Volume 21 (1), pp. 53–61
http://www.orssa.org.za

ORiON
ISSN 0529-191-X
©2005

The rectilinear Steiner ratio

PO de Wet*

Received: 24 December 2004; Revised: 17 March 2005; Accepted: 22 March 2005

Abstract

The rectilinear Steiner ratio was shown to be 3/2 by Hwang [Hwang FK, 1976, *On Steiner minimal trees with rectilinear distance*, SIAM Journal on Applied Mathematics, **30**, pp. 104–114.]. We use continuity and introduce restricted point sets to obtain an alternative, short and self-contained proof of this result.

Key words: Steiner ratio, Steiner tree, spanning tree, rectilinear tree, L_1 norm, NP-complete.

1 Introduction

We consider a set P of n points in the plane and a finite set of arcs which have elements of P as endpoints. (Each arc is a subset of $\mathbb{R}^2$ which is homeomorphic to the closed unit interval $[0, 1]$.) If each arc has two distinct endpoints and if there is at most one arc between two points, then the points and arcs define a graph in a natural way [1]. We call the set of points and arcs a *spanning tree* on P if this graph is connected and acyclic.

We also consider a spanning tree on the set $P \cup S$, where S is also a finite set of points in the plane. We call this a *Steiner tree* on P if all elements of S are endpoints of at least 3 arcs (has degree at least 3). The elements of P are called *terminal points* and those of S are called *Steiner points*.

In what follows we consider all arcs to be finite unions of only vertical and horizontal line segments. Trees consisting of such arcs are called *rectilinear trees*, but for the sake of brevity we shall simply refer to them as trees, spanning trees and Steiner trees and not as rectilinear trees, rectilinear spanning trees and rectilinear Steiner trees. The *length* of a tree is defined as the total length of all its segments. A shortest spanning tree is called a *minimum spanning tree* (MST) and a shortest Steiner tree is called a *Steiner minimal tree* (SMT). In Figure 1 examples of an MST and an SMT for terminal points $\{(-1, 0), (1, 0), (0, -1), (0, 1)\}$ are shown, where we take the x-axis to be horizontal and the y-axis to be vertical. The length of the MST is 6 and that of the SMT is 4. The SMT has one Steiner point, namely $(0, 0)$.

*Department of Decision Sciences, University of South Africa, PO Box 392, Pretoria, 0003, South Africa, email: dwetpo@unisa.ac.za

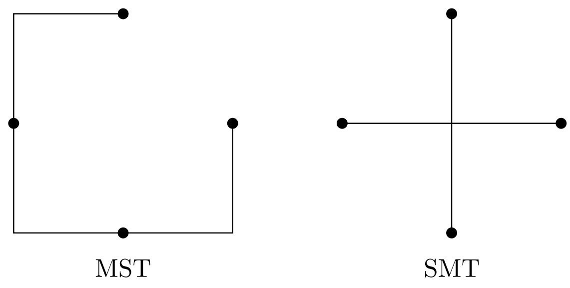


Figure 1

We define the *rectilinear Steiner ratio* ρ_1 as

$$\rho_1 = \max_{P \subset \mathbb{R}^2} \frac{\text{length (MST}(P))}{\text{length (SMT}(P))},$$

where MST(P) is an MST of P and SMT(P) is an SMT of P.

Note that we have elected to define arcs as unions of line segments. A more abstract approach would have been to associate an arc with a pair of points (its endpoints) and to define the length between the points (x_1, y_1) and (x_2, y_2) as $|x_2 - x_1| + |y_2 - y_1|$. The norm used here is known as the L_1 or taxicab norm. Clearly the length of an MST and an SMT for a particular point set P will remain the same if defined in these terms. (In general the L_1 norm for $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$ is defined as $|z| = |z_1| + \ldots + |z_n|$.)

The practical value of studying minimum spanning trees and Steiner minimal trees is obviously related to designing power networks, communication networks and pipelines of minimum cost. In cities, with square blocks, these trees will typically be rectilinear. Another important application is in the design of integrated circuit boards, where shorter networks require less time to charge and discharge, making the circuit board faster. The lines here generally run only vertically and horizontally, thus forming rectilinear trees.

The rectilinear Steiner ratio now gives us an indication of how badly a minimum spanning tree will perform compared to a Steiner minimal tree. In practice a spanning tree may indeed sometimes be used instead of a Steiner tree, because a minimum spanning tree can be constructed in polynomial time [3], whereas no such algorithm is known to exist for Steiner minimal trees. (The associated decision problem is NP-complete [4].)

It is interesting to note in what follows that we restrict ourselves from the beginning to point sets that only have full SMTs. (This concept will be defined in the next section.) It should also be clear from our discussion that an SMT for such a set can be found in polynomial time. Thus the complexity of the problem lies in deciding how to partition the points to allow only full SMTs. There is no efficient method known to achieve this and the number of partitions grows exponentially in n, making it infeasible to consider all possible partitions.

Our main aim will be to prove the following theorem.

Theorem 1 *The rectilinear Steiner ratio is* $3/2$*.* ■

This theorem was originally proved in 1976 by Hwang (1976) by first characterising Steiner trees and then obtaining the Steiner ratio. Richards and Salowe (1991) discussed only the characterisation in 1991 and Salowe (1992) proceeded with the Steiner ratio in 1992. Our proof uses the same two main steps, but we use the notion of continuity to introduce the idea of restricted point sets. This allows us only to have to characterise and consider a dense subset of all possible Steiner trees. Our characterization can then be performed in a more global way and we finally obtain the Steiner ratio by applying Salow's strategy in a complete proof. The result is an alternative, short and self-contained proof of Theorem 1. The value is already achieved in our example (Figure 1). It remains to be shown that ρ_1 will never exceed this value.

2 Restricted point sets

Given $n > 1$ distinct points $(x_1, y_1), \ldots, (x_n, y_n)$ we assume, without loss of generality, that the first point is at the origin, and use the remaining points to define a point $z = (x_2, y_2, \ldots, x_n, y_n) \in \mathbb{R}^{2(n-1)} \backslash \{0\}$. We use the L_1 norm for $\mathbb{R}^{2(n-1)}$ and define

$$f : \mathbb{R}^{2(n-1)} \backslash \{0\} \mapsto \mathbb{R}$$

as the length of an MST of the n points and

$$g : \mathbb{R}^{2(n-1)} \backslash \{0\} \mapsto \mathbb{R}$$

as the length of an SMT of the n points.

Since each terminal point has fewer than n arcs incident to it, we note that $|f(z_2) - f(z_1)| \leq n|z_2 - z_1|$ and $|g(z_2) - g(z_1)| \leq n|z_2 - z_1|$ and thus that f and g are continuous. Since g is never zero we also have that f/g is continuous. Thus, to show that

$$\rho_1 = \max_{z \in \mathbb{R}^{2(n-1)}} \frac{f(z)}{g(z)} \leq \frac{3}{2},$$

it follows (from basic topology [7]) that we only have to consider a dense subset of $\mathbb{R}^{2(n-1)}$ and we consequently assume that no two points have the same x or the same y coordinates. (Note that for any z in an Euclidean space one can find a point arbitrarily close to it such that none of the coordinates of the point are the same.)

If an SMT has a terminal point t with degree larger than one, the removal of t renders connected components Q_i and a partition P_i of the remaining points. If, for all $P_i \cup t$ we have that

$$\text{length } (\text{MST}(P_i \cup t)) \leq \frac{3}{2} \text{ length } (\text{SMT}(P_i \cup t)),$$

then Theorem 1 is proved. The argument remains valid if we remove all terminal points with degree larger than one in this manner. In what follows we need only prove the theorem for point sets for which all SMTs have only terminal points of degree one. Such SMTs are called *full*.

Definition 1 *We say that a terminal point set is a* restricted set *if no two points have the same x or the same y coordinates and if all SMTs on the points are full.* ■

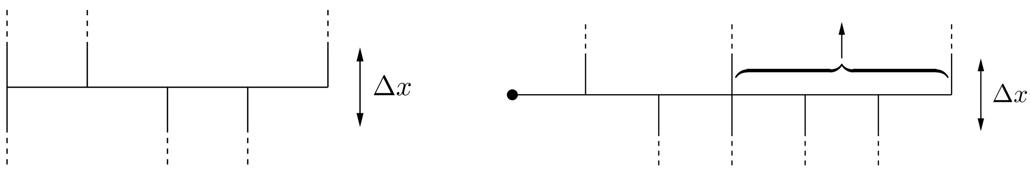
3 Simple SMTs

We collectively refer to terminal points, Steiner points and 90° angles as *points.* A *segment* connects two points without containing any other points. A *line* is a maximal sequence of adjacent collinear segments. In Figure 1 we see that the SMT has 5 points, 4 segments and 2 lines.

Definition 2 *We call an SMT on a restricted set* simple *if each line has a terminal point at one of its endpoints and if the SMT contains no crosses (Steiner points of degree 4).* ■

Lemma 1 *For every restricted point set there exists a simple SMT.*

Proof: First consider an SMT such that the number of horizontal lines that do not end in a terminal point, is a minimum and assume it to be more than zero. Consider the topmost of these lines. Since we have an SMT, this line can be moved up or down by a sufficiently small amount Δx without decreasing (or increasing) the length of the tree (Figure 2(a)).



(a) (b)

Figure 2

We move the line upwards until a terminal point or horizontal line is reached, thus decreasing the number of horizontal lines not ending in terminal points and providing a contradiction. (A terminal point thus reached will have to be at the end of the line not to contradict fullness.) It follows that there is an SMT where each horizontal line has a terminal point as one of its endpoints. Among all such SMTs we may consider one in which the number of vertical lines not ending in terminal points is a minimum. As above, it follows that this number is zero.

Next we consider those SMTs with lines ending in terminal points and the smallest number of crosses and choose one for which the topmost cross is as high as possible (Figure 2(b)).

We move the left or right horizontal arm of the cross (depending on which side does not contain the terminal point) upwards, until we reach another horizontal line or terminal point. (Again the terminal point will have to be at the end of the line.) We still have that all lines end in terminal points. We now either have one fewer cross, or the same number of crosses of which one is higher than before — providing a contradiction. ■

4 Characterisation of a longest simple SMT

Consider a simple SMT (on a given restricted point set) which has a maximum number of points on a line. We call this tree a *longest* simple SMT, the line is called the *trunk* and the subgraphs attached to the trunk are called *branches.* We wish to characterise these trees for $n > 4$. The basic structure is as in Figure 3, where we place the terminal point of the trunk at the top.

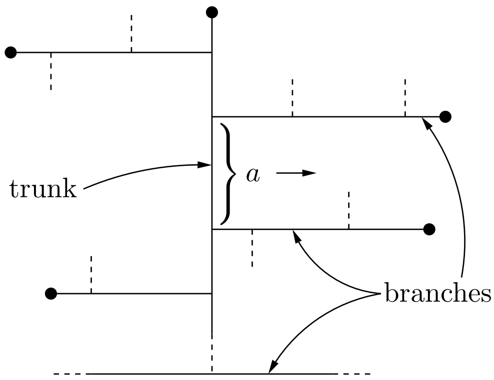


Figure 3

Lemma 2 *Adjacent branches of a longest simple SMT are on opposite sides of the trunk, i.e. branches alternate.*

Proof: If two adjacent branches (neither the lowest) of a longest SMT are on the same side of the trunk, as in Figure 3, a segment a of the trunk can be made to slide between the branches until a terminal point is reached — contradicting fullness. At the lowest branch we may have the situation in Figure 4 (as indicated by the solid lines).

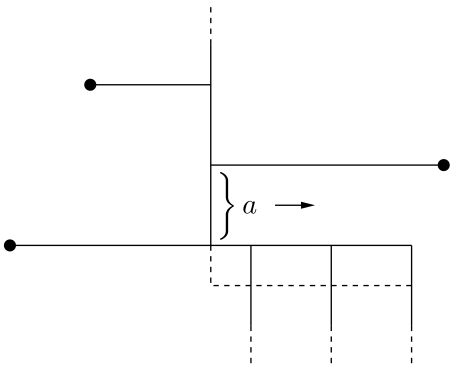
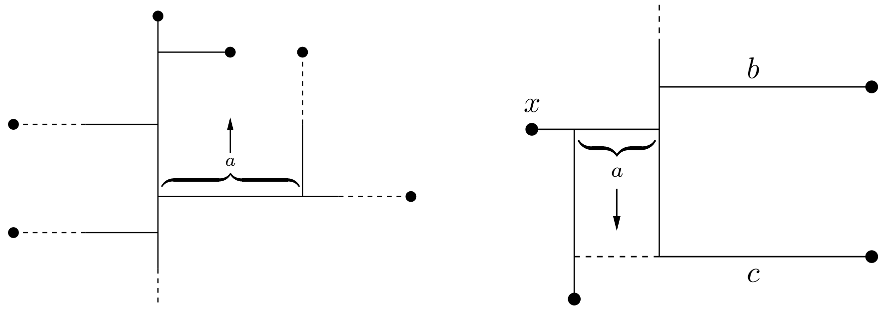


Figure 4

If segment a is slid to reach the terminal point or a vertical segment above the lowest branch, we again contradict our assumptions. (Fullness is contradicted if the terminal point is reached first, and the fact that it is an SMT is contradicted if segment a overlaps another vertical segment.) If this does not happen, we can replace the lowest branch with a new lowest branch (the dotted lower branch in the sketch) to obtain a simple SMT with a longer trunk — contradicting the fact that we had a longest simple SMT. ∎

Lemma 3 *The branches of a longest simple SMT, except possibly for the lowest branch, each consist of only one segment; i.e. branches, except possibly the lowest, have no subbranches.*

Proof: Assume there are subbranches. As before, a segment a (the first segment on the branch) can be slid up or down until a terminal point is reached or another branch is overlapped as in Figure 5(a).



(a) (b)

Figure 5

At the second lowest branch we may, however, reach the end of the trunk. In this case we see (Figure 5(b)) that we have created non-alternating branches b and c, disallowed by Lemma 2.

Note that the branch b must exist, otherwise the line with the most points would have passed through x. ∎

Lemma 4 *The lowest branch of a longest simple SMT, can have at most one subbranch, which consists of a single segment and faces downwards.*

Proof: Assume there are subbranches facing upwards. Consider the first and note that there has to be a single downward going subbranch between it and the trunk to ensure that it is a minimal tree and full (Figure 6).

We replace segments a and b with the parallel dotted segments and let segments c and d slide to obtain a contradiction of fullness. The fact that there is only one downward subbranch and that it consists of a single segment is easily verified by sliding of segments. ∎

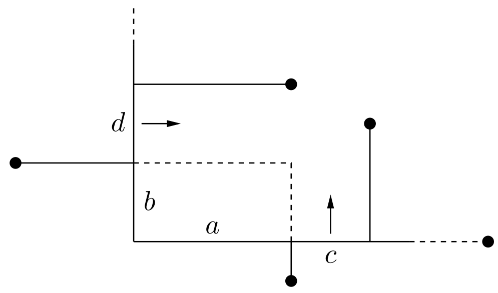


Figure 6

5 The Steiner ratio

According to Lemmas 2–4, a longest simple SMT with $n > 4$ will have the general structure shown in Figure 7, where the subbranch at the bottom may or may not be present.

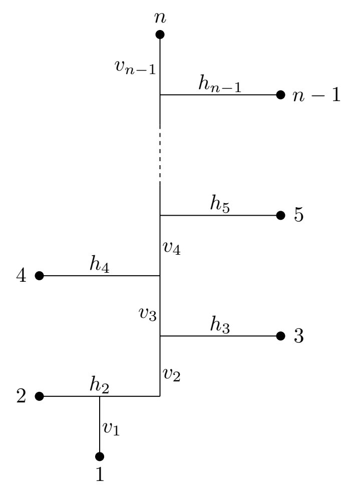


Figure 7

We can now prove that the rectilinear Steiner ratio satisfies $\rho_1 \leq 3/2$.

Proof: We use induction and assume that the theorem is true for all sets with fewer than $n > 4$ terminal points. We only have to consider a longest simple SMT. We first assume that there is a subbranch and number the terminal points and segments as in Figure 7, taking h_n to be zero.

If $h_4 < h_2$, as in Figure 8(a), we consider the rectangle with perimeter $2(h_2+h_3+v_1+v_2+v_3)$ and note that a path W lies on it which connects the terminal points such that

$$\text{length } (W) \leq \frac{3}{4}\ (2(h_2 + h_3 + v_1 + v_2 + v_3)).$$

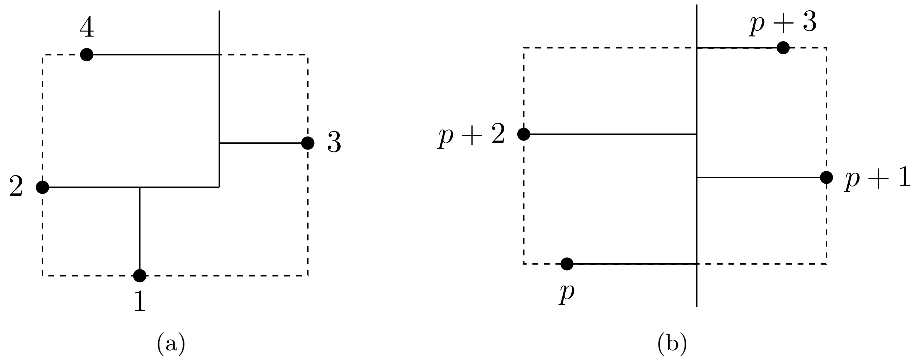


Figure 8

We can now combine W and $\mathrm{MST}(4,\ldots,n)$ to obtain the required MST:

$$\begin{aligned}
\text{length } (\mathrm{MST}(1,\ldots,n)) &\leq \text{length } (W) + \text{length } (\mathrm{MST}(4,\ldots,n)) \\
&\leq \frac{3}{2}\,(h_2 + h_3 + v_1 + v_2 + v_3) + \frac{3}{2}\left(\sum_{i=4}^{n} h_i + \sum_{i=4}^{n-1} v_i\right) \\
&= \frac{3}{2}\text{ length } (\mathrm{SMT}(1,\ldots,n)).
\end{aligned}$$

If $h_4 > h_2$ we find the first p such that $h_{p+3} < h_{p+1}$, as in Figure 8(b). (This must eventually happen, because $h_n = 0$.)

Again we find a path W such that

$$\text{length } (W) \leq \frac{3}{2}\,(h_{p+1} + h_{p+2} + v_p + v_{p+1} + v_{p+2})$$

and combine $\mathrm{MST}(1,\ldots,p)$, W and $\mathrm{MST}(p+3,\ldots,n)$ to obtain the required MST.

We have now shown that the theorem is true for a longest simple SMT on n points with a subbranch, under the hypothesis that the theorem is generally true for fewer than n points. We can repeat the proof for $n+1$ points under the same hypothesis, because we remove 4 points at a time in the proof. By noting that a longest simple SMT without a subbranch on n points can be seen as a longest simple SMT with a subbranch on $n+1$ points, where points 1 and 2 overlap, we see that the theorem is also true for a longest simple SMT on n points without a subbranch. To show that the theorem is true for $n \leq 4$ is straightforward, again using a path W. We omit the detail. ■

References

[1] Diestel R, 1997, *Graph theory*, Springer-Verlag, New York (NY).

[2] Hwang FK, 1976, *On Steiner minimal trees with rectilinear distance*, SIAM Journal on Applied Mathematics, **30**, pp. 104–114.

[3] Papadimitriou CH & Steiglitz K, 1982, *Combinatorial optimization: Algorithms and complexity*, Prentice-Hall, Englewood Cliffs (NJ).

[4] Prömel HJ, Steger A, 2002, *The Steiner tree problem: A tour through graphs, algorithms, and complexity*, Vieweg, Braunschweig.

[5] Richards DS & Salowe JS, 1991, *A simple proof of Hwang's theorem for rectilinear Steiner minimal trees*, Annals of Operations Research, **33**, pp. 549–556.

[6] Salowe JS, 1992, *A simple proof of the planar rectilinear Steiner ratio*, Operations Research Letters, **12**, pp. 271–274.

[7] Sieradski AJ, 1992, *An introduction to topology and homotopy*, PWS-Kent Publishing Company, Boston (MA).