{R}^2 : a \leq x < b\}$. There is not a center point to this fundamental domain, but there does exist a center line, which requires a definition.

Definition 5.4.1. Let F be an infinite strip that forms a fundamental domain of the cylinder. The *center line* of F is the line equidistant to the sides of F.

We further define a *half strip* of the fundamental domain F to be the region between the center line and one of the sides of the strip. With slight modifications to our arguments we can see that the minimal path network connecting n points on a cylinder is contained in the connected union of $e \leq 2n - 3$ rectangular strips and $s \leq n - 2$ fundamental domains where e is the number of edges of the path network and s is the number of generalized Steiner points.

CHAPTER 6. CONCLUSION

If a closed surface of constant curvature X is isometric to a closed surface of constant curvature that has a covering space with a fundamental domain that has a center point, then Section 4.1 provides an algorithm to solve the n-point Steiner problem on X. Clearly this is a subcollection of all closed surfaces of constant curvature. We wish to provide an

algorithm for solving the Steiner problem on all closed surfaces of constant curvature. To do this the following research must be done:

- Provide a classification theorem that determines the existence of isometries of non-orientable closed surfaces of constant negative curvature to an element of $N_g(k)$, defined above, or similar surface.

- Solve the Steiner problem on the flat torus $\mathcal{T}^2(e, f)$ where e and f are not perpendicular as discussed in Section 5.2.

- Solve the Steiner problem on closed surfaces of constant negative curvature that do not have a covering map that provides a fundamental domain with a center point.

The results in this thesis might also extend to other surfaces with a "wrapping" issue as was illustrated in the case of the cylinder. Exploring these extensions may bring fruitful progress in solving the Steiner problem on all surfaces.

Bibliography

[1] A. Baragar, *A survey of classical and modern geometries with computer activities*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, (2001).

[2] M. Brazil, J. H. Rubinstein, D. A. Thomas, J. F. Weng, N. C. Wormald, *Shortest networks on spheres*, DIMACS Series in Discrete Math. and Theoretical Computer Science, **40** (1998) 453–461.

[3] T. Brune, L. Sipe, *Shortest path between two points on the regular tetrahedron*, (2009), available at http://tinyurl.com/BruneSipe2009.lllllllll;

[4] E. Caffarelli, D. M. Halverson, R. J. Jensen, *The Steiner problem on surfaces of revolution*, Graphs Combin. **30**:2 (2014) 315-342.

[5] D. Cieslik, *Steiner Minimal Trees*, Kluwer Academic Publishers, Dordrecht, (1998).

[6] E. J. Cockayne, *On Fermat's problem on the surface of a sphere*, Math. Mag., **45** (1972) 216–219.

[7] R. Courant, H. Robbins, *What is mathematics? An elementary approach to ideas and methods*, Oxford University Press, New York, (1979).

[8] J. Dolan, F. Weiss, J. MacGregor Smith, *Minimal length tree networks on the unit sphere*, Ann. Oper. Fes., **33** (1991) 503–535.

[9] È. N. Gordeev, O. G. Tarastsov, *The Steiner problem. A survey.* Discrete Math. Appl. **3**:4 (1993) 339-364.

[10] D. M. Halverson, A. E. Logan, *The three-point Steiner problem on a cylinder*, Involve **6**:2 (2013) 251-260.

[11] D. M. Halverson, A. E. Logan, *The Steiner problem on a flat torus*, pre-print.

[12] D. M. Halverson, D. March, *Steiner tree constructions in hyperbolic space*, (2005), available at http://tinyurl.com/HalversonMarch2005.

[13] D. M. Halverson, K. Moon, G. Shero, *The Steiner problem on the regular tetrahedron*, Involve, **4**:4 (2011) 365–404.

[14] D. M. Halverson, K. Penrod, *The three point Steiner problem on the flat torus*, (2007) available at http://tinyurl.com/HalversonPenrod2007.

[15] G. A. Hedlund, *Geodesics on a Two-Dimensional Riemannian Manifold With Periodic Coefficients*, The Annals of Mathematics, 2nd Ser. **33**:4 (1932) 719–739.

[16] H. Helmandollar, K. Penrod, *Length minimizing paths in the hyperbolic plane: proof via paired subcalibrations*, Illinois J. Math. **51**:3 (2007), 723–729.

[17] F. K. Hwang, D. S. Richard, P. Winter, *The Steiner Tree Problem*, North-Holland, (1992).

[18] A. O. Ivanov, A. A. Tuzhilin, *Minimal Networks: the Steiner Problem and Its Generalizations*, CRC Press, Inc., (1994).

[19] V. Jarník, M. Kössler, *O minimálních grafech, obsahujících n daných bodů*, Časopis pro Pěstování Matematiky a Fysiky, (1934) 223–235.

[20] A. Lee, J. Lytle, D. Halverson, D. Sampson, *The Steiner problem on narrow and wide cones*, (2011), available at http://tinyurl.com/LeeLytleHalversonSampson2011.

[21] W. S. Massey, *Algebraic Topology: An Introduction*, Reprint of the 1967 edition, Graduate Texts in Mathematics, Vol. 56. Springer-Verlag, New York-Heidelberg, (1977). ISBN 0-387-90271-6.

[22] K. L. May, M. A. Mitchell, *The three point Steiner problem on the flat torus: the minimal lune case*, (2007), available at http://tinyurl.com/MayMitchell2007.

[23] J. R. Munkres, *Elements of algebraic topology,*Addison-Wesley Publishing Company, Menlo Park, CA, (1984), ISBN 0-201-04586-9.

[24] R. Muss, *Graph Theory*, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, (2001). ISBN: 0-471-38925-0

[25] I. M. Roussos, *On the Steiner minimizing point and the corresponding algebraic system*, The College Mathematics Journal, **43**:4 (2012) 305–308.

[26] J. F. Weng, *Steiner trees on curved surfaces*, Graphs Combin., **17**:2 (2001) 353–363.

[27] X. Jiang, *The Steiner problem on a surface*, Applied Mathematics and Mechanics, **8**:10 (1987) 969–974.