2017

Shortest Connection Networks on Triangular Grids

Jie Mei
University of Rhode Island, jie_mei@uri.edu

SHORTEST CONNECTION NETWORKS ON TRIANGULAR GRIDS

BY

JIE MEI

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

COMPUTER SCIENCE

UNIVERSITY OF RHODE ISLAND

2017

MASTER OF SCIENCE THESIS

OF

JIE MEI

APPROVED:

Thesis Committee:

Major Professor Edmund Lamagna

Noah Daniels

William Kinnersley

Nasser Zawia

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2017

ABSTRACT

The shortest path problem, or the Steiner problem, is an interesting problem with numerous real-world applications. Historically the Steiner problem has been studied for the Euclidean plane and for rectilinear distances. Both problems have been proven to be NP-hard. In this research, we look into the Steiner problem on a triangular grid and show that the problem is NP-hard. We explore exact algorithms for constructing a shortest network that optimally interconnects a set of terminal points on a grid. Moreover, we look at a heuristic algorithm to solve the problem and provide a conjecture on the bound of the approximation it produces.

ACKNOWLEDGMENTS

First of all I would like to thank my research advisor, Professor Edmund Lamagna, who proposed this interesting topic to me and who was very patient in discussing and reviewing my work. Besides his guidance on my research, he provided me the opportunity to present my work at the MAA Mathfest conference in 2017. I feel happy to have Professor Lamagna as my advisor.

I would like to thank Professor Hervé for taking a nice picture of the game board of TransAmerica.

I would like to thank my thesis committee members, Professor Noah Daniels, Professor William Kinnersley and Professor Michael Barrus, for their time and effort in reviewing this thesis.

Finally, I want to thank my family for their love and support and I thank myself for producing this nice work.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure		Page

LIST OF TABLES

CHAPTER 1

Introduction

The shortest connection network problem studied here originates from the board game TransAmerica (Figure 1). In the game, each player attempts to connect five cities in different regions of the country with as few train tracks as possible. The map where the players build their railroads is a triangular grid. Since each player can place only two tracks per round, it is advantageous to know the shortest path connecting one's five cities to increase the chances of winning.

Figure 1: The game board for TransAmerica

We describe the shortest connection problem as follows. Given a set of points on a triangular grid, find a path that connects these points with the smallest total length. The points to be connected are referred to as *terminal points*. The paths lie along the grid lines and additional junction points may be introduced into the network. These junction points are called *Steiner points* and the resulting connected network is called a *Steiner tree*. A minimum Steiner tree (SMT) has the

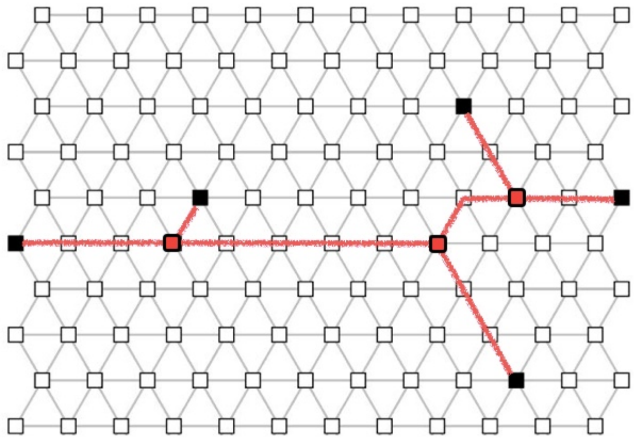


Figure 2: A Steiner minimum tree for five terminal points (black). The red points are Steiner points.

shortest length among all Steiner trees (Figure 2). A vertex in a Steiner tree can be either a terminal point or a Steiner point.

Steiner trees and Steiner points are named after Jakob Steiner, a 19th-century Swiss mathematician who studied the problem of using a single junction point to optimally interconnect a set of terminal points [1]. In the 17th century the minimum path problem for three terminals was studied by mathematicians including Fermat, Cavalieri and Torricelli. The Steiner point for the three-point case is also called the Fermat point or Torricelli point. All angles at this junction are 120°. When Richard Courant and Herbert Robins included Steiner's problem in their book, *What Is Mathematics* [2], they formalized the general form of the Steiner tree problem as finding a shortest possible network for a set of points with extra vertices as junctions. The Steiner tree problem has become popular since then, and it has been studied for different geometries, metrics and cost functions.

On the Euclidean plane, Melzak (1961) [3] first proposed an algorithm to solve the Euclidean Steiner problem. The algorithm constructs Steiner trees for all possible topologies and selects the shortest one to be the SMT. Melzak's algorithm is a brute-force approach and takes exponential time in that the number of topologies is large. In fact the Euclidean Steiner problem was proven to be NP-hard by Garey et al. (1977) [4] and Rubinstein et al. (1997) [5].

Another geometry where the Steiner problem has been widely studied is the rectilinear grid. This version is important because of its application to integrated circuit routing design. The problem is also known as the Manhattan distance Steiner problem. In 1966 Hanan [6] showed that the Steiner points can be chosen from a predetermined set of points, and a rectilinear SMT can therefore be computed by an exhaustive search. The rectilinear Steiner problem (RST) was shown to be NP-hard by Garey and Johnson in 1977 [7]. They demonstrated that the RST problem can be reduced from the vertex cover problem, which is known to be NP-complete.

To the best of our knowledge, there has been far less research on the Steiner tree problem on a triangular grid or a hexagonal grid, a similar geometry. There are interesting differences between locating Steiner points on a triangular grid and the other geometries mentioned. For example, when connecting three terminal vertices on a triangular grid, there may be multiple choices for the Steiner point, whereas there is only one choice for the Euclidean and rectangular problems. In this research, we provide a comprehensive study of the triangular Steiner problem's computational complexity and several algorithms to solve it.

The main body of the thesis is divided as follows. Chapter 2 introduces the metric system and looks into the simplest non-trivial case of the Steiner problem on a triangular grid. Chapter 3 focuses on constructing exact solutions to the Steiner problem. Chapter 4 studies the computational complexity and gives a proof that the triangular Steiner problem is NP-hard. Chapter 5 discusses an approximation solution to the problem and bounds on the goodness of the approximation. Chapter 6 provides a summary of this research.

CHAPTER 2

The Coordinate System

2.1 The triangular grid coordinate system

A coordinate system based on triangular grid lines is used as the reference system in this research.

Let O be the origin. The axes, x_1, x_2, x_3, are the three directions along the triangular grid lines at O. x_1, x_2, x_3 are 120° apart. Each axis separates the plane into a positive half and a negative half. A counterclockwise rotation is defined to be the positive direction. The half plane initially scanned by rotating an axis counterclockwise is the positive half; the other half is the negative half plane (Figure 3).

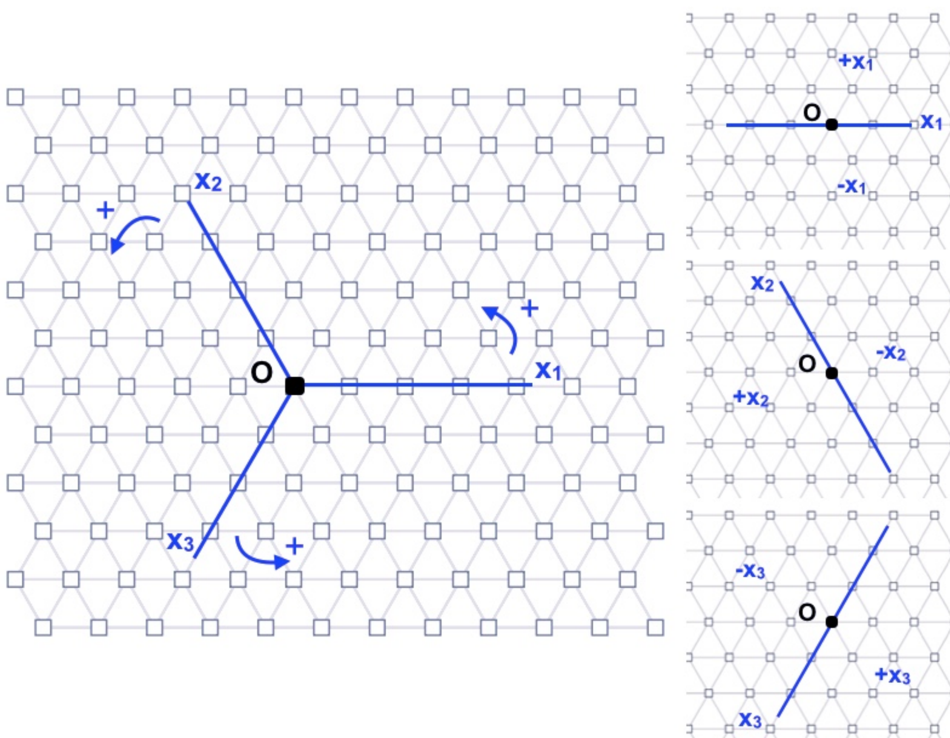


Figure 3: The triangular coordinate system and half planes.

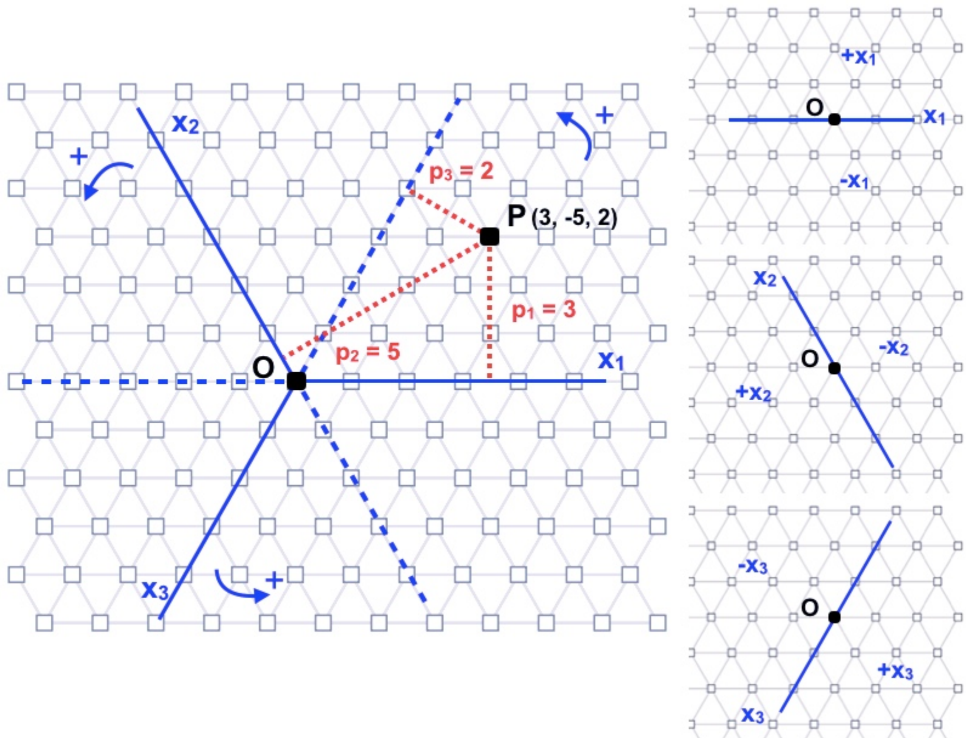


Figure 4: The coordinates of a point.

Parallel grid lines are one unit length apart in all three directions. For any point P, let p_1, p_2, p_3 be the distances to the axes. Depending on whether P sits on the positive or the negative half plane of x_1, its x_1 coordinate is either $+p_1$ or $-p_1$. The same applies to the other two coordinates of P. For example, in Figure 4, the coordinates of P are $(+p_1, -p_2, +p_3) = (+3, -5, +2)$. Note that the coordinates of the points and distances discussed throughout this research are constrained to be integers. Though only discretized problems are considered here, almost all the definitions and theorems also apply to continuous coordinates due to the linearity of the metric system.

We next list some of the basic properties of the triangular coordinate system, providing proofs for those that are not obvious.

Lemma 2.1. The three coordinates of any point add to 0 (constant).

Proof. Consider a point $P(p_1, p_2, p_3)$ in the triangular coordinate system. Let the polar coordinates of P be (r, θ).

Then

$$p_1 = r\sin\theta,\ p_2 = r\sin(\theta - \frac{2\pi}{3}),\ p_3 = r\sin(\theta - \frac{4\pi}{3})$$
$$p_1 + p_2 + p_3 = r\sin\theta + r\sin(\theta - \frac{2\pi}{3}) + r\sin(\theta - \frac{4\pi}{3}) = 0$$

□

The distance between two points

Definition (Distance). Consider two points, $A(a_1, a_2, a_3)$ and $B(b_1, b_2, b_3)$. The distance from A to B is the fewest number of steps taken walking from A to B along grid lines.

The axis lines passing through A and B enclose three parallelograms having A and B as diagonally opposite vertices. In one of these parallelograms, the angles at both A and B are 60°. This is the *distance parallelogram* of A and B. The distance between A and B is the sum of the lengths of any two adjacent edges of their distance parallelogram (Figure 5).

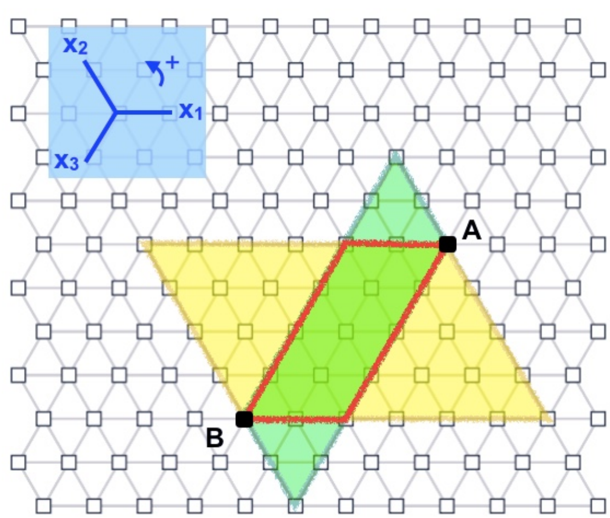


Figure 5: Three parallelograms enclosed by the axes lines passing through A and B. The one in the red frame is the distance parallelogram of A and B.

Lemma 2.2. Let $d_1 = |a_1 - b_1|$, $d_2 = |a_2 - b_2|$ and $d_3 = |a_3 - b_3|$. The *distance* between A and B is $\min\{d_1 + d_2,\ d_1 + d_3,\ d_2 + d_3\}$. It is also $\max\{d_1,\ d_2,\ d_3\}$, and equal to $\frac{1}{2}(d_1 + d_2 + d_3)$.

Proof. Without loss of generality, suppose $d_3 = \max\{d_1, d_2, d_3\}$ and $b_3 > a_3$. Because $a_1 + a_2 + a_3 = b_1 + b_2 + b_3$, $(a_1 - b_1) + (a_2 - b_2) = b_3 - a_3 = d_3$, we must have $a_1 - b_1 \geqslant 0$ (otherwise $d_2 \geqslant a_2 - b_2 = d_3 + (b_1 - a_1) > d_3$, which contradicts to our assumption). Similarly $a_2 - b_2 \geqslant 0$. Thus we have $d_1 + d_2 = d_3$ and $d_1 + d_2 + d_3 = 2d_3$. Therefore, $\min\{d_1 + d_2,\ d_1 + d_3,\ d_2 + d_3\} = d_1 + d_2 = d_3 = \frac{1}{2}(d_1 + d_2 + d_3)$ is the smallest sum of any two d_i.

□

Lemma 2.3. All walks from A to B within the distance parallelogram (DP), with directions restricted to being parallel to two edges of the DP, moving only in the "forward" directions (never doubling-back in any coordinate), yield the same total length.

This is an obvious result from Figure 5.

Circle and sectors

Definition (Circle). Given a point P and length l, the set of all points at distance l from P forms a regular hexagon. We call these points the *circle* with center P and radius l on the triangular grid (Figure 6).

Consider a point $P(p_1, p_2, p_3)$. Draw lines through P that are parallel to the three axis lines. These lines divide the plane to three *sectors* around P. Call sector I the infinite region between directions x_2 and x_3, sector II the region between x_1 and x_3, and sector III the region between x_1 and x_2 (Figure 7a).

Let us revisit the distance between two points using the notion of sectors. Consider another point $Q(q_1, q_2, q_3)$ on the grid. If Q is in sector I of P, the

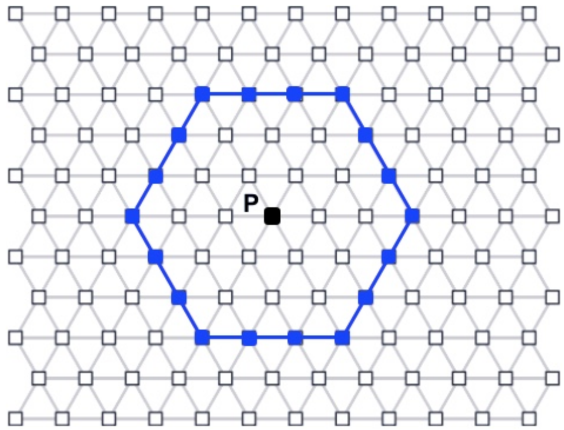


Figure 6: A circle on a triangular grid. P is the center; the radius is 3.

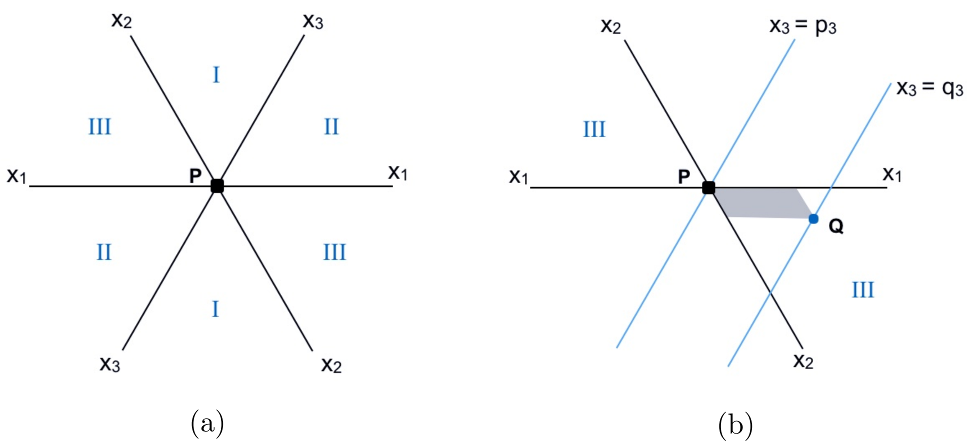


Figure 7: (a) Sectors of a vertex P on a triangular grid. (b) Distance between two points based on sectors. Q is in sector III of P. $|PQ|$ is the distance between lines $x_3 = p_3$ and $x_3 = q_3$. The gray shaded parallelogram is the distance parallelogram of P and Q.

distance $|PQ|$ is the difference of their first coordinates, i.e., $|PQ| = |p_1 - q_1|$. A similar result holds for points in sectors II and III. To understand this definition visually on the grid, depending on which sector of P the point Q is located, draw a line parallel to that axis through Q. $|PQ|$ is the distance between these parallel lines (Figure 7b).

2.2 Properties

Three terminal vertices are the simplest nontrivial case for the Steiner problem on triangular grids. In the previous section we described how to construct the triangular coordinate system and provided some basic concepts and definitions. Based on these foundations, we discuss in this section properties related to three points on triangular grids as well as properties of triangular Steiner trees.

2.2.1 Equilateral triangles aligned to the grid

Consider three vertices that form an equilateral triangle with edges along grid lines. Let the vertices be $A(0,0,0)$, $B(0,-a,a)$, and $C(a,-a,0)$. a is the side length of the triangle (Figure 8).

Lemma 2.4. For any point in or on the equilateral triangle, the sum of the distances to each of the three edges adds to a.

Proof. Let $P(p_1,p_2,p_3)$ be a point in or on $\triangle ABC$. The line equations of the three sides are $x_1 = 0$, $x_2 = -a$ and $x_3 = 0$. This implies that $0 \leqslant p_1 \leqslant a$, $-a \leqslant p_2 \leqslant 0$, $0 \leqslant p_3 \leqslant a$. The distance to edge AB ($x_1 = 0$) is p_1; the distance to edge BC ($x_2 = -a$) is $p_2 - (-a) = p_2 + a$; the distance to edge AC ($x_3 = 0$) is p_3. So the sum of the distances is $p_1 + p_2 + a + p_3 = a$ (Figure 8). $\square$

Lemma 2.5. The sum of the distances from any point P in or on the equilateral triangle to the three vertices is $2a$ (constant).

Proof.

$$|PA| = \frac{1}{2}(|p_1| + |p_2| + |p_3|) = \frac{1}{2}(p_1 - p_2 + p_3)$$

$$|PB| = \frac{1}{2}(|p_1| + |p_2 + a| + |p_3 - a|) = \frac{1}{2}(p_1 + (p_2 + a) + (a - p_3))$$

$$|PC| = \frac{1}{2}(|p_1 - a| + |p_2 + a| + |p_3|) = \frac{1}{2}((a - p_1) + (p_2 + a) + p_3)$$

$$|PA| + |PB| + |PC| = 2a$$

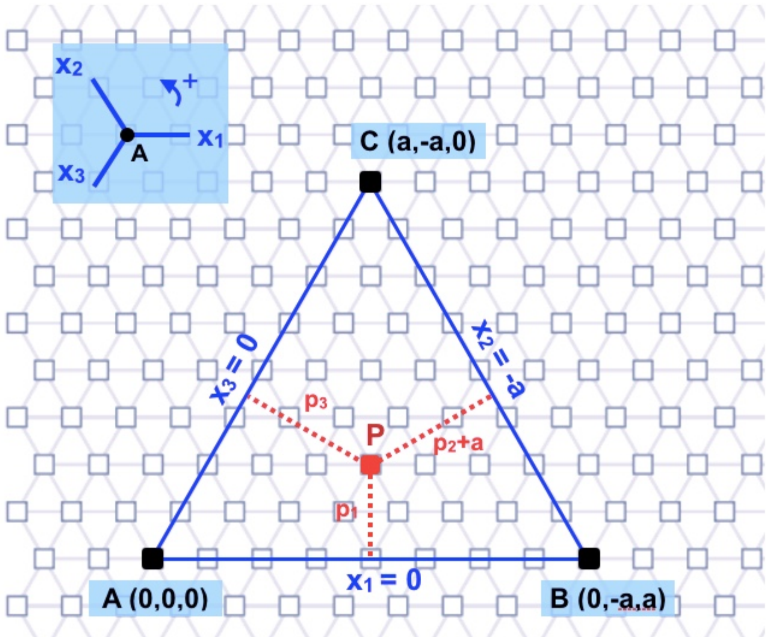


Figure 8: Equilateral triangle aligned with grids. P is an internal point.

□

Lemmas 2.4 and 2.5 are useful in finding solutions to the triangular Steiner problem for three terminal vertices. A detailed characterization is provided in Theorem 2.1.

2.2.2 Steiner trees on triangular grids

The optimal solution to finding a shortest connection network for a set of terminal points is a Steiner tree. Additional junction points that might be introduced in a Steiner tree are called Steiner points. A terminal point serving as a junction point (degree of two or more) is not, in general, considered a Steiner point. In this situation, the terminal point is considered to coincide with one or more Steiner points and it can be regarded as a terminal point of degree 1, Steiner point(s), and some zero-length edge(s). In this section, we provide some important properties of Steiner trees and Steiner minimum trees (SMT) on triangular grids.

Lemma 2.6. For n terminal vertices on a triangular grid, there are at most $n-2$ Steiner points in the Steiner minimum tree.

Proof. Let s be the total number of Steiner points and $d_1, d_2, \cdots, d_s$ be the degrees for every Steiner point in the Steiner tree.

The degree of any Steiner point must be greater than two, i.e., $d_i \geqslant 3$ for any $d_i \in \{d_1, d_2, \cdots, d_s\}$. Otherwise, if a Steiner point had degree 2, it could simply be eliminated. So there is no need to consider cases where a Steiner point connects only two vertices.

Let e be the number of edges in the SMT. Because the sum of the degrees of all vertices in any graph is twice the number of edges and that the sum of the degrees of all terminal points is larger than the number of terminals,

$$2e \geqslant d_1 + d_2 + \cdots + d_s + n = \sum_i^s d_i + n,$$

and for a tree

$$e = s + n - 1$$

$$\Longrightarrow \qquad 2e = 2(s+n-1) \geqslant \sum_i^s d_i + n \geqslant 3s + n$$

Thus

$$n - 2 \geqslant s.$$

□

Lemma 2.6 describes a common property shared by the Euclidean SMT and the rectilinear SMT. A Steiner tree that has $n-2$ Steiner points is called a *full Steiner tree*. In a full Steiner tree, the terminals and Steiner points as well as the Steiner points themselves are distinguished from each other; and a terminal point has degree 1.

Lemma 2.7. The degree of any Steiner point in a Steiner minimum tree is at most 4.

Proof. We prove this by contradiction.

Suppose a Steiner point S in some Steiner minimum tree T of terminal set $\mathcal{X}$ has degree 6. Let A, B, C, D, E, F be all immediate neighbors of S. Note that $A \cdots F$ are not necessarily vertices of T. Remove all edges from S to $A \cdots F$ and connect AB, BC, CD, DE, EF to transform T to a new tree T'. The connectivity of $\mathcal{X}$ does not change, however, the length of T' is shorter than T. Thus the original tree is not an SMT (Figure 9a).

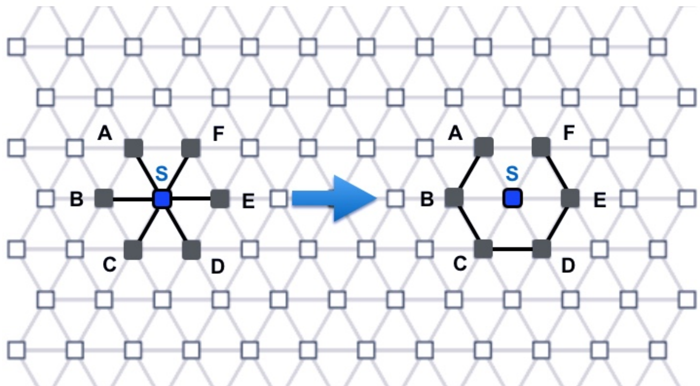
Next suppose a Steiner point S in some SMT T of terminal set $\mathcal{X}$ has degree 5. Let A, B, C, D, E be the immediate neighbors of S. Transform T to T' as above and we have a new Steiner tree that is shorter than the original one (Figure 9b). $\square$

A Steiner point is either of degree 3 or 4 in a Steiner tree. We can comfortably regard a Steiner point of degree 4 as two Steiner points of degree 3 joining by an imaginary zero-length edge.

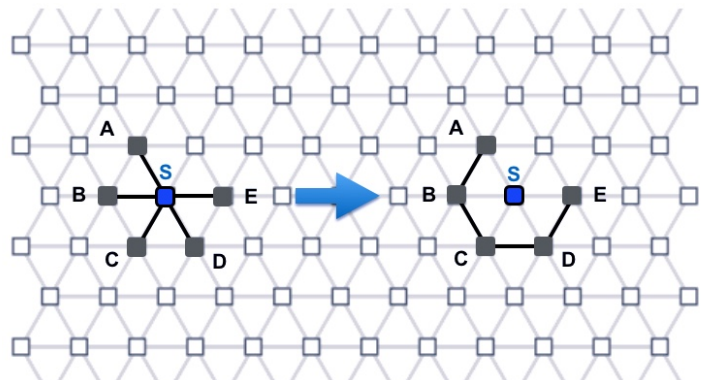
Definition (**Median triangle**)**.** Consider three terminal vertices $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$ and $C(c_1, c_2, c_3)$ on a triangular grid. Let the median values in each direction be $m_1 = \text{median}\{a_1, b_1, c_1\}$, $m_2 = \text{median}\{a_2, b_2, c_2\}$, $m_3 = \text{median}\{a_3, b_3, c_3\}$. The region enclosed by lines $x_1 = m_1$, $x_2 = m_2$, and $x_3 = m_3$ is an equilateral triangle. This follows naturally because the lines are $120°$ apart. This equilateral triangle is the *median triangle* for the terminal set $\{A, B, C\}$. Lines $x_1 = m_1$, $x_2 = m_2$ and $x_3 = m_3$ are the *median lines* in each direction (Figure 10).

In some cases, the median triangle may degenerate to a single point. These cases will be considered later in Lemma 2.8.

Based on the definition of the median triangle for three terminal vertices, we are able to provide a theorem that solves the Steiner problem for three terminal



(a)



(b)

Figure 9: (a) Tree transformation for a Steiner point of degree 6. (b) Tree transformation for a Steiner point of degree 5.

vertices.

Theorem 2.1. To optimally interconnect three terminal points on a triangular grid, there is one Steiner point and the possible location of this Steiner point may not be unique.

1. If the terminal points are collinear (all lie on the same grid line), there is only one possible choice for the Steiner point and it is the terminal point that lies between the other two (Figure 11a).

2. If the terminal points are noncollinear, every point in or on the median triangle is a possible choice for the Steiner point and all yield Steiner minimum

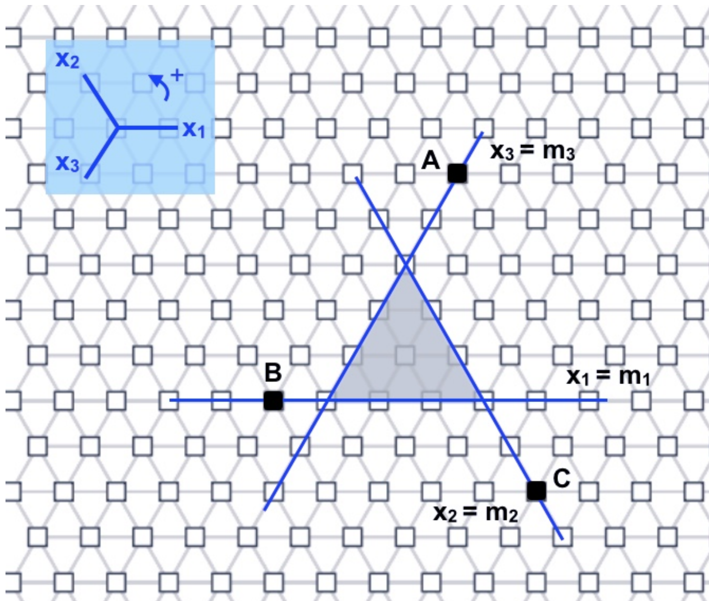
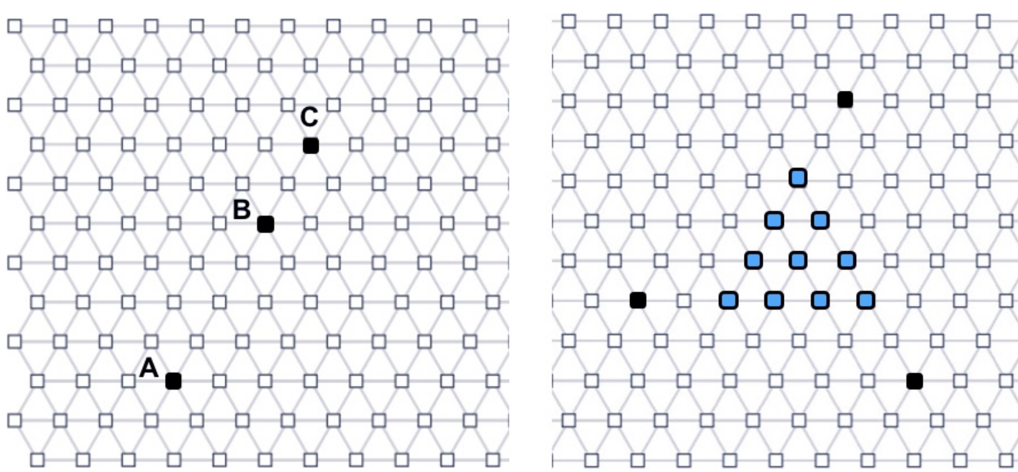


Figure 10: The gray shaded equilateral triangle is the median triangle for terminal set $\{A, B, C\}$. The median lines are $x_1 = m_1 = b_1$, $x_2 = m_2 = c_2$ and $x_3 = m_3 = a_3$.

trees of the same shortest length. The region for the Steiner points is called the *Steiner region* (Figure 11b).



(a) (b)

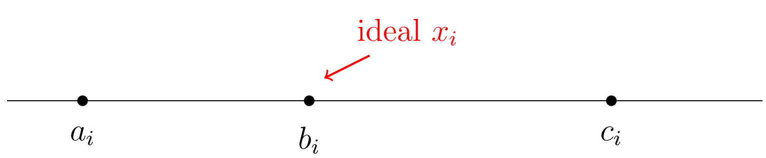
Figure 11: (a) Three collinear terminal points. Terminal B serves as the Steiner point. (b) Three noncollinear terminal points (black). Each blue point can serve as a Steiner point and the median triangle forms the Steiner region.

Proof. Let $X(x_1, x_2, x_3)$ be an arbitrary point on the triangular grid. The total distance from X to the three terminal vertices is

$$\begin{aligned} S &= |XA| + |XB| + |XC| \\ &= \frac{1}{2}(|x_1 - a_1| + |x_2 - a_2| + |x_3 - a_3|) + \frac{1}{2}(|x_1 - b_1| + |x_2 - b_2| + |x_3 - b_3|) \\ &\quad + \frac{1}{2}(|x_1 - c_1| + |x_2 - c_2| + |x_3 - c_3|) \end{aligned}$$

$$\begin{aligned} 2S &= \sum_{a,b,c} \sum_{i=1}^{3} |x_i - a_i| \\ &= \sum_{i=1}^{3} (|x_i - a_i| + |x_i - b_i| + |x_i - c_i|) \end{aligned}$$

Let $m_i = \text{median}\{a_i, b_i, c_i\}$, $d_i = \max\{a_i, b_i, c_i\} - \min\{a_i, b_i, c_i\}, i \in \{1, 2, 3\}$. On triangular grids, the median lines $x_1 = m_1$, $x_2 = m_2$ and $x_3 = m_3$ enclose the median triangle. In order to minimize S, in each direction, we prefer to choose x_i as the median among a_i, b_i, c_i, $i \in \{1, 2, 3\}$.



When the median lines intersect at a single point, i.e., $m_1 + m_2 + m_3 = 0$, clearly the intersection (m_1, m_2, m_3) is the only point that minimizes S. Because $|x_i - a_i| + |x_i - b_i| + |x_i - c_i| = d_i$ for $i \in 1, 2, 3$, is minimized for all directions, $S = \frac{1}{2}(d_1 + d_2 + d_3)$.

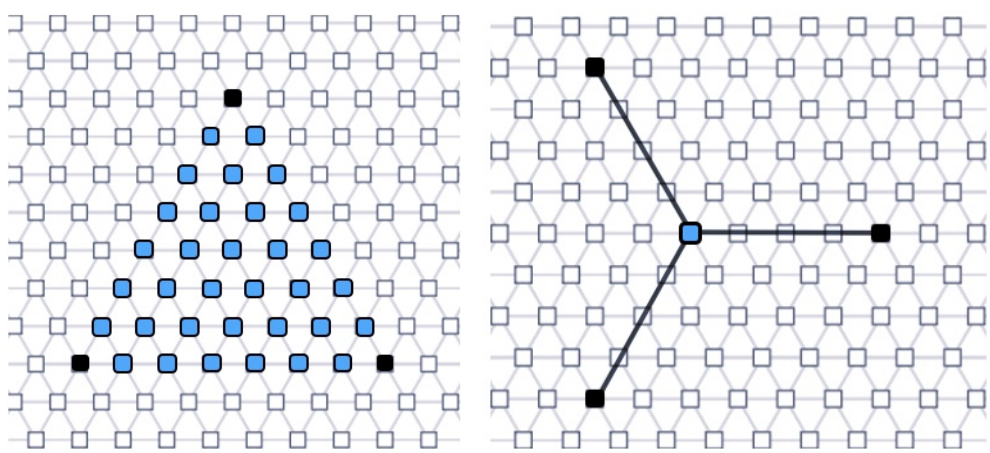
When the median lines do not intersect at one point within the median triangle,

$$|x_i - a_i| + |x_i - b_i| + |x_i - c_i| = |x_i - m_i| + d_i,$$

$$2S = \sum_{i=1}^{3} (|x_i - m_i| + d_i) = \sum_{i=1}^{3} |x_i - m_i| + \sum_{i=1}^{3} d_i.$$

Note that $\sum_{i=1}^{3} |x_i - m_i|$ is the total distance from point X to the three edges of the median triangle, which is a constant for all points in the triangle according to Lemma 2.5. $\square$

Lemma 2.8. *Special cases:* (a) For three terminal points that form an equilateral triangle along the grid lines, the Steiner region is coincident with the equilateral triangle itself (Figure 12a). (b) For three terminal points, if their median lines meet at the same point, the Steiner region degenerates to a single point. There is a *unique* Steiner minimum tree for this terminal set. Call this terminal set the *spinner set* and this unique Steiner minimum tree the *spinner tree* (Figure 12b).



(a) (b)

Figure 12: (a): The Steiner region is the entire equilateral triangle determined by the terminal points. (b): The spinner set has a unique Steiner minimum tree, the spinner tree. Black denotes terminal points; blue denotes Steiner points.

CHAPTER 3

Exact Algorithms

3.1 Four terminal vertices on a triangular grid

Introducing a fourth terminal vertex to the three-point Steiner problem brings uncertainties when we consider the topology of Steiner trees. For three terminal vertices, there is only one Steiner point and we simply connect it to each terminal to obtain the Steiner tree. The position of the Steiner point and the length of the Steiner tree can be calculated from Theorem 2.1 without much effort. For four terminal vertices, there are two Steiner points and a Steiner tree is constructed so that one Steiner point connects two terminal points, the other connects the remaining two terminals, and the two Steiner points are connected to each other. The complexity arises since we need to consider all possible pairings among the terminal vertices, resulting different Steiner tree topologies, in order to find the minimum Steiner tree. The solution to a four-point Steiner problem will provide the foundation for finding a general solution to an n-point triangular Steiner problem.

Consider four terminal points, $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$, $C(c_1, c_2, c_3)$ and $D(d_1, d_2, d_3)$. There will be two Steiner points in a Steiner tree. Steiner point S_1 connects A and B. Steiner point S_2 connects C and D. S_1 is connected with S_2 (Figure 13).

Lemma 3.1. S_1 is the Steiner point for $\{A, B, S_2\}$ and S_2 is the Steiner point for $\{C, D, S_1\}$.

Proof. We prove this by contradiction.

The total length of the tree is $|AS_1|+|BS_1|+|S_1S_2|+|CS_2|+|DS_2|$. Suppose S_1 is not a Steiner point for $\{A, B, S_2\}$. Then we can find another Steiner point S' for

$\{A, B, S_2\}$ such that $|AS'|+|BS'|+|S'S_2| < |AS_1|+|BS_1|+|S_1S_2|$, a contradiction. Therefore, the satisfying point S_1 must be a Steiner point for $\{A, B, S_2\}$. The same is true for S_2. $\square$

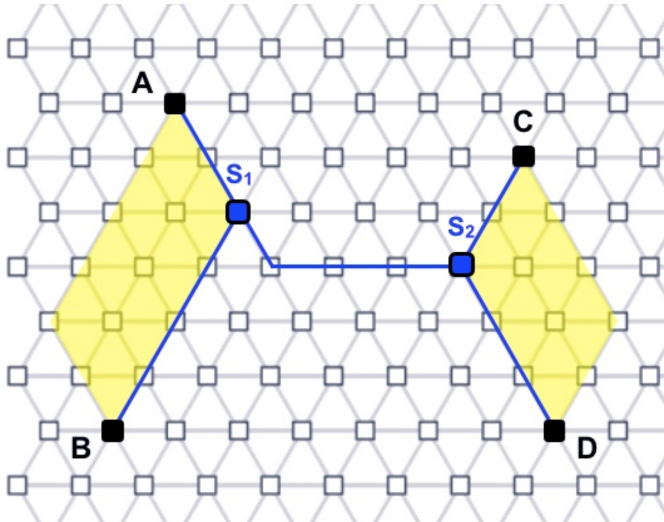


Figure 13: SMT for four terminal vertices.

The following lemma determines whether a point is inside the boundary of three terminal vertices. In the implementation of Algorithm 3.1, it is used to check if S_1 is inside the median triangle of $\triangle ABS_2$ and S_2 is inside the median triangle of $\triangle CDS_1$.

Lemma 3.2. Consider a terminal set $\{A, B, C\}$ and an arbitrary point P on the triangular grid. The coordinates of the points are $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$, $C(c_1, c_2, c_3)$, and $P(p_1, p_2, p_3)$. To find $P = \alpha A + \beta B + \gamma C$, we solve the system

$$\alpha a_1 + \beta b_1 + \gamma c_1 = p_1$$

$$\alpha a_2 + \beta b_2 + \gamma c_2 = p_2$$

$$\alpha + \beta + \gamma = 1$$

for α, β and γ. Then P is inside $\triangle ABC \iff \alpha \geqslant 0, \beta \geqslant 0, \text{and } \gamma \geqslant 0$. $\square$

Lemma 3.1 is a necessary condition for the Steiner point pairs. From its proof we also see that by satisfying this condition we will find the Steiner minimum tree for a particular topology. However, when searching for the Steiner pairs, we need to narrow our choices based on some criteria rather than brute-forcing all pairs of points on the grid. Therefore we introduce the concept of *candidate points* for a Steiner point.

Definition (Candidate Points/Set)**.** Consider two points A and B on the grid. If A and B are noncollinear, the *candidate points* for A and B are the four vertices of their distance parallelogram. Otherwise if A and B are collinear, the *candidate points* are A and B themselves. The set of candidate points we shall call the *candidate set*.

The construction of a Steiner tree can also be considered from the viewpoint of the distance parallelograms (DP). A and B are connected by some route within their DP, DP_1. C and D are also connected by some route within their DP, DP_2. These two components are then connected by the shortest path between DP_1 and DP_2, yielding a locally minimum tree. The two ends of the path touching DP_1 and DP_2 are Steiner points S_1 and S_2. As a result, S_1 and S_2 are on the outer edges of the distance parallelograms. This observation leads to the following lemma.

Lemma 3.3. Consider the shortest path between two parallelograms that do not overlap. There must exist a shortest path that includes a vertex of either parallelogram.

Corollary 3.4. At least one of S_1 and S_2 belongs to the vertices of its corresponding distance parallelogram, i.e., the candidate set.

Corollary 3.4 is an immediate result of Lemma 3.3 and here we illustrate the correctness of Lemma 3.3 with figures.

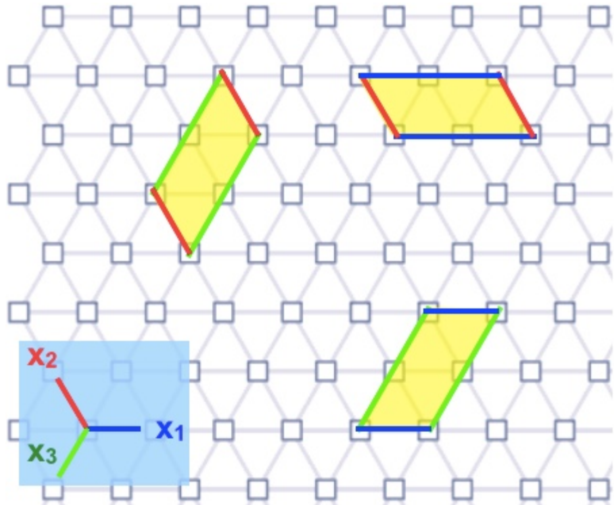


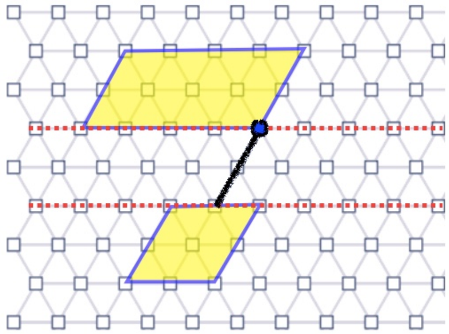
Figure 14: Three orientations of a parallelogram on a grid.

A parallelogram on the grid can be in any of the three orientations shown in Figure 14. For two parallelograms on the grid, there are two relative positions between them according to how many pairs of parallel edges these two parallelogram share. In Figure 15a, where there are two pairs of parallel edges, we find the closest pair that do not cut through the parallelograms and the shortest path that ends at a vertex. In Figure 15b, where there is only one pair of parallel edges, the closest vertex from one parallelogram to the other will be included in the shortest path.

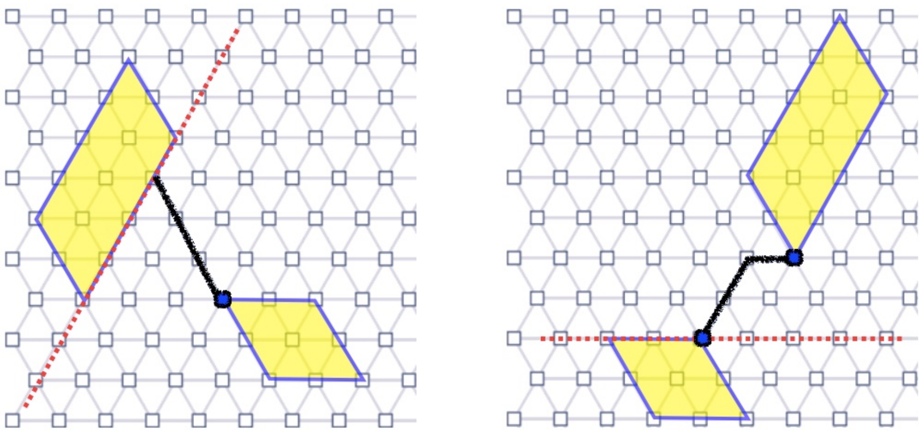
Lemma 3.5. Let $\mathcal{S}_1$ be the candidate set of A and B, $\mathcal{S}_2$ be the candidate set of C and D. The Steiner point pair (S_1, S_2) for this particular Steiner tree topology satisfies either (1) $S_1 \in \mathcal{S}_1$, $S_2 \in \mathcal{S}_2$, or (2) $S_1 = S_2$, $S_1 \in \mathcal{S}_1 \bigcup \mathcal{S}_2$.

Corollary 3.6. It is sufficient to search for Steiner pairs within $\{(P, Q) | P, Q \in \mathcal{S}_1 \bigcup \mathcal{S}_2\}$ to obtain a Steiner tree.

Proof. Without loss of generality, assume A and B are as shown in Figure 16. Draw three grid lines through A and B respectively. The grid plane is divided into 16 regions surrounding the distance parallelogram of A and B. Label eight of



(a)



(b)

Figure 15: Relative positions and a shortest path between two parallelograms.

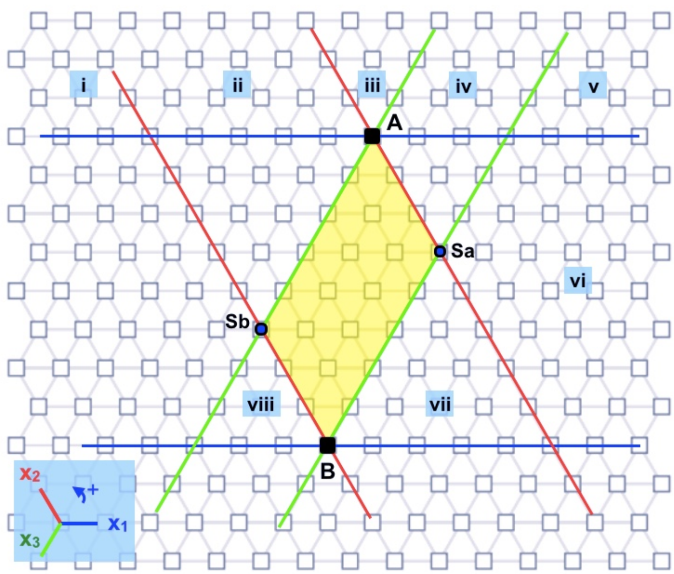


Figure 16: Regions surrounding the distance parallelogram of A and B.

these regions i through $viii$, with the others being their symmetric halves.

The candidate set $\mathcal{S}_1$ for A and B is $\{A, B, Sa, Sb\}$. We want to show that, *no matter where the other Steiner point S_2 is, Steiner point S_1 can always be found in* $\mathcal{S}_1 \bigcup \{S_2\}$, i.e., there exists $S_1 \in \mathcal{S}_1 \bigcup \{S_2\}$ such that S_1 is in the median triangle of $\triangle ABS_2$.

We do a case analysis by assuming S_2 to be in regions i through $viii$ (Figure 17).

1. S_2 in region i. A provides two median lines and B provides the third median line. Choose $S_1 \in \{A, S_b\}$.

2. S_2 in region ii. A provides two median lines and S_2 provides the third median line. Choose $S_1 \in \{A\}$.

3. S_2 in region iii. All three median lines pass through A and it is the only choice for S_1. Choose $S_1 \in \{A\}$.

4. S_2 in region iv. A provides two median lines and S_2 provides the third median line. Choose $S_1 \in \{A\}$.

5. S_2 in region v. A provides two median lines and B provides the third median line. Choose $S_1 \in \{A, S_a\}$.

6. S_2 in region vi. A, B, S_1 each provide a median line. Choose $S_1 \in \{S_a\}$.

7. S_2 in region vii. S_2 provides two median lines and B provides the third median line. Choose $S_1 \in \{S_2\}$.

8. S_2 in region $viii$. S_2 provides two median lines and B provides the third median line. Choose $S_1 \in \{S_2\}$.

In cases 7 and 8, S_1 coincides with S_2. By Corollary 3.4, if S_1 is not in $\mathcal{S}_1$, then S_2 must be a vertex in $\mathcal{S}_2$. $\square$

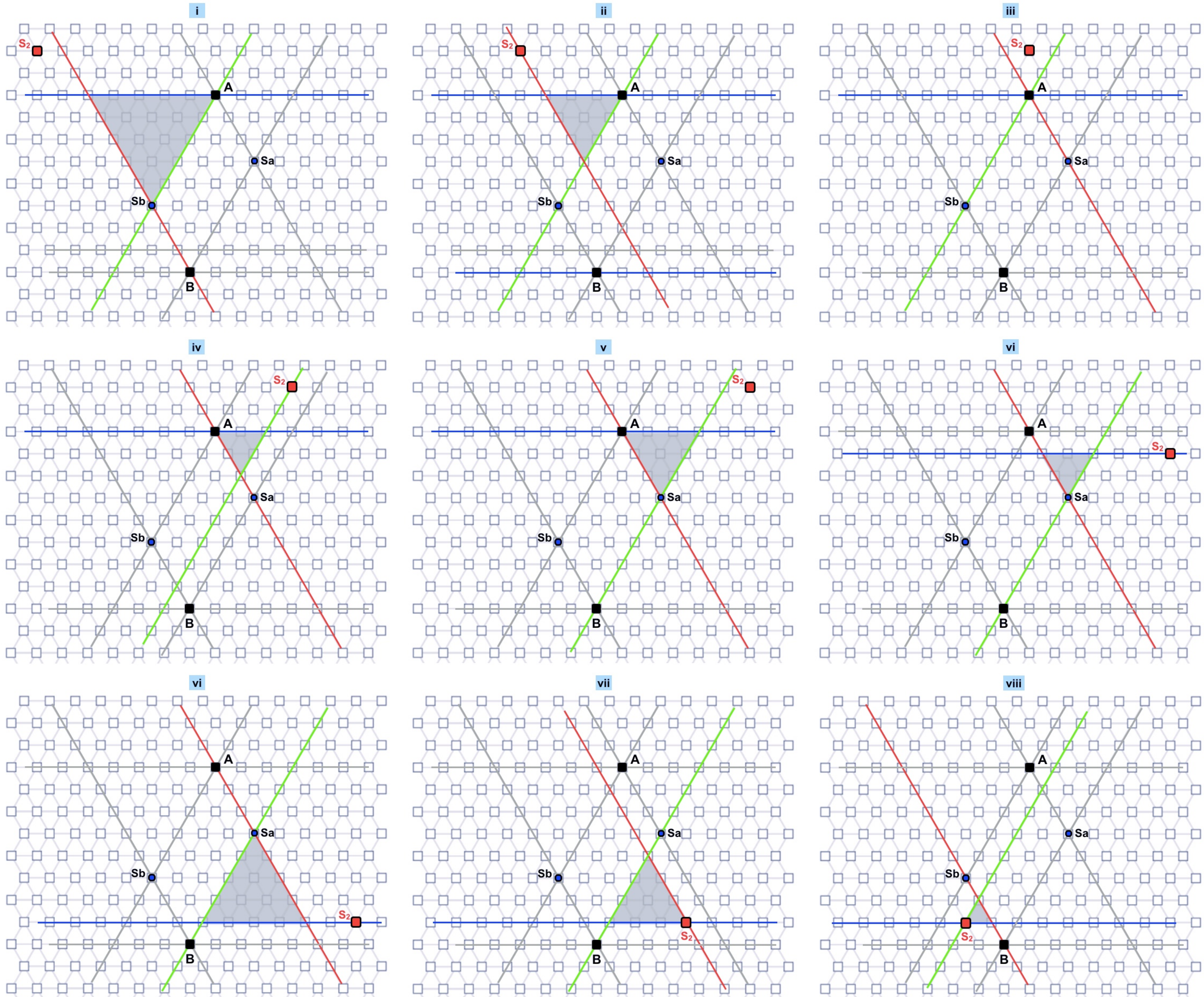


Figure 17: S_2 in different regions. The gray shaded triangle is the median triangle of $\triangle ABS_2$.

With Lemma 3.5 and Corollary 3.6 in hand, it is easy to compute the local minimum Steiner tree for a given topology. All topologies should be considered in order to find the global Steiner minimum tree. In Algorithm 3.1 we provide a procedure for finding the SMT for four terminal points.

Algorithm 3.1 (Steiner Minimum Tree for 4 points)

Input: A set of terminal points $\mathcal{X} = \{A, B, C, D\}$

Output: Length of the Steiner minimum tree.

Let $\mathcal{P}$ be all possible 2-2 partitions of $\mathcal{X}$,

$\mathcal{P} = \{\{(A, B), (C, D)\}, \{(A, C), (B, D)\}, \{(A, D), (B, C)\}\}$

for each partition $\{(p_1, p_2), (p_3, p_4)\}$ in $\mathcal{P}$:

  for each candidate s_1 of (p_1, p_2):   // 4 choices

    for each candidate s_2 of (p_3, p_4):   // 4 choices

      Check if s_1 is inside the median triangle of $\{s_2, p_1, p_2\}$

      Check if s_2 is inside the median triangle of $\{s_1, p_3, p_4\}$

      // s_1 connecting p_1, p_2; s_2 connecting p_3, p_4

      $\text{local1} = |s_1p_1| + |s_1p_2| + |s_2p_3| + |s_2p_4| + |s_1s_2|$

      // special case: degenerate Steiner points at s_1

      $\text{local2} = |s_1p_1| + |s_1p_2| + |s_1p_3| + |s_1p_4|$

      // special case: degenerate Steiner points at s_2

      $\text{local3} = |s_2p_1| + |s_2p_2| + |s_2p_3| + |s_2p_4|$

      $\text{global\_min} = \min(\text{global\_min}, \text{local1}, \text{local2}, \text{local3})$

return global_min.

3.2 The recursive algorithm

The five-point Steiner problem is considered before we tackle the general case with n terminal vertices. Suppose T is a Steiner minimum tree for a five-terminal set $\mathcal{X} = \{A, B, C, D, E\}$. If vertices A and B are connected to Steiner point S_1, then $T - S_1A - S_1B$ must be the SMT for four-terminal set $\{S_1, C, D, E\}$. Clearly $S_1A + S_1B$ is the shortest distance connecting A and B. We can choose S_1 to be in the candidate set of these two terminal vertices and we say A and B are *merged* to their candidate point S_1. The following lemma shows that a Steiner point can be obtained by merging two vertices to their candidate set when we construct a Steiner minimum tree.

Lemma 3.7. A Steiner point can be found in the candidate set of some vertex pair in a Steiner minimum tree.

Proof. Suppose in a Steiner minimum tree, vertices A and B are joined to Steiner point S and S is not in the candidate set of (A, B), denoted by $Candidate(A, B)$. We show by case analysis that we can choose S from the candidate set of some vertex pair to obtain an SMT with the same length.

1. S is a terminal point. Merging *terminal* vertex pair (S, A) to their candidate set then merging (S, B) shall we find this topology.

2. S is not a terminal and it has degree 3. Let S connect to A, B and X. S must be the Steiner point for set $\{A, B, X\}$. From the proof of Corollary 3.6, we can choose S from $Candidate(A, B) \bigcup \{X\}$.

3. S is not a terminal and it has degree 4. Let S connect to A, B, X and Y. From Lemma 3.5, S can be chosen from $Candidate(A, B) \bigcup Candidate(X, Y)$. Since $S \notin Candidate(A, B)$, it can be chosen from $Candidate(X, Y)$.

Since the degree of a Steiner point can only be 3 or 4 (see Lemma 2.7), this concludes our proof. □

Now the idea for a recursive algorithm to solve an n-point Steiner problem becomes straightforward. In each round we choose two terminal points P and Q and merge them into their candidate set $\mathcal{C}$. For every member C in $\mathcal{C}$, we compute the Steiner tree for the $n-1$ terminal set, comprised of the original terminal set with P and Q replaced by C.

Algorithm 3.2 (Steiner Minimum Tree for n points)

Input: A set of terminal points $\mathcal{X}$

Output: Length of Steiner minimum tree.

function SMT($\mathcal{X}$):

n = number of elements in $\mathcal{X}$

if $n \leqslant 3$:

return $\mathrm{SMT}_3(\mathcal{X})$ // one-step SMT calculation for 3 points

for each pair (p_1, p_2) in $\mathcal{X} \times \mathcal{X}$, $p_1 \neq p_2$: // $\binom{n}{2}$ pairs

for each candidate s of (p_1, p_2):

let $\mathcal{X}' = (\mathcal{X} - \{p_1, p_2\}) \bigcup \{s\}$

local_tree_length = $|sp_1| + |sp_2| + \mathrm{SMT}(\mathcal{X}')$

global_min =min(global_min, local_tree_length)

return global_min

Runtime analysis

We did a rough estimate of the runtime of the basic recursive algorithm. In each terminal reducing step, each pair of two points is selected to be merged to its candidate set (maximum size of 4). So the runtime is governed by a recurrence relation of the form $t(n) = 4 \cdot \binom{n}{2} \cdot t(n-1)$, whose solution is $O(2^n (n!)^2)$, where n is the number of terminal vertices to be connected.

The recursive algorithm was implemented in Maple. Though effective in finding the shortest route, the program is inefficient and can be used to find the SMT for only up to 6 points within a reasonable amount of time. The results are presented in Table 1.

Table 1: Runtimes of recursive algorithm implemented in Maple

No. terminals	3	4	5	6	7
Runtime	0.6 ms	0.14 s	5.5 s	360 s	58875 s (16 h)

3.3 The binary tree model

The principal redundancy of the basic recursive algorithm comes mainly from reconsidering terminal pairs. An example of a Steiner tree for a five-point set $\{1, 2, 3, 4, 5\}$ is shown in Figure 18. In the basic recursive program, this topology will be considered at least twice: (1) merge terminal 1 and 2 to S_1 $\rightarrow$ set reduced to $\{S_1, 3, 4, 5\}$ $\rightarrow$ merge terminal 3 and 4 to S_2 $\rightarrow$ set reduced to $\{S_1, S_2, 5\}$; (2) merge terminal 3 and 4 to S_2 $\rightarrow$ set reduced to $\{1, 2, S_2, 5\}$ $\rightarrow$ merge terminal 1 and 2 to S_1 $\rightarrow$ set reduced to $\{S_1, S_2, 5\}$.

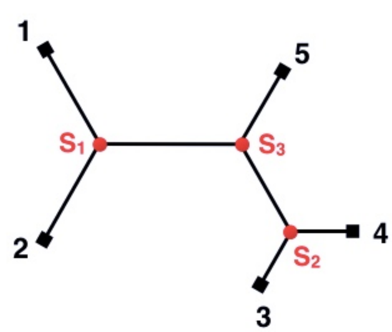


Figure 18: Diagram of a Steiner tree for 5 terminal vertices. 1...5 are the terminals; S_1, S_2 and S_3 are Steiner points.

To eliminate counting a tree topology multiple times, we use a binary tree model to represent a Steiner tree topology and consider pairing terminals[1] at the

1 In our recursive program, terminals not only include those in the original terminal set, but also refer to merged points that are treated as new terminal points in the next recursive step.

beginning of the program. Figure 19 shows the binary tree representation of the example in Figure 18. In a binary tree, the leaves correspond to terminal points and all internal vertices excluding the root correspond to Steiner points. Note that the root does not represent any vertex in the Steiner tree and its children are simply the two components it connects.

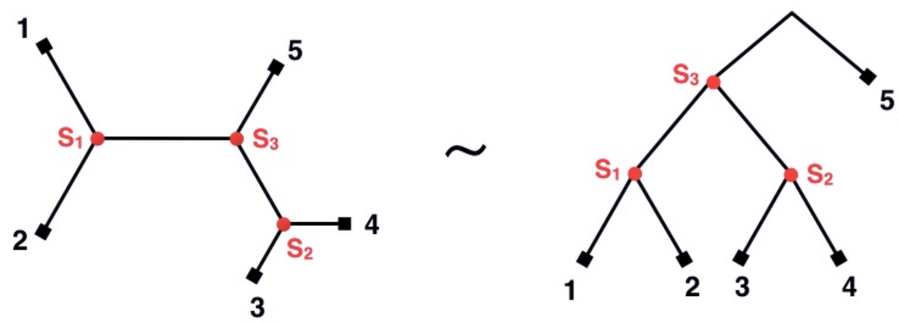


Figure 19: A Steiner tree and its binary tree representation.

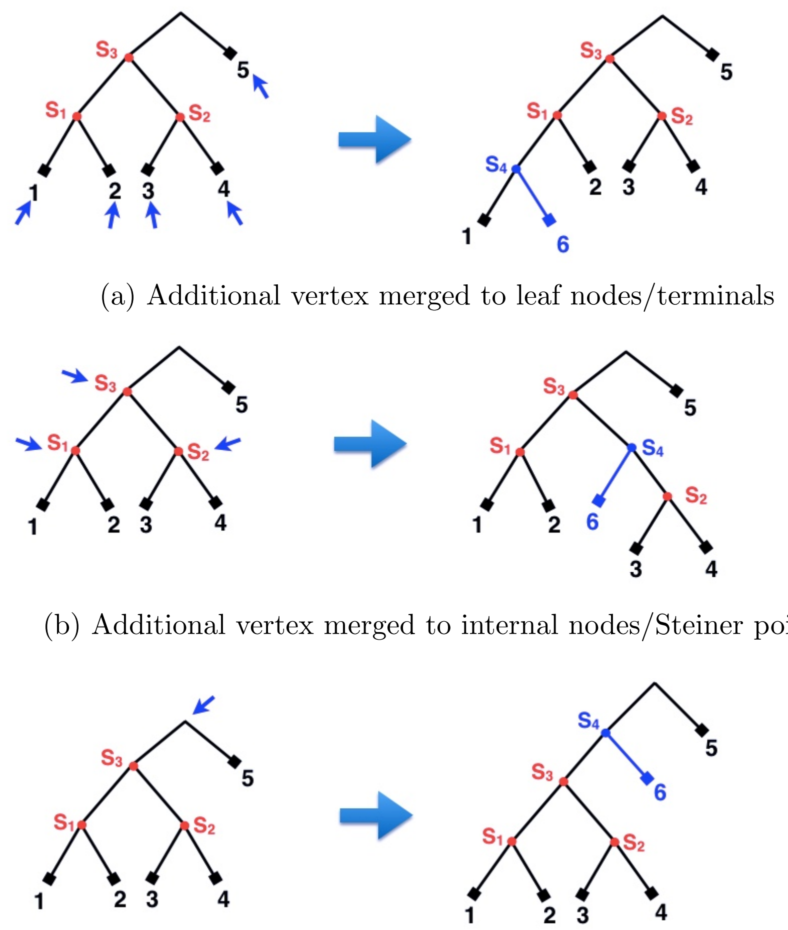
Lemma 3.8. For n terminal points, there are $(2n-3)!!$ binary trees.

Proof. Consider how many ways we can add a terminal point to a given topology. The additional point can be merged to either the leaf nodes or the internal vertices including the root (Figure 20). So the recurrence relation for the number of binary trees is:

$$T(n) = (\underbrace{n-1}_{\text{leaves}} + \underbrace{n-2}_{\text{internal vertices}}) \cdot T(n-1)$$

$$\begin{aligned} \Longrightarrow \qquad T(n) &= (2n-3)\cdot(2n-5)\cdots 5 \cdot T_3 \\ &= (2n-3)!! \quad \text{since} \quad T_3 = 3 \end{aligned}$$

□



(a) Additional vertex merged to leaf nodes/terminals

(b) Additional vertex merged to internal nodes/Steiner points

(c) Additional vertex merged to root

Figure 20: Adding an additional terminal vertex to a binary tree.

Implementation and runtime analysis

Similar to the original recursive program, the binary tree model has the runtime recurrence relation $t(n) = 4 \cdot (2n-3) \cdot t(n-1)$, and

$$t(n) \;=\; 4^{n-3} \cdot (2n-3)!! \;=\; 4^{n-3} \cdot \frac{(2n-2)!}{2^{n-1}\,(n-1)!} \;=\; 2^{n-5} \cdot \frac{(2n-2)!}{(n-1)!}.$$

So asymptotically the runtime is $O(2^n \dfrac{(2n)!}{n!})$, where n is the size of the initial terminal set. Compared to the basic recursive program, the improvement brought about by the binary tree model is huge as one can see simply from the two recurrence relations. The basic algorithm is quadratically dependent on the previous term while the binary tree algorithm is linearly dependent on the previous term.

However, there still remains some redundancy with the binary tree model. Take the simplest Steiner tree consisting of three terminal points as an example. There are three binary tree representations while only one topology exists for this terminal set (Figure 21). This three-body redundancy exists for all Steiner points connecting three components. So, in order to efficiently implement the recursive binary tree algorithm, we used a global hash table to store all previously computed SMTs to avoid any recalculation, at the expense of requiring additional space. This procedure is given in Algorithm 3.3.

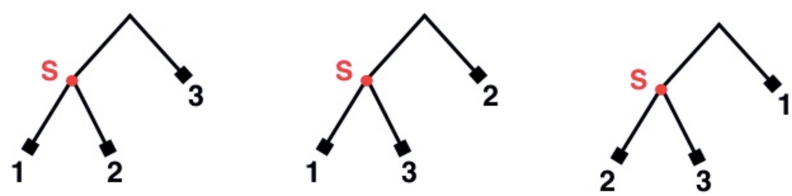


Figure 21: Binary trees for three terminal points.

Algorithm 3.3 (Steiner Minimum Tree for n points, improved implementation)

Input: A set of terminal points $\mathcal{X}$

Output: Length of Steiner minimum tree.

global H = Hashtable(point set $\rightarrow$ SMT length)

function SMT($\mathcal{X}$):

    n = number of elements in $\mathcal{X}$

    if $n \leqslant 3$:

        if H has no entry $\mathcal{X}$:

            $H[\mathcal{X}] = \mathrm{SMT}_3(\mathcal{X})$ // one-step SMT calculation for 3 points

        return $H[\mathcal{X}]$

    for each pair (p_1, p_2) in $\mathcal{X} \times \mathcal{X}$, $p_1 \neq p_2$: // $\binom{n}{2}$ pairs

for each candidate s of (p_1, p_2):

let $\mathcal{X}' = (\mathcal{X} - \{p_1, p_2\}) \bigcup \{s\}$

if H has entry $\mathcal{X}'$:

local_tree_length = $|sp_1| + |sp_2| + H[\mathcal{X}']$

else

$H[\mathcal{X}'] = \mathrm{SMT}(\mathcal{X}')$

local_tree_length = $|sp_1| + |sp_2| + H[\mathcal{X}']$

global_min =min(global_min, local_tree_length)

return global_min

The results of the improved implementation for computing Steiner trees are shown in Table 2, including a comparison with the original basic program. The speedup is significant and the limits of its feasibity has been pushed to 8 terminal points from 6. However, this performance improvement involves a tradeoff between space and time. Part of the factorial cost growth is transferred to space usage of the hash table, and how efficiently Maple implements a hash table may also be a limit of computing power.

Table 2: Runtimes of binary tree (BT) vs basic recursive program in Maple

No. terminals	4	5	6	7	8
BT runtime	0.13 s	3.2 s	36 s	389 s	5083 s
Basic runtime/BT runtime	1.1	1.7	10	151	-

3.4 Pruning trees

The following lemma provides a pruning criteria to cut down the number of terminal pairs considered when computing an SMT.

Lemma 3.9. In a Steiner minimum tree, a terminal point is not necessarily inside the distance parallelogram of any grouped pair of terminal vertices.

Proof. Consider a Steiner minimum tree T of terminal set $\mathcal{X}$. Suppose terminals A and B are grouped and merged to Steiner point S. S connects to the remaining component of T with edge SO. Terminal C is inside the distance parallelogram of A and B.

Remove edges SA, SB and SO in T. Connect edges CA and CB. We obtain a new Steiner tree T' of $\mathcal{X}$. The length of T' is $|T'| = |T| - |SA| - |SB| - |SO| + |CA| + |CB| = |T| - |SO| \leqslant |T|$. □

Thus, there is no need to consider grouping two terminals (A and B) if any other terminal point (C) is found to be inside their DP. In any event, the optimum Steiner tree will be considered by grouping either (A, C) or (B, C) (Figure 22).

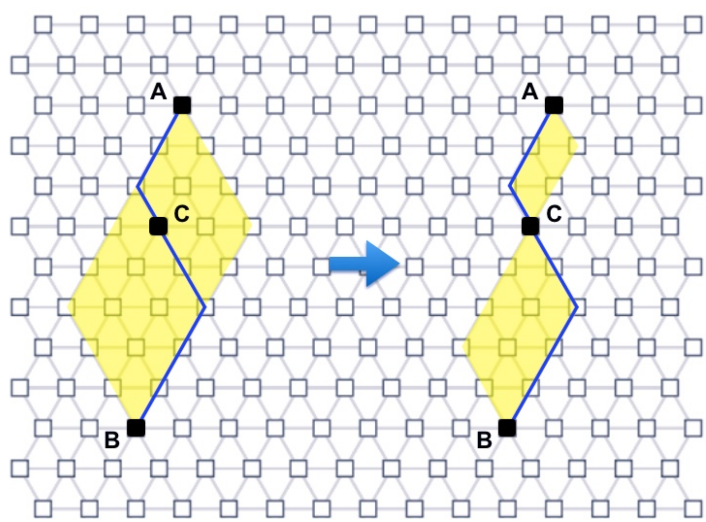


Figure 22: Discard grouping of terminals A and B when terminal C is found inside the DP of A and B.

The pruning method has been implemented for both the basic recursive program and the binary tree version. Before each exponential SMT calculation step, we check each terminal pair to see if another terminal point is inside their DP. This condition can easily be checked in linear time. There are large speedups on the basic program (Table 3) but no obvious improvement on the binary tree recursive program (not shown).

We conclude our discussion on exact algorithms for computing Steiner minimum trees with a performance comparison of the three algorithms implemented in Maple (Table 3). As this research focuses on the development of algorithms, and Maple is not the most efficient language for computing purposes, we are satisfied with the data presented and leave more in-depth analyses for future work.

Table 3: Performance comparison of the three SMT programs in Maple.

No. terminals	4	5	6	7	8
Basic runtime	0.14 s	5.5 s	360 s	58875 s	-
Basic with pruning (BP)	-	2.7 s	143 s	6905 s	-
Binary tree runtime (BT)	0.13 s	3.2 s	36 s	389 s	5083 s
Basic runtime/BP runtime	-	2.0	2.5	8.5	-
Basic runtime/BT runtime	1.1	1.7	10	151	-

CHAPTER 4

Computational Complexity

The Steiner problem on the Euclidean plane [4] and on the rectilinear plane [7] have both been proven to be NP-hard. It is intuitive to imagine that on the triangular plane, the triangular Steiner problem (TRISMT) is also NP-hard. To show the NP-hardness of the TRISMT problem it is sufficient to find a special class of triangular SMTs so that the computation of such trees is NP-hard. On the Euclidean plane, such a class of Euclidean SMTs was found by Rubinstein et al [5]. In this construction, terminals are constrained to be on two parallel lines. Based on this configuration, Weng [8] shows that the Euclidean Steiner minimal tree (ESMT) problem is NP-hard by transforming the SUBSET SUM problem into the ESMT problem. Using a construction similar to Rubinstein's and Weng's, we are able to find a special class of terminal sets on the triangular grid that can be polynomially transformed from the SUBSET SUM problem.

An instance of the SUBSET SUM problem is described as follows. *Given a set of positive integers* $S = \{d_1, d_2, \cdots, d_n\}$ *and an integer* s *(*$0 \leqslant s \leqslant D = \sum_{1 \leqslant i \leqslant n} d_i$*), is there a subset* $J \subset \{1, 2, \cdots, n\}$ *such that* $\sum_{i \in J} d_i = s$*?*

Our proof develops in these steps:

1. We describe the TRISMT problem as a decision problem. *Given a set of terminals* $\mathcal{X}$ *on a triangular grid and some integer* l*, is there a Steiner tree* T *spanning* $\mathcal{X}$ *such that* $|T| \leqslant l$*?*

2. Construct an instance of the TRISMT problem from an instance of SUBSET SUM.

3. Show that if the SUBSET SUM instance has a "yes" solution, then we can

find an SMT of length l for the terminal set in the TRISMT instance within polynomial time.

4. Show that if we can find an SMT of length l for the TRISMT instance, then we can find the solution to the SUBSET SUM instance within polynomial time.

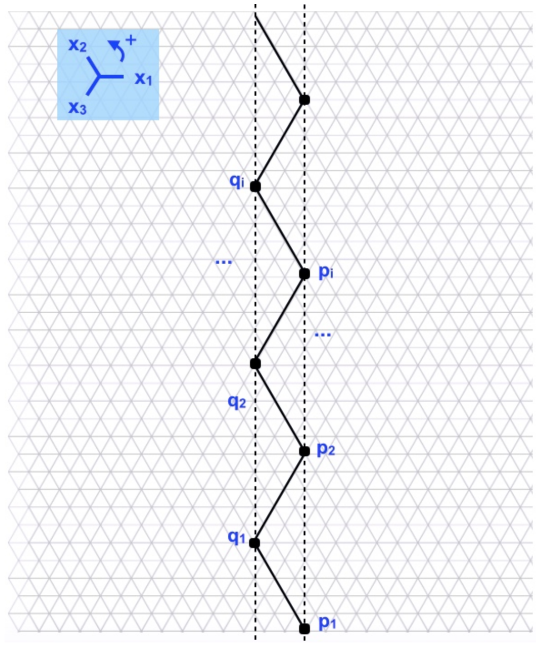
First we give lemmas to describe some properties of SMTs for terminals that sit on two parallel lines. Note that we choose the two auxiliary lines to be vertically parallel to each other and the grid points on these lines are 2 unit lengths apart.

Lemma 4.1. Consider $2n$ terminal vertices on a triangular grid lying on two vertical lines, with n points on each line, such that every point and its neighboring point on the other vertical line are collinear along axis-aligned lines. Connect all the collinear pairs. The resulting zigzag path aligned to the grid is the SMT for this $2n$ terminal set (Figure 23a). $\square$

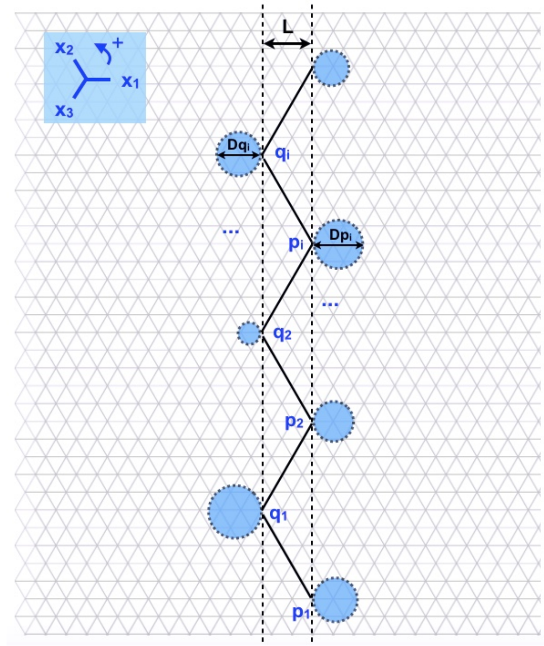
Corollary 4.2. Suppose each terminal point in Lemma 4.1 is replaced by an "island" of terminal points, where the diameter of the island (the furthest distance between any two points) is much smaller than the distance between the two vertical lines. The topology of the SMT for the new set is such that the islands are connected by a zigzag path and within each island, the terminal points are connected with local SMTs (Figure 23b). $\square$

Next we construct an instance of TRISMT (Figure 24). Let $L_2 \gg L_1 \gg D$.

1. Draw two vertical lines l_1 and l_2 at a distance L_2. Let u_i $(i = 0, 1, \cdots n)$ be grid points on l_2 and x_i $(i = 0, 1, \cdots n)$ be grid points on l_1 such that $\{u_i, x_i\ (i = 0 \cdots n)\}$ forms a terminal set as described in Lemma 4.1. Connect these points with a zigzag path.



(a) A zigzag path spanning a set of terminal points lying on two vertical lines.



(b) A zigzag path with turning points that connect to local SMTs. Each blue balloon suggests a local SMT for the terminals inside. L is much greater than the diameter of any disk (Dp_i and Dq_i).

Figure 23: Zigzag path as an SMT and its "island" derivative.

2. Draw a third vertical line l_0 to the left of l_1, where the distance between l_0 and l_1 is L_1. Extend $u_n x_0$ to intersect l_0 at v. Place triples of vertices, $\{v_i, v_i', v_i''\ (i = 1 \cdots n)\}$, on l_0 such that: (1) v_i is $2d_i$ below x_i, (2) v_i' is $2d_i$ above x_i, (3) v_i'' is $4d_i$ above v_i'.

   Consider the terminal set $\mathcal{X} = \{v, u_0, u_i, v_i, v_i', v_i''\ (i = 1 \cdots n)\}$. By Corollary 4.2, the SMT of $\mathcal{X}$, $T_{\mathcal{X}}$, is the zigzag path $u_0 x_1 u_1 x_2 \cdots u_n x_0 v$, plus local SMTs connecting $v_i, v_i', v_i'', x_i\ (i = 1 \cdots n)$. Let the length of $T_{\mathcal{X}}$ be M.

3. Let v_0 be a grid point on l_0 that is $8s$ below v in the x_1 direction. The instance terminal set is $\mathcal{X}' = \{v_0, u_0, u_i, v_i, v_i', v_i''\ (i = 1 \cdots n)\}$, i.e. $\mathcal{X}' = (\mathcal{X} - \{v\}) \bigcup \{v_0\}$.

Let the SMT of $\mathcal{X}'$ be $T_{\mathcal{X}'}$. Similar to $T_{\mathcal{X}}$, $T_{\mathcal{X}'}$ is composed of a zigzag path

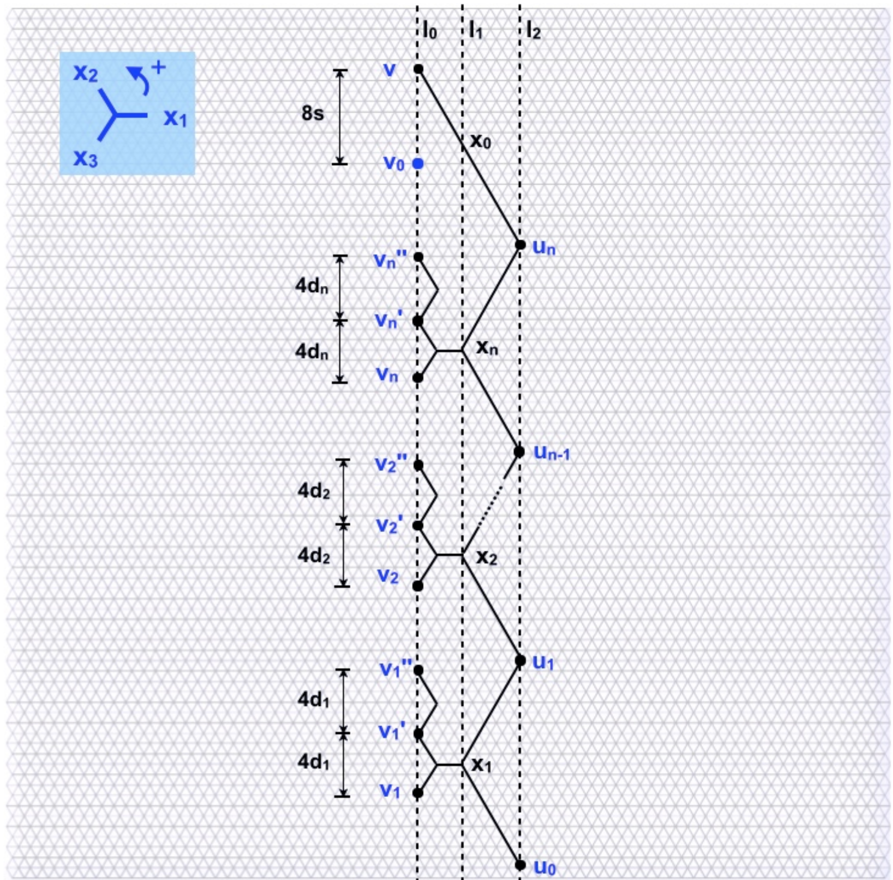


Figure 24: Terminal set $\mathcal{X} = \{v, u_0, u_i, v_i, v_i', v_i''\ (i = 1 \cdots n)\}$ and its SMT. $x_i\,(i = 0 \cdots n)$ are Steiner points. Terminal set $\mathcal{X}' = (\mathcal{X} - \{v\}) \bigcup \{v_0\}$.

joining local SMTs of the triple set terminals $\{v_i, v_i', v_i''\ (i = 1 \cdots n)\}$. Next we show that the SUBSETSUM instance has a "yes" solution $\Longleftrightarrow |T_{\mathcal{X}'}| = M - 6s$.

For a given subset $J \subset \{1, 2, \cdots, n\}$, construct $T_{\mathcal{X}'}$ as follows (Figure 25):

1. Join every island terminal set $\{v_i, v_i', v_i''\ (i = 1 \cdots n)\}$ with path $v_i v_i' v_i''$ between lines l_0 and l_1. Let w_i be the vertex of the distance parallelogram of v_i and v_i', and w_i' be the vertex of the DP of v_i' and v_i''.

2. For $i \in J$, the island set joins the zigzag path with w_i'; for $i \notin J$, the island set joins the zigzag path with w_i.

3. Starting from terminal u_0, use a zigzag path to join u_i, islands of $\{v_i, v_i', v_i''\}$

$(i = 1 \cdots n)$ and v_0. The new turning points in this zigzag path are x_i' and u_i' $(i = 1 \cdots n)$. Note that for the last segment $u_n' v_0$, the two end points are simply connected and we do not know if they are collinear.

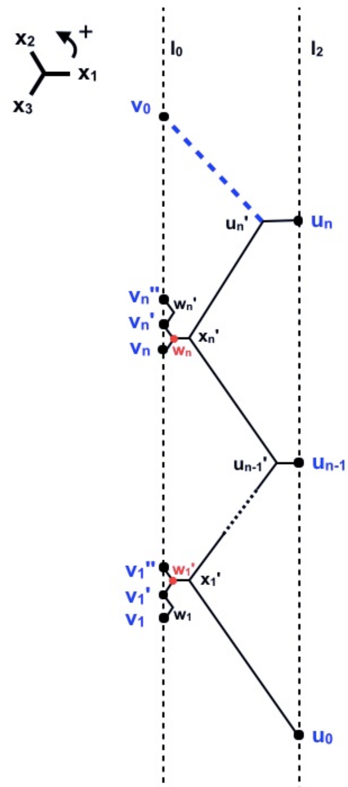


Figure 25: Terminal set $\mathcal{X}'$ and $T_{\mathcal{X}'}$. $1 \in J$ and $n \notin J$ in this example.

We now consider the length of $T_{\mathcal{X}'}$.

Lemma 4.3. For $1 \leqslant m \leqslant n$, let $J_m = J \bigcap \{1, 2, \cdots, m\}$. Horizontally, x_m' lies to the left of x_m by $2 \sum_{i \in J_m} d_i$. u_m' lies to the left of u_m by $4 \sum_{i \in J_m} d_i$.

Proof. Figure 26 shows a partial view of $T_{\mathcal{X}}$ and $T_{\mathcal{X}'}$. Let the left shift amount from x_i to x_i' be δx_i, and from u_i to u_i' be δu_i for $i = 1, 2, \cdots, n$.

For $i \in J$:

$$\delta x_i = \delta u_{i-1} + 2d_i,\ \delta u_i = \delta x_i + 2d_i$$

For $i \notin J$:

$$\delta x_i = \delta u_{i-1},\ \delta u_i = \delta u_{i-1}$$

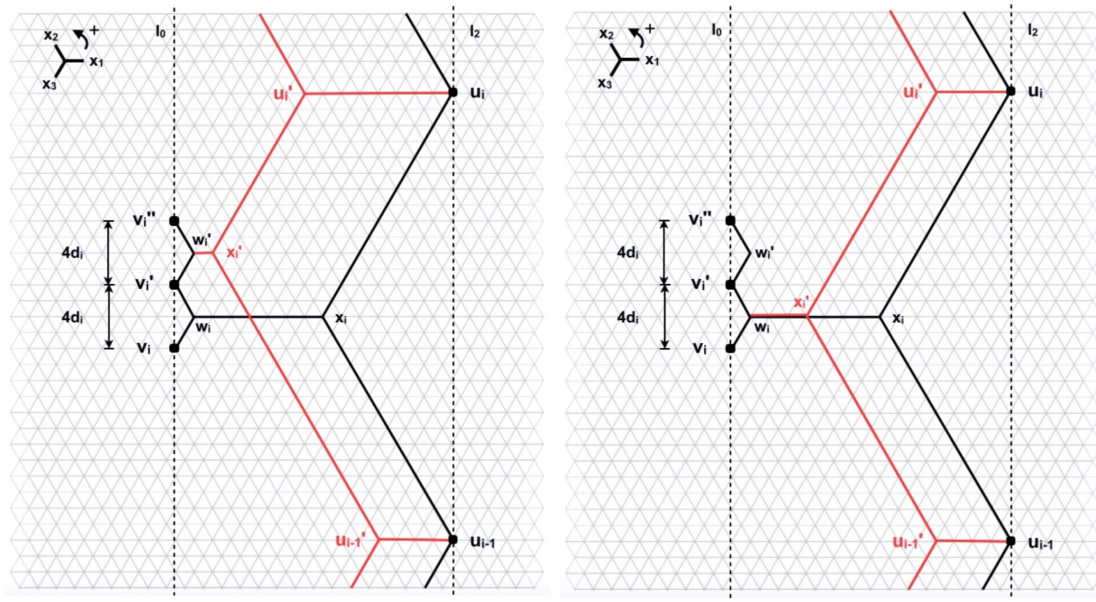
The initial conditions are

$$\delta u_0 = 0$$

and

$$\delta x_1 = 0 (1 \notin J) \text{ or } \delta x_1 = 2d_i (1 \in J).$$

The lemma follows naturally by solving the recurrences. $\square$



(a) $i \in J$ (b) $i \notin J$

Figure 26: Partial view of $T_{\mathcal{X}'}$ (red route) and $T_{\mathcal{X}}$ (black route).

Claim 1. $\sum_{i \in J} d_i = s \implies |T_{\mathcal{X}'}| = M - 6s.$

Proof. To compute $|T_{\mathcal{X}'}|$, we first consider the difference between the lengths of $T_{\mathcal{X}'}$ and $T_{\mathcal{X}}$ from stage $i-1$ to i (Figure 27). Let Δ_i be this difference.

For $i \in J$:

$$\Delta_i = (|u'_{i-1}x'_i| + |x'_i w'_i| + |x'_i u'_i| + |u'_i u_i|) - (|u_{i-1}x_i| + |x_i w_i| + |x_i u_i|) = 2d_i$$

For $i \notin J$:

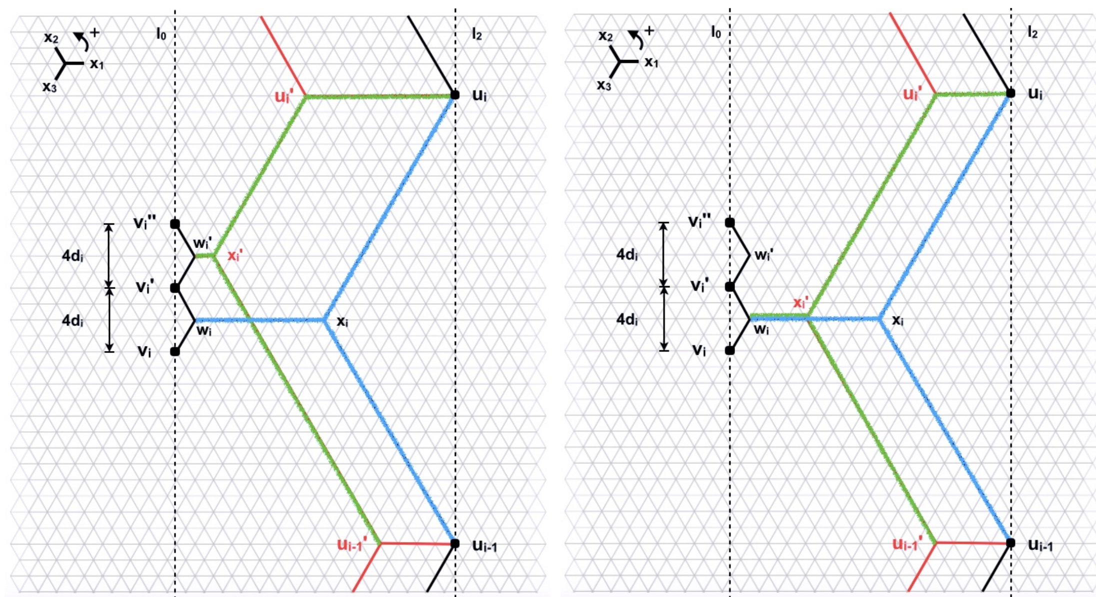
$$\Delta_i = (|u'_{i-1}x'_i| + |x'_i w_i| + |x'_i u'_i| + |u'_i u_i|) - (|u_{i-1}x_i| + |x_i w_i| + |x_i u_i|) = 0$$

The initial condition is $\Delta_0 = 0$.

$$\sum_{i=0}^{n} \Delta_i = \sum_{i \in J} 2d_i = 2s$$

$$|T_{\mathcal{X}'}| - |T_{\mathcal{X}}| = \sum_{i=0}^{n} \Delta_i + (|u'_n v_0| - |u_n v|) \tag{$\star$}$$

From Lemma 4.3 we know that $|u'_n u_n| = \delta u_n = 4 \sum_{i \in J} d_i = 4s$. Because $|vv_0|$ is $8s$, the last segment $u'_n v_0$ must be aligned to the grid in the x_2 direction (Figure 28). Thus, $|u_n v| - |u'_n v_0| = 8s$ and $|T_{\mathcal{X}'}| = M - 6s$. $\square$



(a) $i \in J$ (b) $i \notin J$

Figure 27: Difference in length of $T_{\mathcal{X}'}$ (green route) and $T_{\mathcal{X}}$ (blue route) from stage $i-1$ to i.

Claim 2. The SMT of $\mathcal{X}'$ is of length $M - 6s \implies$ There exists $J \subset \{1, \cdots, n\}$, such that $\sum_{i \in J} d_i = s$.

Proof. Let $T_{\mathcal{X}'}$ be the SMT of $\mathcal{X}'$. By Corollary 4.2, T has a zigzag topology plus locally minimal Steiner trees connecting $\{v_i, v'_i, v''_i\}$. According to Lemma 3.9 (the tree pruning criteria), it is not necessary to join terminal v_i and v''_i. So the locally

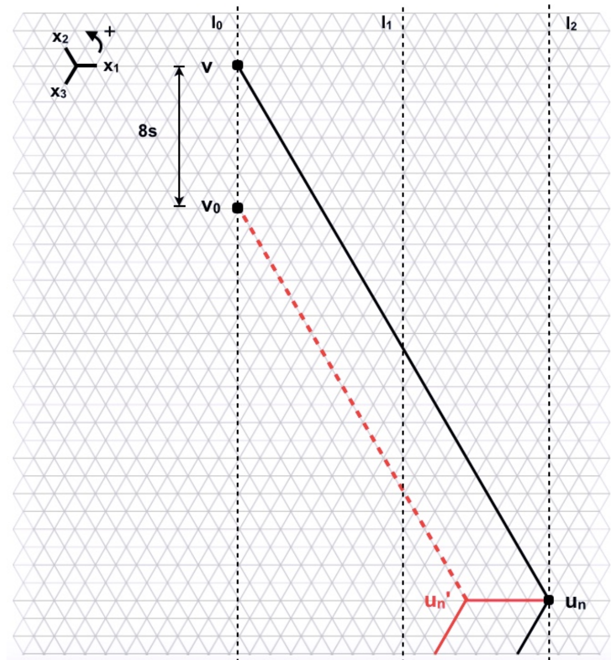


Figure 28: Last segments of $T_{\mathcal{X}}$ (black route) and $T_{\mathcal{X}'}$ (red route).

minimal Steiner trees can only be like the one in either Figure 29a or Figure 29b to join the zigzag path. And T must be as in Figure 25, with some $i \in J$ connecting the upper part of the island and some $i \notin J$ connecting the lower part.

(a) upper

(b) lower

Figure 29: Local SMTs joining the zigzag path.

Consider the last segment $u_n' v_0$.

1. If it is aligned with the grid line in the x_2 direction, then $|T_{\mathcal{X}'}| = M - 6 \sum_{i \in J} d_i$. From our assumption that $T_{\mathcal{X}'} = M - 6s$, $\sum_{i \in J} d_i = s$.

2. If $u_n' v_0$ is not aligned with the grid line, let the ray from u_n' in the x_2 direction

intersect the line l_0 at v_0'. When v_0 is above v_0', edge $u_n'v_0$ in $T_{\mathcal{X}'}$ has length $|u_n'v_0'| + |v_0'v_0|$ (Figure 30a),

Equation ($\star$) then becomes

$$
\begin{aligned}
|T_{\mathcal{X}'}| - |T_{\mathcal{X}}| &= \sum_{i=0}^{n} \Delta_i + (|u_n'v_0| - |u_n v|) \\
&= 2\sum_{i\in J} d_i + |u_n'v_0'| + |v_0'v_0| - |u_n v| \\
&= 2\sum_{i\in J} d_i - 8s \\
M - 6s - M &= 2\sum_{i\in J} d_i - 8s \\
\implies \sum_{i\in J} d_i = s
\end{aligned}
$$

3. If $u_n'v_0$ is not aligned with the grid line and v_0 is below v_0', edge $u_n'v_0$ in $T_{\mathcal{X}'}$ is a bow leg in Figure 30b.

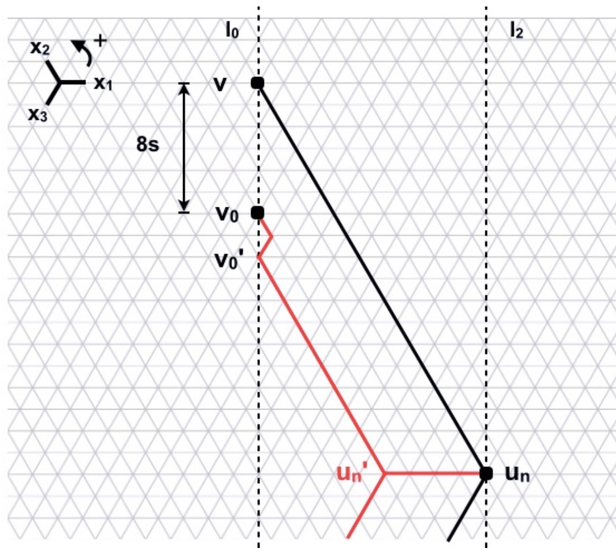
   Equation ($\star$) then becomes

$$
\begin{aligned}
|T_{\mathcal{X}'}| - |T_{\mathcal{X}}| &= \sum_{i=0}^{n} \Delta_i + (|u_n'v_0| - |u_n v|) \\
&= 2\sum_{i\in J} d_i + |tv_0| - 8s \\
&= 2\sum_{i\in J} d_i + (4s - |u_n'u_n|) - 8s \\
M - 6s - M &= -2\sum_{i\in J} d_i - 4s \\
\implies \sum_{i\in J} d_i = s
\end{aligned}
$$

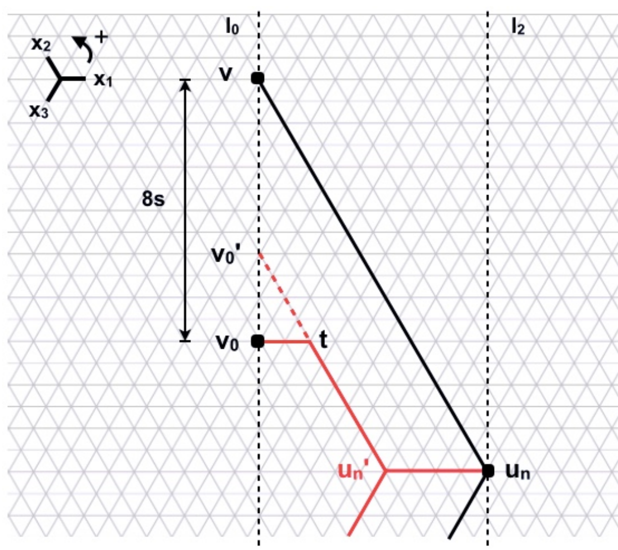
From the above case analysis on segment $u_n'v_0$ in $T_{\mathcal{X}'}$, we see the claim is correct. It also implies that v_0 and v_0' are the same point. □

Theorem 4.1. From Claim 1 and Claim 2, it follows that the solution to our instance of SUBSETSUM is "yes" $\iff$ SMT of $\mathcal{X}'$ has length $M - 6s$. So the TRISMT problem is NP-hard.

The theorem concludes our discussion for this chapter.



(a)



(b)

Figure 30: The situations for $u_n'v_0$, the last segment of $T_{\mathcal{X}'}$.

CHAPTER 5

The Steiner Ratio

5.1 The Steiner ratio conjecture

Given a terminal point set $\mathcal{X}$, a *minimum spanning tree* (MST) is the shortest connection network for $\mathcal{X}$ where all vertices belong to $\mathcal{X}$. A spanning tree differs from a Steiner tree in that no additional vertices are introduced in a spanning tree. An example is illustrated in Figure 31. Since there are no known polynomial time algorithms for computing a Steiner minimum tree (SMT), an MST can conveniently be used to approximate an SMT because there exist fast algorithms (Prim's and Kruskal's [9]) to compute an MST. Given a terminal set $\mathcal{X}$ on a triangular grid, we can find the distance between every pair of vertices in $\mathcal{X}$, and then apply either Prim's or Kruskal's algorithm to this complete graph. Since both Prim's and Kruskal's algorithms run in polynomial time, it is possible to find an MST on a triangular grid in a time that is polynomial in the number of terminal points.

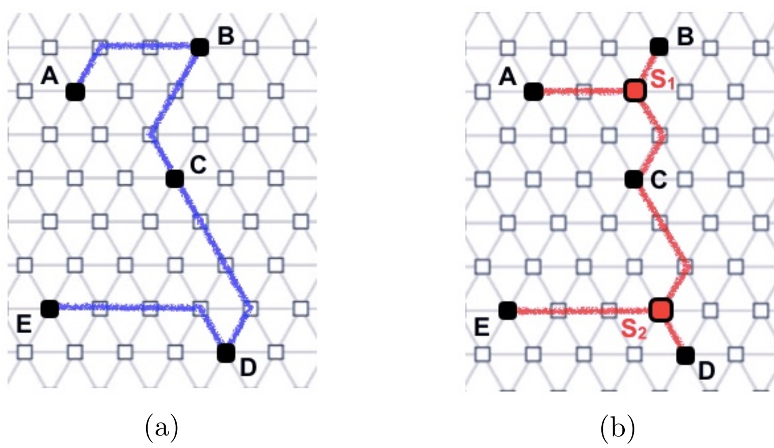


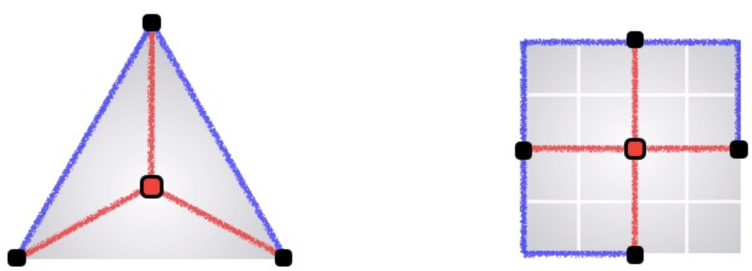
Figure 31: Terminal set $\mathcal{X} = \{A, B, C, D, E\}$. (a) MST for $\mathcal{X}$ with length 14. $V(MST) = \mathcal{X}$. (b) SMT for $\mathcal{X}$ with length 12. $V(SMT) = \mathcal{X} \bigcup \{S_1, S_2\}$.

The Steiner ratio has been studied for both the Euclidean and the rectilinear Steiner problems. Although this ratio has been defined in different ways, it is meant to quantify how well we can estimate an SMT with an MST. In this chapter our discussion of the Steiner ratio will be based on the following definition.

Definition (Steiner ratio)**.** Let $\rho = l_{MST}/l_{SMT}$ denote the ratio between the length of a minimum spanning tree and that of a Steiner minimum tree for a particular graph. The upper bound of ρ over all graphs is the *Steiner ratio.*

Lemma 5.1. $\rho \geqslant 1$ because spanning trees form a subset of Steiner trees. □

In the Euclidean plane, the Steiner ratio is conjectured to be $2/\sqrt{3}$ and this is achieved when three terminal points form an equilateral triangle (Figure 32a). This Steiner ratio conjecture on the Euclidean plane was proposed in 1968 by Gilbert and Pollak [10] and allegedly proven in 1990 by Du and Hwang [11]. Their claim was disproved by Ivanov and Tuzhilin in 2012 and the conjecture remains an open problem [12]. On a rectilinear grid, the Steiner ratio is 3/2 when four terminals are aligned in a cross shape (Figure 32b). The Steiner ratio for the rectilinear Steiner problem was proposed and proved by Hwang in 1976 [13].



(a) Euclidean plane (b) Rectilinear plane

Figure 32: Steiner ratio achieved for the Euclidean and rectilinear Steiner problems. Black vertices denote terminal points; red denote Steiner points. The red route gives the Steiner minimum tree; the blue route is the minimum spanning tree.

Conjecture 5.1 (Steiner ratio conjecture)**.** The Steiner ratio for the triangular Steiner problem is 4/3, i.e. $\rho \leqslant \frac{4}{3}$. The ratio is achieved when three terminals form a spinner set (defined in Lemma 2.8) and the three arms of the spinner tree are of equal length (Figure 33).

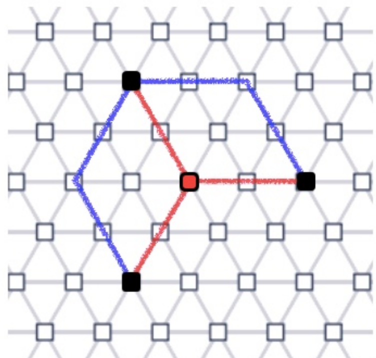


Figure 33: A spinner set (black points) that has a Steiner minimum tree with equal length arms.

In the remaining sections of this chapter, we will prove our conjecture for three and four point terminal sets.

5.2 Proof of the three point case

We describe two techniques, one graphic and the other algebraic, to prove the Steiner ratio conjecture for three terminal points.

Lemma 5.2. When one or more terminal points can serve as a Steiner point, there exists a Steiner minimum tree that is identical to the minimum spanning tree and $l_{SMT} = l_{MST}$ for the terminal set.

This is called a *degenerate* Steiner minimum tree (Figure 34). The terminal point that serves as a Steiner point provides at least two median lines for the median triangle.

Lemma 5.3. Given a spinner set of three terminals on a triangular grid, $\rho \leqslant 3/4$.

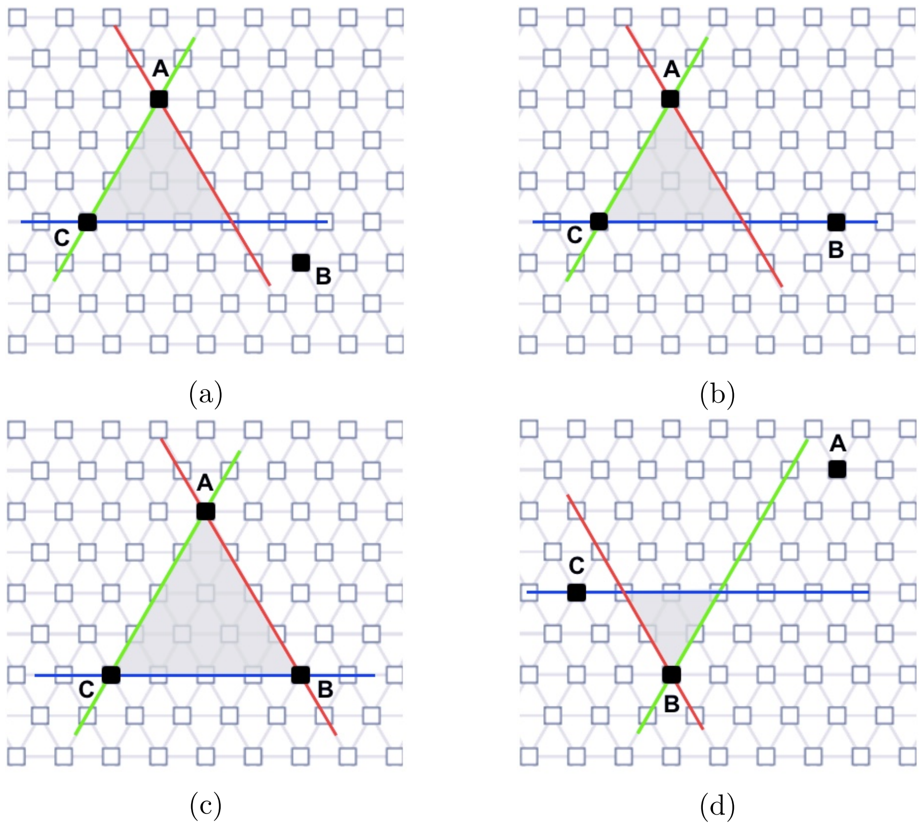


Figure 34: Different situations where an SMT degenerates. Terminal points are shown in black. The gray shaded triangles are median triangles. In (a), (b) and (c), at least two terminal points are collinear. In (d), all vertices are non-collinear.

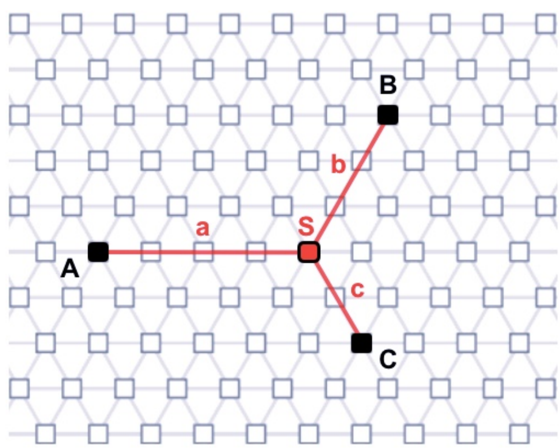


Figure 35: Spinner set $\mathcal{X} = \{A, B, C\}$ and the spinner tree.

Proof. Suppose the spinner set is $\mathcal{X} = \{A, B, C\}$. According to Lemma 2.8, there is a unique Steiner minimum tree and Steiner point S for $\mathcal{X}$ (Figure 35).

Let $|SA| = a$, $|SB| = b$, $|SC| = c$. Without loss of generality, suppose $a \geqslant b \geqslant c$. The total length of the SMT is $l_{SMT} = a + b + c$. In a metric system, $|AB| \leqslant |SA| + |SB| = a + b$. Similarly, $|AC| \leqslant a + c$ and $|BC| \leqslant b + c$.
The total length of the minimum spanning tree (MST) is

$$\begin{aligned} l_{MST} &= \min\{|AC| + |BC|,\ |AB| + |AC|,\ |AB| + |BC|\} \\ &\leqslant \min\{a + b + 2c,\ 2a + b + c,\ a + 2b + c\} = a + b + 2c \end{aligned}$$

Therefore, $\rho = l_{MST}/l_{SMT} \leqslant \dfrac{a+b+2c}{a+b+c} = 1 + \dfrac{c}{a+b+c} \leqslant 1 + \dfrac{c}{c+c+c} = \dfrac{4}{3}$. □

Lemma 5.4. Any non-spinner terminal set can be transformed to a spinner set without changing the total length of the Steiner minimum tree.

Proof. Consider a non-spinner terminal set $\{A, B, C\}$ on a grid. Choose one vertex of the median triangle, S, to construct a Steiner minimum tree (Figure 36). S is at the intersection of two median lines, one passing through A and the other through C in this example. Terminal B and S are non-collinear and will be transformed.

To construct the new set,

1. Draw a circle with center S and radius $|SB|$. The blue dashed line in Figure 36 shows a portion of this circle.

2. Draw the remaining axis line passing through S (green dashed line, a ray that does not run through the distance parallelogram of terminals A and C). The intersection of this line with the circle is terminal B'. $\{A, B', C\}$ is the transformed set.

Terminal set $\{A, B', C\}$ is a spinner set because all median lines intersect at the same point S. l_{SMT} for $\{A, B', C\}$ is the same as l_{SMT} for $\{A, B, C\}$ because $|SB| = |SB'|$. □

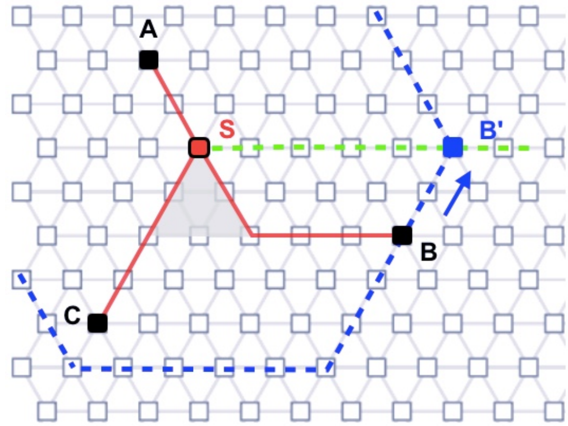


Figure 36: Terminal set $\{A, B, C\}$ transforms to spinner set $\{A, B', C\}$ on Steiner point S. S is the unique Steiner point for the new set.

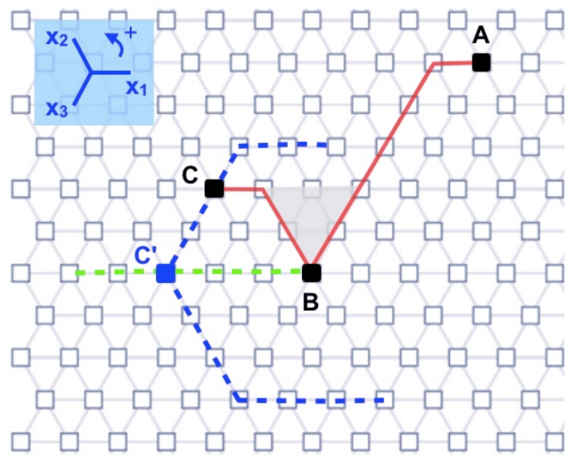


Figure 37: A degenerate non-spinner SMT transforms to a spinner tree. $\{A, B, C\}$ is the original terminal set and $\{A, B, C'\}$ is the spinner set after transformation.

The transformation can also be done on degenerate SMTs when one terminal point serves as the Steiner point. Take the case shown in Figure 34d as an example. In Figure 37, terminal B is the Steiner point for $\{A, B, C\}$ and B provides two median lines along axes x_2 and x_3. The intersection of axis x_1 passing through B and circle B with radius $|BC|$ is C'. $\{A, B, C\}$ can be transformed to spinner set $\{A, B, C'\}$ and their SMTs are of the same length.

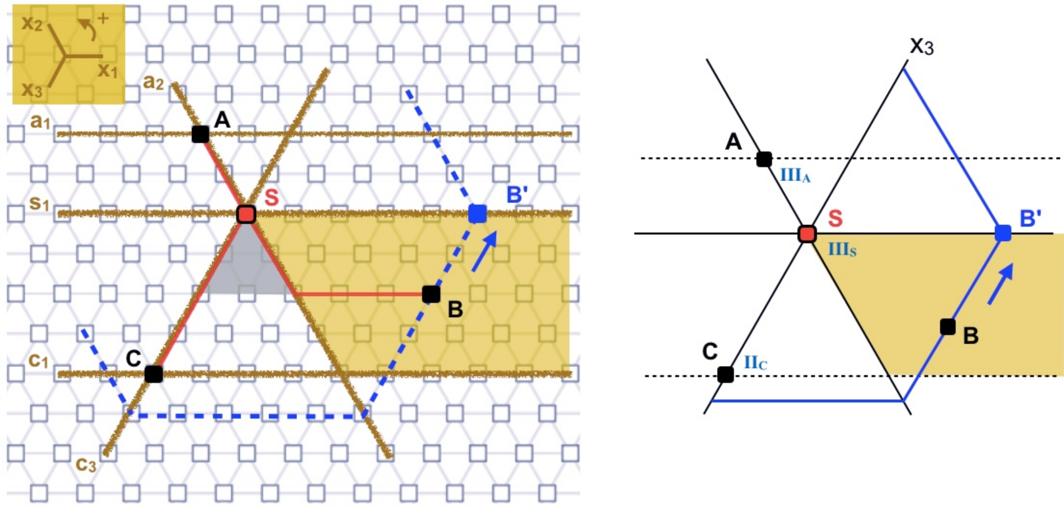
Lemma 5.5. l_{MST} for the transformed terminal set (the spinner set) from Lemma 5.4 is greater than or equal to l_{MST} for the terminal set before transformation.

Proof. Suppose a non-spinner set $\mathcal{X}$ transforms to a spinner set $\mathcal{X}'$. $l_{MST}(\mathcal{X})$ denotes the length of the minimum spanning tree of $\mathcal{X}$ and $l_{SMT}(\mathcal{X})$ denotes the length of the Steiner minimum tree of $\mathcal{X}$. We would like to show $l_{MST}(\mathcal{X}') \geqslant l_{MST}(\mathcal{X})$.

First consider the case when $\mathcal{X}$ has a degenerate SMT. Terminal point P can serve as the Steiner point for $\mathcal{X}$ and $l_{MST}(\mathcal{X}) = l_{SMT}(\mathcal{X})$. Transform $\mathcal{X}$ to $\mathcal{X}'$ on P. Because $P \in \mathcal{X}'$ and P is the unique Steiner point for $\mathcal{X}'$, $l_{MST}(\mathcal{X}') = l_{SMT}(\mathcal{X}')$. From the transformation we know $l_{SMT}(\mathcal{X}) = l_{SMT}(\mathcal{X}')$, and therefore $l_{MST}(\mathcal{X}') = l_{MST}(\mathcal{X})$.

Next we consider cases when $\mathcal{X}$ does not have a degenerate SMT. This suggests that each terminal provides a median line to form the median triangle. We use the example in the proof of Lemma 5.4 to continue our discussion. It shows a general situation under this assumption (Figure 38). Median line a_2 passes through A. C is in the positive half plane so B must lie in the negative half. Median line c_3 passes through C. A is in the negative half plane so B must lie on the positive half. B provides the median line in the remaining direction (x_1 in the example shown), separating A and C into each half plane. Therefore, B should lie between lines a_1 and c_1. Because of symmetry, B is further restricted to be between lines s_1 and c_1. The resulting region, the golden area shown in Figure 38, is where B can be located.

In the transformation, as B moves along the circumference of circle S to B' in direction x_3, B is in sector III of S and is also in the same sector of A. This suggests B is also moving along the circumference of circle A, with $|AB|$ being its radius. Thus, $|AB'| = |AB|$. Meanwhile, B is in sector II of C, moving away from C to increase $\angle CSB$ to $120°$ so $|B'C| \geqslant |BC|$. $|AC|$ remains the same after the transformation. Therefore, $l_{MST}(\mathcal{X}') \geqslant l_{MST}(\mathcal{X})$. □



(a) (b)

Figure 38: (a) Terminal set $\{A, B, C\}$ transforms to spinner set $\{A, B', C\}$ on Steiner point S. A, B, and C each provide a median line. The golden shaded area is where B can be located under this assumption. (b) Layout view of sectors of S.

The following theorem is a natural consequence of Lemmas 5.3 to 5.5. We also present below an alternative method to prove the ratio.

Theorem 5.1. The Steiner ratio for three arbitrary points on a triangular grid is 4/3.

Proof. Suppose the terminal vertices are $A(a_1, b_1, c_1)$, $B(a_2, b_2, c_2)$, and $C(a_3, b_3, c_3)$. Consider a Steiner tree for $\{A, B, C\}$ where the Steiner point is $X(x_1, x_2, x_3)$. Let $M_i = \max\{a_i, b_i, c_i\}$, $m_i = \min\{a_i, b_i, c_i\}$, $d_i = M_i - m_i$, $md_i =$ median$\{a_i, b_i, c_i\}$ for $i \in \{1, 2, 3\}$.

From previous results (see proof of Theorem 2.1), the total length of a tree is

$$l = |XA| + |XB| + |XC| = \frac{1}{2}\sum_{i=1}^{3}(|x_i - md_i| + d_i) = \frac{1}{2}(\sum_{i=1}^{3}|x_i - md_i| + \sum_{i=1}^{3} d_i).$$

So

$$l_{SMT} \geqslant \frac{1}{2}\sum_{i=1}^{3} d_i$$

Now consider the minimum spanning tree.

$$\begin{aligned} l_{MST} &\leqslant \text{twice the average of } |AB|, |BC|, |AC| = \frac{2}{3}(|AB| + |BC| + |AC|) \\ &= \frac{2}{3} \cdot (\frac{1}{2}\sum_{i=1}^{3}|a_i - b_i| + \frac{1}{2}\sum_{i=1}^{3}|b_i - c_i| + \frac{1}{2}\sum_{i=1}^{3}|a_i - c_i|) \\ &= \frac{1}{3}\sum_{i=1}^{3}(|a_i - b_i| + |b_i - c_i| + |a_i - c_i|) \\ &= \frac{1}{3}\sum_{i=1}^{3}2(M_i - m_i) = \frac{2}{3}\sum_{i=1}^{3}d_i \end{aligned}$$

$$\therefore \quad \frac{l_{MST}}{l_{SMT}} \leqslant \frac{4}{3}.$$

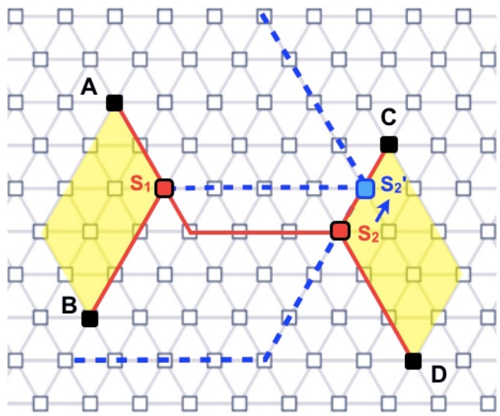
□

5.3 Proof of the four point case

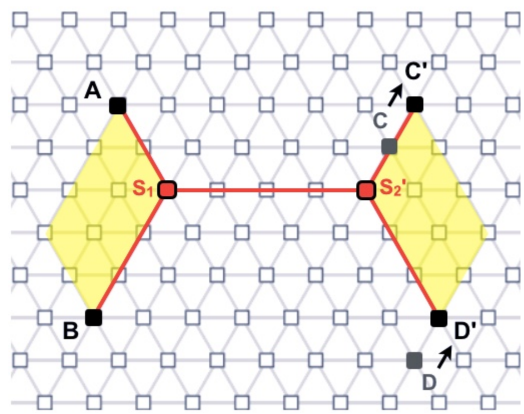
Consider the full Steiner tree for a four-terminal set $\mathcal{X} = \{A, B, C, D\}$ in Figure 39a. S_1 and S_2 are two Steiner points. From the Steiner tree construction described in Algorithm 3.1, assume edges S_1A, S_1B, S_2C and S_2D all align with grid lines. Edge S_1S_2 may not align with the grid, in which case $\{A, B, S_2\}$ and $\{C, D, S_1\}$ are not spinner sets. Using the transformation described in Lemma 5.4, transform S_2 to S_2'. Consequently, terminals C and D are transformed to C' and D' (Figure 39b). The resulting terminal set $\mathcal{X}' = \{A, B, C', D'\}$ has the same Steiner tree length as $\mathcal{X}$ and $l_{MST}(\mathcal{X}') \geqslant l_{MST}(\mathcal{X})$. Therefore, $\rho(\mathcal{X}) \leqslant \rho(\mathcal{X}')$. It would be sufficient to prove $\rho(\mathcal{X}') \leqslant 4/3$ for $\mathcal{X}'$, which has a double spinner Steiner tree. For the remainder of this section, a Steiner tree for four terminal points refers to a double spinner Steiner tree for four terminals.

Spanning trees

Spanning trees consist only of edges between terminal pairs. Given the topology of a double spinner Steiner tree for a terminal set $\mathcal{P} = \{A, B, C, D\}$, the edges joining a pair of terminal vertices can be categorized as one of the three types:

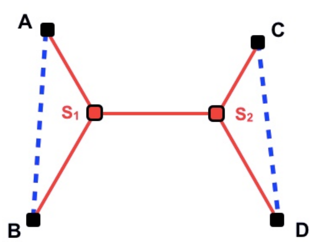


(a)

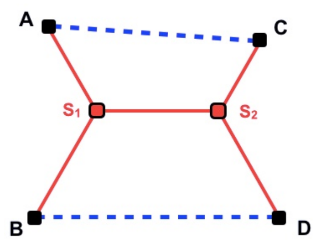


(b)

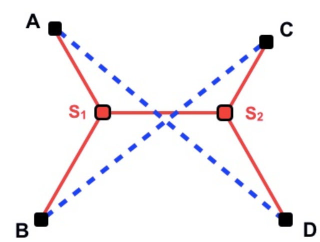
Figure 39: Terminal set $\{A, B, C, D\}$ transforms to $\{A, B, C', D'\}$ of the same Steiner tree length. The Steiner tree for $\{A, B, C', D'\}$ is a double spinner tree.



(a) neighbor



(b) bridge

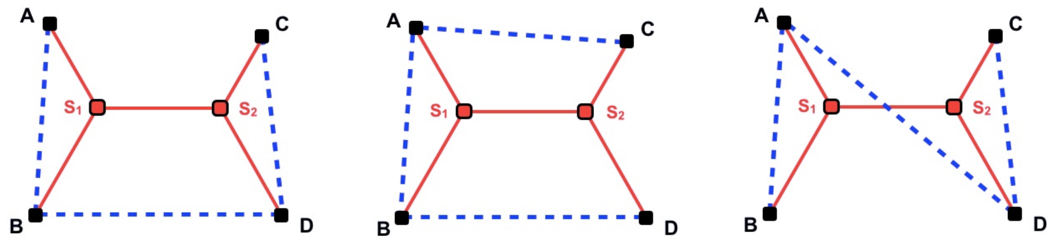


(c) cross

Figure 40: Types of edges joining two terminal points.

neighbors, bridges and crosses (Figure 40). A neighbor edge joins two terminals that are joined to the same Steiner point in a Steiner tree. A bridge edge joins two terminals that are directly connected to different Steiner points, and the bridge edge does not intersect the Steiner tree. A cross edge joins two terminals that are directly connected to different Steiner points and the cross edge intersects the Steiner tree. Examples of these types of spanning tree edges are given in Figure 40 in blue dashed lines and they reflect only the ways of joining the terminal pairs.

Based on the edge types in a spanning tree, there exist a total of seven different shaped spanning trees for $\mathcal{P}$. We name and describe three that will be useful in our discussion: U-shaped, C-shaped and N-shaped spanning trees (Figure 41).



(a) U-shaped (b) C-shaped (c) N-shaped

Figure 41: Different shaped spanning trees.

A U-shaped spanning tree contains two neighbor edges and a bridge edge; a C-shaped spanning tree contains two bridge edges and a neighbor edge; an N-shaped spanning tree contains two neighbor edges and a cross edge.

Steiner ratio

Lemma 5.6. Given a double spinner tree for four terminal points, $l_{MST}/l_{SMT} \leqslant 4/3$.

Proof. The proof is by contradiction.

Assume $l_{MST} > 4/3\, l_{SMT}$. This implies the length of *any* spanning tree must be greater than $4/3\, l_{SMT}$. It is sufficient to show any contradiction by choosing any spanning tree(s).

Suppose $\{A, B, C, D\}$ is a four-point terminal set and it has a double spinner Steiner tree with Steiner points S_1 and S_2. Let the lengths of the edges be $|AS_1| = a$, $|BS_1| = b$, $|S_1S_2| = s$, $|CS_2| = c$ and $|DS_2| = d$. So $l_{SMT} = a + b + c + d + s$. Without loss of generality, assume $a = \max\{a, b, c, d\}$. In order to express the length of the possible components of the spanning tree, we divide the proof into two cases.

1. Case $b \leqslant d$ (Figure 42).

   The distances between the terminal points are

   $|AB| = a + b,\ |BD| = s + d,\ |CD| = c + d,\ |BC| = s + b + c,\ |AC| = s + a$

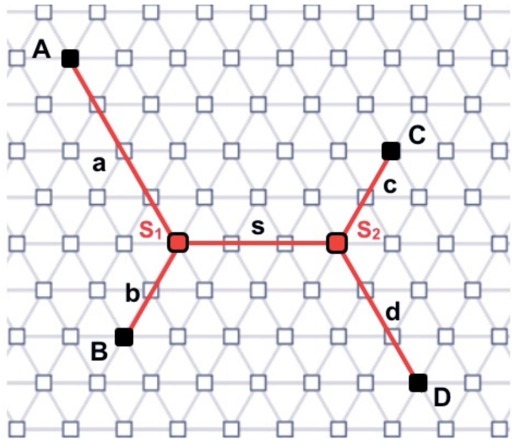


Figure 42: A double spinner tree for four terminal points. Case $b \leqslant d$.

Suppose

- for a U-shape, $(|AB| + |BD| + |CD|) > \frac{4}{3} l_{SMT}$

$$\implies 3(a + b + s + d + c + d) > 4(a + b + c + d + s)$$

$$\implies 2d > a + b + c + s \tag{1}$$

- for a C-shape, $(|AC| + |CD| + |BD|) > \frac{4}{3} l_{SMT}$

$$\implies 3(s + a + c + d + s + d) > 4(a + b + c + d + s)$$

$$\implies 2s + 2d > a + 4b + c \tag{2}$$

- for an N-shape, $(|AB| + |BC| + |CD|) > \frac{4}{3} l_{SMT}$

$$\implies 3(a + b + s + b + c + c + d) > 4(a + b + c + d + s)$$

$$\implies 2b + 2c > a + d + s \tag{3}$$

$(1) + (2) + (3) \implies d > a + b$, which contradicts $a > d$.

2. Case $b > d$ (Figure 43).

The distances between the terminal points are

$$|AB| = a + b,\ |BD| = s + b,\ |CD| = c + d,\ |AC| = s + a$$

Suppose

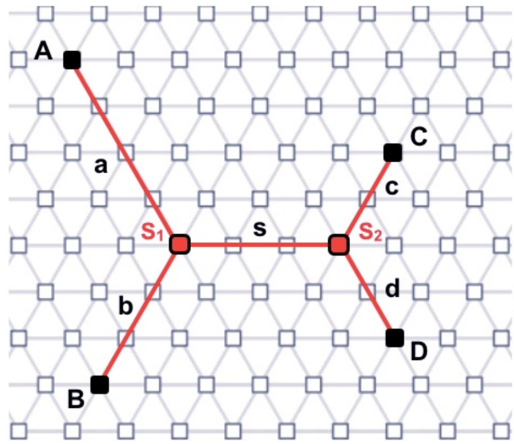


Figure 43: A double spinner tree for four terminal points. Case $b > d$.

- for a U-shape, $(|AB| + |BD| + |CD|) > \dfrac{4}{3}\, l_{SMT}$

$$\implies 3(a + b + s + b + c + d) > 4(a + b + c + d + s)$$

$$\implies 2b > a + c + d + s \tag{1}$$

- for a C-shape, $(|AC| + |CD| + |BD|) > \dfrac{4}{3}\, l_{SMT}$

$$\implies 3(s + a + c + d + s + b) > 4(a + b + c + d + s)$$

$$\implies 2s > a + b + c + d \tag{2}$$

$(1) \times 2 + (2) \implies b > a + c + d$, which contradicts $a > b$.

□

Theorem 5.2. The Steiner ratio for four points on a triangular grid is 4/3.

Proof. Lemma 5.6 shows that $l_{MST}/l_{SMT} \leqslant 4/3$. We need only prove that 4/3 is a tight upper bound.

Construct the Steiner tree in Figure 42 with $a = b = s = m$, $c = d = 1$. Then

$$l_{SMT} = a + b + c + d + s = 3m + 2,\ l_{MST} = |AB| + |BD| + |CD| = 4m + 2$$

The Steiner ratio is $l_{MST}/l_{SMT} = \dfrac{4m+2}{3m+2} \rightarrow \dfrac{4}{3}$ as $m \rightarrow \infty$.

□

Note that the bound of $\frac{4}{3}$ can be extended to any number of points. For example, in Figure 44, from the spinner set $\{1, 2, 3\}$ with an equal-arm spinner tree, add points one unit length apart along ray $S3$ to grow the set. m is the arm length of the spinner tree; n is the total number of terminal points. Let $m \gg n$, then $l_{MST}/l_{SMT} = \frac{4m+(n-3)}{3m+(n-3)} \to \frac{4}{3}$.

Our future work on the Steiner ratio includes generalizing the proof of the conjecture for four-terminal sets to n points using induction. Hwang [13] proved the Steiner ratio on a rectilinear grid by first showing all Steiner trees can be transformed to a fur Steiner tree shape (Figure 45a) and then conducting his proof on the fur tree through induction. Intuitively, we hope to prove Conjecture 5.1 for n points in a similar fashion: constrain the topology of a Steiner tree to a zigzag tree (Figure 45b) and prove the Steiner ratio by induction.

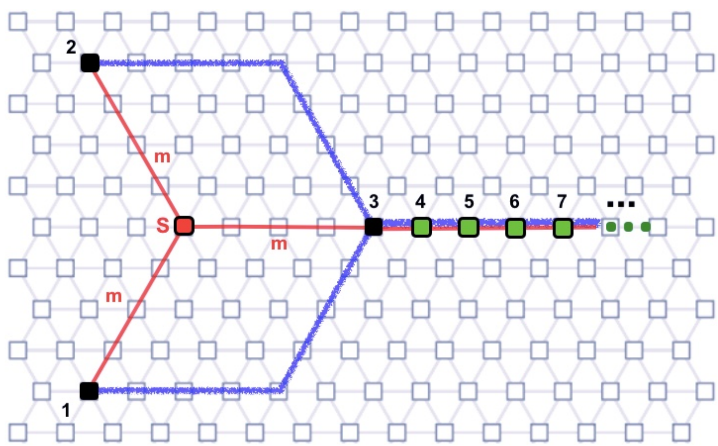
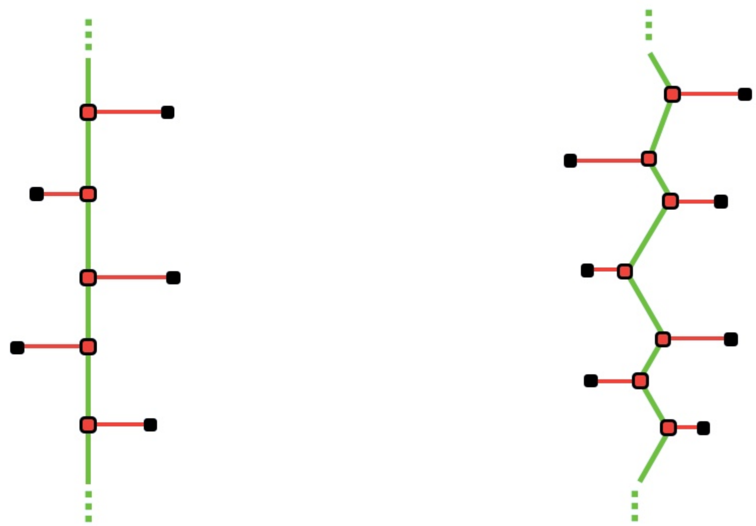


Figure 44: A Steiner tree (the red route) of $n \geqslant 4$ terminal points with $\rho \to \frac{4}{3}$. The black and green points denote terminals. The red point S is the only Steiner point in the Steiner tree. The blue route denotes the minimum spanning tree. $m \gg n$.



(a) A fur Steiner tree on a rectilinear grid. (b) A zigzag Steiner tree on a triangular grid.

Figure 45: Steiner trees on different geometric planes. Both consists of a backbone (green) and branches sticking out of both sides . Red denotes Steiner points; black denotes terminal points.

CHAPTER 6

Summary

In this thesis we investigated the shortest connection network problem on a triangular grid. The problem originates from the board game TransAmerica, where players build a railroad network linking five cities on a map that is a triangular grid. Previously the shortest path problem had been studied for the Euclidean plane and for a rectilinear grid. In both these cases, the problem was shown to be NP-hard.

Chapter 2 provides a foundation for this research. We introduced the triangular grid coordinate system and presented some key concepts and results which were drawn upon in later discussions. The solution to the three-point case was given. This important result, which is specific to the triangular metrics, was used in finding a general solution for any number of points, and in studying the computational complexity of the problem and approximation solutions.

In Chapter 3 we proposed several exact algorithms to solve the Steiner problem. Our first approach was an elementary recursive algorithm based on merging two terminal points to their candidate points to reduce the problem size. Groups of terminal points are enumerated multiple times in the basic algorithm. In order to reduce this redundancy, we introduced a binary tree model to represent the topology of a Steiner tree. In the actual implementation, we used a hash table to store all computed results to further improve efficiency. With these efforts, we were able to increase the limit of computation from 6 to 8 terminal points in Maple programs. A tree pruning technique was also developed and it significantly reduced the number of terminal pairs that need to be considered.

In Chapter 4 we showed that the triangular Steiner (TRISMT) problem is

NP-hard. Inspired by Rubinstein and Weng who proved the NP-hardness of the Euclidean Steiner problem by reducing it from the SUBSET SUM problem, we constructed a similar transformation for the triangular grid problem. To do this, we used a special class of triangular Steiner trees whose construction can be done from an instance of SUBSET SUM in polynomial time.

Since it is unlikely to find an efficient solution to the TRISMT problem, we considered approximating a Steiner minimum tree (SMT) with a minimum spanning tree (MST). An MST differs from an SMT in that it does not introduce additional vertices into the network. The goodness of this approximation, called the Steiner ratio, is an upper bound of the length of an MST to an SMT. We conjectured the Steiner ratio to be 4/3 and provided proofs for the three and four terminal points. The proof for four points shows promise of being generalized to n terminal points. This will be investigated in the future.

LIST OF REFERENCES

[1] M. W. Bern and R. L. Graham, "The shortest-network problem," *Scientific American*, vol. 260, pp. 84–89, January 1989.

[2] R. Courant and H. Robins, *What is Mathematics?: An Elementary Approach to Ideas and Methods.* Oxford University Press, 1941.

[3] Z. A. Melzak, "On the problem of steiner," *Canadian Mathematical Bulletin*, vol. 4, pp. 143–148, March 1961.

[4] M. R. Garey, R. L. Graham, and D. S. Johnson, "The complexity of computing steiner minimal trees," *SIAM Journal on Applied Mathematics*, vol. 32, pp. 835–859, 1977.

[5] J. H. Rubinstein, D. A. Thomas, and N. Wormald, "Steiner trees for terminals constrained to curves," *SIAM Journal on Discrete Mathematics*, vol. 10, pp. 1–17, 1997.

[6] M. Hanan, "On steiner's problem with rectilinear distance," *SIAM Journal on Applied Mathematics*, vol. 14, pp. 255–265, March 1966.

[7] M. R. Garey and D. S. Johnson, "The rectilinear steiner tree problem is np-complete," *SIAM Journal on Applied Mathematics*, vol. 32, pp. 826–834, June 1977.

[8] J. F. Weng, *Steiner trees, coordinate systems, and NP-hardness.* Springer, Boston, MA: Combinatorial Optimization, 2000, vol. 6, pp. 63–80.

[9] H. J. Prmel and A. Steger, *The Steiner Tree Problem.* Berlin, Germany: Vieweg, 2002.

[10] E. N. Gilbert and H. O. Pollak, "Steiner minimal trees," *SIAM Journal on Applied Mathematics*, vol. 16, pp. 1–29, January 1968.

[11] D. Z. Du and F. K. Hwang, "The steiner ratio conjecture of gilbert and pollak is true," in *Proceedings of the National Academy of Sciences of the United States of America*, December 1990, pp. 9464–9466.

[12] A. O. Ivanov and A. A. Tuzhilin, "The steiner ratio gilbert-pollak conjecture is still open," *Algorithmica*, vol. 62, pp. 630–632, 2012.

[13] F. K. Hwang, "On steiner minimal trees with rectilinear distance," *SIAM Journal on Applied Mathematics*, vol. 30, pp. 104–114, January 1976.

BIBLIOGRAPHY

Bern, M. W. and Graham, R. L., "The shortest-network problem," *Scientific American*, vol. 260, pp. 84–89, January 1989.

Brazil, M. and Zachariasen, M., *Optimal Interconnection Trees in the Plane.* Algorithms and Combinatorics, 2015, vol. 29.

Courant, R. and Robins, H., *What is Mathematics?: An elementary approach to ideas and methods.* Oxford university press, 1941.

Du, D. Z. and Hwang, F. K., "The steiner ratio conjecture of gilbert and pollak is true," in *Proceedings of the National Academy of Sciences of the United States of America*, December 1990, pp. 9464–9466.

Garey, M. R., Graham, R. L., and Johnson, D. S., "The complexity of computing steiner minimal trees," *SIAM Journal on Applied Mathematics*, vol. 32, pp. 835–859, 1977.

Garey, M. R. and Johnson, D. S., "The rectilinear steiner tree problem is np-complete," *SIAM Journal on Applied Mathematics*, vol. 32, pp. 826–834, June 1977.

Gilbert, E. N. and Pollak, H. O., "Steiner minimal trees," *SIAM Journal on Applied Mathematics*, vol. 16, pp. 1–29, January 1968.

Hanan, M., "On steiner's problem with rectilinear distance," *SIAM Journal on Applied Mathematics*, vol. 14, pp. 255–265, March 1966.

Hwang, F. K., "On steiner minimal trees with rectilinear distance," *SIAM Journal on Applied Mathematics*, vol. 30, pp. 104–114, January 1976.

Ivanov, A. O. and Tuzhilin, A. A., "The steiner ratio gilbert-pollak conjecture is still open," *Algorithmica*, vol. 62, pp. 630–632, 2012.

Melzak, Z. A., "On the problem of steiner," *Canadian Mathematical Bulletin*, vol. 4, pp. 143–148, March 1961.

Prmel, H. J. and Steger, A., *The Steiner Tree Problem.* Berlin, Germany: Vieweg, 2002.

Rubinstein, J. H., Thomas, D. A., and Wormald, N., "Steiner trees for terminals constrained to curves," *SIAM Journal on Discrete Mathematics*, vol. 10, pp. 1–17, 1997.

Weng, J. F., *Steiner trees, coordinate systems, and NP-Hardness.* Springer, Boston, MA: Combinatorial Optimization, 2000, vol. 6, pp. 63–80.