The Fermat Problem

Stefanie Streck

Pacific Lutheran University

Advisors: Amy Shell-Gellasch and Jessica Sklar

Email: strecksl@gmail.com

8751 Town and Country Blvd. Apt. F, Ellicott City, MD 21043

Abstract

The Fermat Problem was published in the early 1600s by Pierre de Fermat but was not studied or applied substantially until the twentieth century. We trace the history of the problem's development over the last four hundred years, addressing its initial use as a geometric exercise and later its progression due to computers and its applications to networks.

1 The Fermat Problem and Trees

In the early 1600s, Pierre de Fermat (1601-1665) made the following proposition at the end of an essay on maxima and minima, *Method for determining Maxima and Minima and Tangents to Curved Lines* [16]: "Let he who does not approve of my method attempt the solution of the following problem: *Given three points in a plane, find a fourth point such that the sum of its distances to the three given points is a minimum!*" [7]

In order to discuss the *Fermat Problem* there are some basic definitions and explanations about trees to consider. For instance, a **tree** is defined as follows:

Definition 1 *A **tree** is a connected graph that has no cycles. [10]*

This means that any two distinct vertices within a graph, G, are connected by a unique path, and the paths do not create any cycles within G. (See Figure 1)

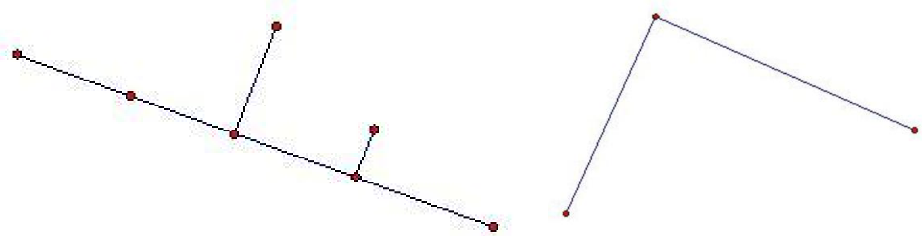

Figure 1: Tree, and Minimal Spanning Tree of Three Points

Definition 2 *Let G be a connected graph. A **spanning tree** in G is a subgraph of G that includes every vertex of G and is also a tree. [10]*

Definition 3 *A **minimal spanning tree** for a graph G is the spanning tree for G that has the shortest length.*

So, a minimal spanning tree for any three points in the plane is simply the tree that is made by connecting the points using the two shortest edges. (See Figure 1)

The objective of the problem posed by Fermat is to find a tree smaller than the minimal spanning tree for three points given in the plane. To do this, the goal is to find another point in the plane such that the tree connecting each of the original points to this point will be minimal. So in adding another point, our tree should contain three line segments whose combined distances between the points is shorter than (or equal to) that of the minimal spanning tree for the three points.

2 History of the Fermat Problem

Evangelista Torricelli (1608-1647) was the first to give a geometric solution to this problem and did so around 1640 however, his solution was not published until 1919 as part of a compilation of his works in *Opere di Evangelista Torriceilli* [14]. His construction (see Figure 2) can be made using a few simple steps:

1. Given a triangle formed by points A, B, and C, construct equilateral triangles ACE, ABG, and BCF so that each contains one of the edges from triangle ABC.
2. Circumscribe each equilateral triangle. [11]

The point of intersection of these three circles is the point that yields the minimal path between points A, B, and C and it has come to be called the Torricelli Point which will be denoted by P.

Soon after, Bonaventura Francesco Cavalieri (1598-1647) found that the line segments from the original three points to the Torricelli Point form angles of 120° with each other. That is, the edges AB, AC, and BC subtend angles of 120°. Cavalieri is said to have worked out his solution independently of Torricelli's findings and he published his solution in his book *Excercitationes Geometricae* [13] right before his death in 1647.

Thomas Simpson (1710-1761) contributed an alternate geometric construction in 1750. He found that it was possible to find the Torricelli Point by constructing the equilateral triangles

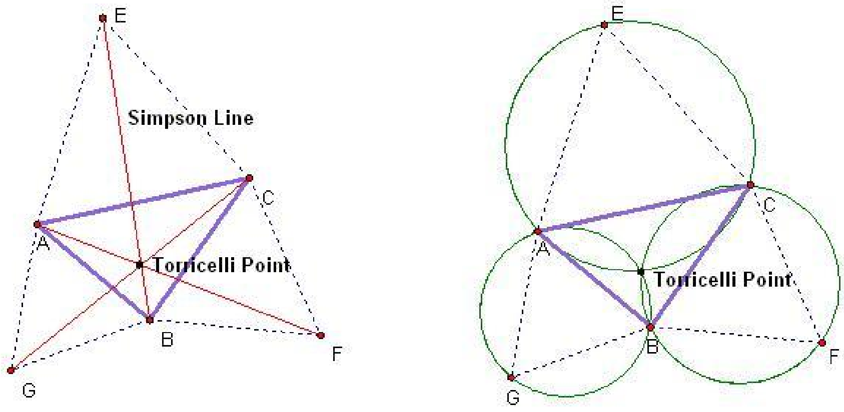

Figure 2: Simpson's Proof and Torricelli's Proof

and then drawing a line segment from each of the original vertices to the opposite vertex of the opposite equilateral triangle (see Figure 2). These lines have come to be called Simpson Lines in his honor. He included the problem as an exercise in his book, *The Doctrine and Application of Fluxions* [15], drawing the attention of more prominent mathematicians, including the Swiss Mathematician Jakob Steiner.

The problem resurfaced in 1834 when Franz Heinen discovered that the Simpson Lines are all of equal length and that the length of a Simpson Line is equal to the length of the minimal path created by finding the Torricelli Point. That is, $AP + BP + CP = EB = FA = GC$. He also noted that if a vertex of the original three points has edges creating an angle greater than or equal to 120° that vertex becomes the minimizing point for the tree, making the shortest network the minimal spanning tree of the original three points. Since Heinen was the first to address this "trivial case" of the Fermat Problem in which the shortest path between the points is actually the minimal spanning tree, he is credited as the first person to solve Fermat's Problem in its entirety.

Very little significant work was published concerning the Fermat Problem between 1834-1941. The majority of research on the problem during the twentieth century was prompted by

Richard Courant and Herbert Robbins after they published their book *What is Mathematics?* [8] in 1941. It was after this publication that the problem became known as the Steiner Problem, because Courant and Robbins present the problem as a special case of a general problem that Steiner had worked on in his studies of combinatorial optimization. Due to the popularity of their book, this name has been applied to the Fermat Problem ever since, the tree creating the shortest path between the original vertices has come to be called the Steiner Tree, and the Torricelli Points are now referred to as Steiner Points. In their book, Courant and Robbins give a rigorous geometric proof of how to find the Torricelli Point for the Fermat Problem (the case in which the number of points given in the plane initially is n=3) and they also prove the findings of Heinen. That is, that P is the point from which each of the three sides of triangle ABC subtends an angle of 120°. The proof will be presented in this paper in Section 3 with several details added that the authors proved in different sections of their book or that they felt were clear enough without further explication.

3 Proof of a Solution to the Fermat Problem: Courant and Robbins

The problem and answer to the Fermat Problem as posed by Courant and Robbins are as follows:

Problem: *Three points A, B, C are given in a plane, and a fourth point, P, in the plane is sought so that the sum AP+BP+CP is minimal, where AP, BP, and CP denote the three distances from P to A, B, C respectively.*

Answer: *If in the Triangle ABC all angles are less than* $120°$*, then P is the point from which each of the three sides,* $\overline{AB}$*,* $\overline{BC}$*,* $\overline{CA}$*, subtends an angle of* $120°$*. If, however, an angle of triangle ABC (for example, the angle at C) is equal to or larger than* $120°$*, then the point P coincides with the vertex of that angle (C).* ([8] pg. 354)

Proof:

Case 1: P coincides with one of the vertices A, B, or C. In this case, clearly P must be the vertex of the largest angle of triangle ABC because the side opposite the largest angle must have the greatest length of all the triangle's sides. Thus, the minimal spanning tree yields the shortest path.

Case 2: P differs from A, B, C. For this case, we must prove that P is the point from which each of the three sides, AB, BC, CA, subtends an angle of 120°.

Let K be the circle about one of the vertices, say C, having radius of length c. We want to construct the point P on K such that $AP + PB$ will be minimal. (See Figure 3)

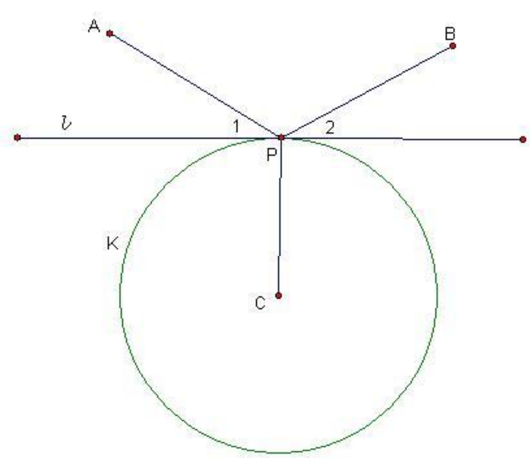

Figure 3: Courant and Robbins' Proof

To find the location of point P on K Heron's Theorem for the Extremum Property of light rays will be used. Heron's Problem states: "Given a line l and two points, A and B, on the same side of l, for what point, P, on l is $AP + PB$ the shortest path from A to l to B?"

([8] pg. 330) Basically, the goal is to find out where to place a tangent line on K so that the point of tangency will be P with the properties sought. First, reflect A across l to obtain point A' (see Figure 4). It is clear that the shortest distance from A' to l to B is the straight line $A'B = A'P + PB$. The point of intersection of $\overline{A'B}$ with l is P since we reflected A across l, making $A'P = AP$ and since $A'B$ gives the shortest distance from A' to l to B then we have

$$A'P + PB = AP + PB$$

Also, note that m$\angle$1=m$\angle$3 since we reflected A across l. Since $\overline{A'B}$ is a straight line, this means that $\angle 2$ and $\angle 3$ are vertical angles and thus have the same measure. So by transitivity, m$\angle$1=m$\angle$2. To prove uniqueness of P, let P' be any point on l such that $P \neq P'$. Then we have that

$$AP = A'P$$

and

$$AP' = A'P'$$

so

$$AP + PB = A'P + PB = A'B$$

and

$$AP' + P'B = A'P' + P'B$$

However, $A'P' + P'B > A'B$ since the sum of the lengths of two sides of a triangle is greater than the length of the third side. That is,

$$A'P' + P'B > AP + PB$$

Thus P is unique.

So we have proven that P must be a point on the circle K such that the tangent line, l, to K through P will have the property that the measure of the angle created by segment $\overline{AP}$ and l will be equal to the measure of the angle created by segment $\overline{BP}$ and l. The result is that

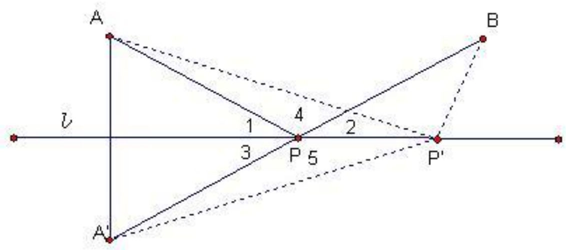

Figure 4: Heron's Proof

the distance $AP + BP$ is minimal. Now note that we said previously that the radius of K is some constant c. The same construction as in figure 3 could be done to find points with the same properties as P for circles around either of the other two vertices, A or B. Since the radius is an arbitrary constant for each circle, then we have really found the line on which P will exist. That is, the radius $\overline{CP}$ can be extended as the line that is perpendicular to l (since l is tangent to K) and P will be some point on this line. The exact P will thus be the point of intersection of the lines extended from each of the radii that can be constructed for a circle around A or B on which this point P has been found using Heron's Theorem.

To use Heron's Theorem we had to assume that both A and B were on the same side of l, in our case not inside of K. Now suppose that one of the points A or B is not outside of K but instead is on or inside of K. Without loss of generality, let this point be A. Then since $P \neq A$ and $P \neq B$ we have that $PA + PB > AB$ since the sum of the lengths of two sides of a triangle is greater than the length of the third side. But $PC \geq AC$ since A lies on or inside of K. So $PA + PB + PC > AB + AC$, which implies that the shortest path would be found if $P = A$ which is a contradiction. The same can be shown for B and therefore both A and B lie outside of K. $\blacksquare$

4 The Steiner Tree Problem After Courant and Robbins

Courant and Robbins prompted interest in the Fermat Problem and its general form among many mathematicians of the twentieth century. This boom was also a result of the advent of computers. The goal of much of the study of this problem after the 1940s was to develop algorithms by which to find solutions for the General Fermat Problem (where any number of points, n, in the plane are given as opposed to just $n = 3$) for as many points given as possible and later to implement computers in doing this. Other research had to do with the finding how much shorter the Steiner Tree was than the Minimal Spanning Tree. The remainder of this paper will be used to focus on several of the achievements of the twentieth century and to give a brief overview on where research on the problem has been directed.

4.1 Hybrid Geometric Construction

After the different geometric constructions presented in the previous sections became more well known, a sort of hybrid geometric construction began to emerge in which the original forms were merged together. One example that became widely implemented can be seen in Figure 5. This construction uses one equilateral triangle, one circle, and one Simpson Line to locate the Torricelli Point. To construct the point, construct an equilateral triangle from one edge of triangle ABC so that the two triangles do not overlap, creating for example triangle AEC. Then, circumscribe the triangle and construct the Simpson Line by inserting the segment from point E to point B. The point at which the circle and the Simpson Line intersect is the Torricelli Point. Melzak made use of this construction in his work, as will be seen in Section 4.2.

4.2 Melzak's Algorithm: 1961

In 1961, Melzak published an article in which he developed an algorithm for finding the Steiner Minimal Tree for a set of n points in the plane in a finite number of steps. He

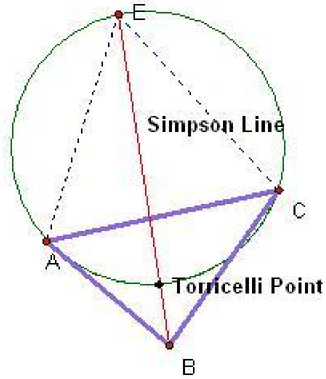

Figure 5: Hybrid Construction

called the problem (S_n) meaning the Steiner Problem for n points given in the plane stating, "Given n points $a_1, ..., a_n$ in the plane, $n \geq 3$, to construct the shortest tree(s) whose vertices contain these n points" ([2] pg. 144). This is the problem known as the Steiner Problem. Melzak gave several more definitions that are necessary for one to follow his work so these shall follow ([2] pg. 144):

Definition 4 *Tree: Given a set of N points $b_1, ..., b_N$ in the plane, a tree U on the vertices $b_1, ..., b_N$ is any set consisting of some of the closed straight segments b_ib_j such that any two vertices can be joined by a sequence of segments belonging to U in only one way (acyclic).*

Definition 5 *Branch: A segment b_ib_j in U.*

Definition 6 *Length of Tree U: L(U) is the sum of the lengths of its branches.*

Definition 7 *$\{b_i\}$: the set of all vertices sending branches to the vertex b_i. (see Figure 6)*

Definition 8 *$w(b_i)$: the weight of a vertex b_i is the number of vertices sending branches to vertex b_i. (see Figure 6)*

Melzak proved the following theorem:

Theorem 1 *For every $n \in \mathbb{Z}^+$, $n \geq 2$, there exists a finite sequence of Euclidean constructions yielding all the minimizing trees of the problem (S_n). [2]*

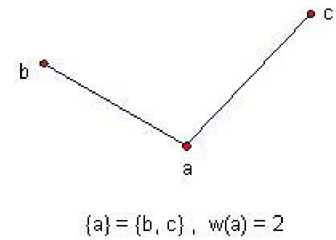

Figure 6: Branches and Weight of Vertices

Note that a Euclidean construction is defined as a construction that uses only a straightedge and compass. Basically, this theorem states that there exist a finite number of Euclidean constructions that will yield all possible minimizing trees of the Steiner Problem for n points in the plane. Melzak gives six characteristics for any minimizing tree, T, of (S_n), although the wording and symbols used here may differ from his in order to make his work easier to follow ([2] pg. 146):

Let T be a minimizing tree of (S_n). Then,

1. Tree T is made up of the set of the original n vertices, A=$\{a_1,...,a_n\}$, plus the set of Steiner Points, S_k^n=$\{s_1,...,s_k\}$ as well as the edges connecting them.

2. Tree T is acyclic (non-selfintersecting).

3. The weight of each Steiner Point, $w(s_i)$, is 3 for 1≤i≤k.

4. Each Steiner Point, s_i, 1≤i≤k, is the Steiner Point of the triangle $\{s_i\}$ (the triangle made of the set of all vertices sending branches to the point s_i).

5. The weight of each of the original vertices, $w(a_j)$, is less than or equal to 3 for 1≤j≤n.

6. The number of Steiner Points in T is less than or equal to the number of original vertices minus two, that is, 0≤k≤n-2.

Any tree that satisfies these six properties is a Steiner Tree, or an S-tree, with k Steiner Points, none of which are the same as the original vertices. We call such a tree an S_k-tree if it has $n+k$ vertices. When $k=0$, the only vertices of tree T are the set of the original vertices, A. Conditions (2)-(5) are evident from (S_3) and (6) is not going to be proven here. The set of vertices of an S_k-tree not including those in A is called an S_k^n-set which basically means the set of Steiner Points in an S_k-tree. For example, for the Fermat Problem, S_1^3 is the set of one Steiner Point for the three original vertices.

Lemma 1 *The number N(n,k) of S_k^n-sets is finite for every n and k, and every such set can be obtained by a Euclidean Construction. [2]*

In short, the lemma states that the number of different sets of Steiner Points in an S_k-tree is finite and every set can be found via Euclidean construction.

Proof by Induction ([2] pg.146-47):

First, N(n, 0)=1 (the empty set), by definition. Now, let k=1. Then any S_1^n-set has a single point, s_1, that by (4) is a Steiner Point of some triangle whose vertices are in A. Therefore, $N(n,1) \leq \begin{pmatrix} n \\ 3 \end{pmatrix}$ and each S_1^n-set can be found through Euclidean Construction of the problem (S_3).

Now, suppose the lemma is true for $k = 1, ..., K$ where $K \leq n-3$ for every n. Consider a particular set of Steiner Points S_{K+1}^n-set, Y. There must exist some point s in Y such that s includes at least two points of A, call them a_1 and a_2. Let b be the third point in s and note that b can either be a point in A or in Y. That is, the third point b could be one of the original vertices or a Steiner Point. Using the hybrid construction, let a be the point of intersection of the circle C through the vertices a_1, a_2 and s and the extension of the straight segment $\overline{sb}$ beyond s (see Figure 7).

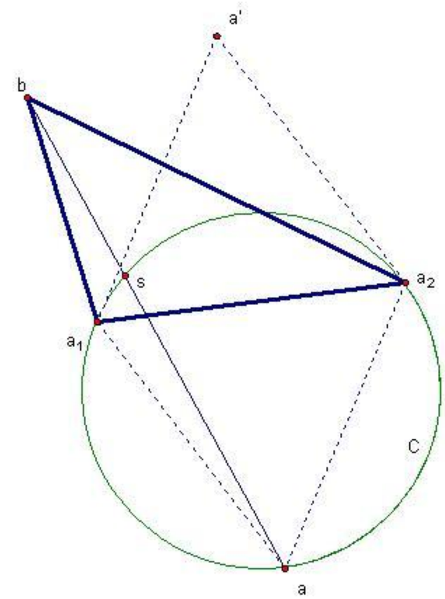

Figure 7: Construction for Lemma 1

Then it follows, as in the hybrid construction for (S_3) that a is the third vertex of the equilateral triangle with two vertices a_1 and a_2. More accurately, it is the third vertex of one of two such triangles (one overlaps the original triangle). Since there are two such triangles and since there are $\begin{pmatrix} n \\ 2 \end{pmatrix}$ ways of selecting a_1 and a_2 in A, it follows that there are $2\begin{pmatrix} n \\ 2 \end{pmatrix} = 2\frac{n!}{(n-2)!2!} = 2\frac{n(n-1)}{2} = n(n-1)$ possibilities for a. So a can be found by Euclidean constructions based on the set A and $N(n, K+1) \leq n(n-1$. So it is evident that the K members of Y, other than s, form an S_K^{n+1}-set for the points $a_1, ..., a_n, a$. Thus $N(n, K+1) \leq n(n-1)N(n+1, K)$. Since $N(n, 1) \leq \begin{pmatrix} n \\ 3 \end{pmatrix}$ then it follows that $N(n, k) \leq \begin{pmatrix} n+k-1 \\ 3 \end{pmatrix} \frac{(n+k-2)!(n+k-3)!}{(n-1)!(n-2)!}$, and each of these $N(n, k)$ sets can be found by Euclidean constructions. ■

The next lemma that Melzak proved was:

Lemma 2 *There is a finite number of S-trees and each one can be obtained by a Euclidean construction.*

Proof ([2] pg. 147-48):

Consider a particular S_k^n-set with vertices $\{s_1, ..., s_k\}$. Combined with set A, with vertices $\{a_1, ..., a_k\}$, we have the set of $n+k$ vertices. It was proven by [19] that for N distinct vertices there are N^{N-2} trees and so for $n+k$ vertices there are $(n+k)^{N+k+2}$ trees. From lemma 1 it is known that there are $N(n,k)$ possibilities for the S_k^n-set so combined there are

$$\sum_{k=0}^{n-2} N(n,k)(n+k)^{n+k-2}$$

trees to examine. However, given any one of these trees it is possible by Euclidean constructions to decide whether it is a Steiner tree or not. ■

The proof of the Theorem follows as a result of these two lemmas. First, construct all of the Steiner trees; by lemma 2 this is possible to do. Euclidean constructions can be done to obtain the minimum length of a finite number of trees and to find the length of a Steiner tree. Since the Steiner trees are the minimizing tree(s) of the problem (S_n) the proof is complete. ■

Melzak realized that his algorithm was not practical due to the increase in the number of sets of Steiner trees as the number of vertices (n) increased. He seemed enthusiastic about the possibility of developing a better algorithm that a computer could use to solve the problem (S_n) but after years of research and development after Melzak's work, it is now evident that this problem is not solvable in real time for large numbers of vertices.

4.3 Gilbert and Pollak's Steiner Ratio Conjecture: 1968

In 1968, Gilbert and Pollak published a paper on Steiner Minimal Trees, called *Steiner Minimal Trees* [5], in which they examined much of the previous work done, as well as made some new inferences. For the purposes of this paper, our focus will be on their Steiner Ratio Conjecture. They conjectured that for any set A of points given in the plane, $L_s(A) \geq (\sqrt[2]{3}/2)L_m(A)$ where $L_s(A)$ and $L_m(A)$ denote the lengths of the Steiner Minimal Tree and the Minimum Spanning Tree of A, respectively. In other words, they conjectured that the ratio of the Steiner Minimal Tree to the Minimum Spanning Tree is always greater than or equal to $\sqrt[2]{3}/2$. They based their conjecture on their previous study of the behavior of trees in the Euclidean Plane and did not submit a proof for their hypothesis but a proof was published in 1990 by Du and Hwang. [3]

It is interesting to address this conjecture because it prompts the question of why they would want to come up with this ratio in the first place. The answer to this question is that the ratio tells us how close of an approximation to the shortest network (i.e., the Steiner Tree) can be obtained through the use of other minimal spanning trees. That is, how accurate of a solution to (S_n) can be obtained using this approximation rather than computing the actual solution.

Why would we want an approximation if we can calculate the actual solution? The answer to this question refers back to Melzak's Algorithm and also to Torricelli's original construction. How efficiently can the Steiner Minimal Tree be obtained for a large set of points in the plane? With the development of the computer, it was found that with the algorithms developed, the problem of finding the Steiner Minimal Tree for a large number of points is an NP-hard Problem. A problem is assigned to the NP (nondeterministic polynomial time) class if it is solvable in polynomial time by a nondeterministic Turing machine (a theoretical computing machine that can run forever) [12] and to be NP-hard means that it is theorized that even a nondeterministic Turing machine couldn't solve it. What this means is that for a set of points that is really large (in 1968, even 30 points may have been too many) a computer running the algorithm for finding the Steiner Minimal Tree can't find a complete solution in real time. As algorithms have improved, more and more solutions are being found but for a large number of points, complete solutions are still unobtainable. So since it is not realistic to use Steiner Trees for most real life scenarios, Gilbert and Pollak's conjecture still made it possible to apply the concept of finding a tree smaller than the Minimal Spanning Tree to optimize networking.

5 Conclusion

The research described here is only a small part of the extensive work that has been done on the Fermat Problem. Further research is recommended for those interested in the development and use of algorithms for computers. This paper focused much more on the geometric aspects of the problem but there is an extensive extension of this problem into Graph Theory. Hwang's book ([6]) is the only compilation that I found that was extensive enough to cover most of the mathematical extensions of the problem and is an excellent resource for further study.

References

[1] M. Bern and R. Graham, *The Shortest-Network Problem*, Scientific American **260** (January 1989), 84-89.

[2] Z. Melzak, *On the Problem of Steiner*, Canadian Mathematical Bulletin **4(2)** (May 1961), 143-148.

[3] D. Du and F. Hwang, *The Steiner Ratio Conjecture of Gilbert and Pollak is True*, Proceedings of the National Academy of Sciences, USA, **87** (December 1990), 9464-9466.

[4] E. Cockayne and D. Hewgill, *Exact Computation of Steiner Minimal Trees in the Plane*, Information Processing Letters **22** (1986), 151-156.

[5] E. Gilbert and H. Pollak, *Steiner Minimal Trees*, SIAM Journal of Applied Mathematics **16(1)** (1968), 1-29.

[6] F. Hwang et al., *The Steiner Tree Problem*, Elsevier Science Publishers, Amsterdam, 1992.

[7] J. Krarup, *On Torricelli's Geometrical Solution to a Problem of Fermat*, IMA Journal of Mathematics Applied in Business and Industry **8** (1997), 215-224.

[8] R. Courant and H. Robbins, *What is Mathematics?*, Oxford University Press, New York, 1961.

[9] R. Balakrishnan and K. Ranganathan, *A Textbook of Graph Theory*, Springer-Verlag, New York, 2000.

[10] Aldous and Wilson, *Graphs and Applications: An Introductory Approach*, The Open University, Berlin, 2000.

[11] http://www-history.mcs.st-and.ac.uk/Biographies/Torricelli.html

[12] http://mathworld.wolfram.com/NP-Problem.html

[13] B. Cavalieri, *Excercitationes Geometricae*, 1647.

[14] E. Torricelli, *Opere di Evangelista Torriceilli*, 1640.

[15] T. Simpson, *The Doctrine and Application of Fluxions*, 1750.

[16] P. Fermat, *Method for determining Maxima and Minima and Tangents to Curved Lines*, 1636.

[17] F. Heinen, 1834.

[18] V. Jarník and O. Kössler, *O minimálních grafech obsahujících n daných bodu, Ĉas*, Pêstování Mat. **63** (1934), 223-235.

[19] J. Riordan, *Introduction to Combinatorial Analysis*, New York, 1959.