J. SIAM Appl. Math.
Vol. 14, No. 2, March, 1966
Printed in U.S.A.

ON STEINER'S PROBLEM WITH RECTILINEAR DISTANCE*

M. HANAN†

1. Introduction. This paper is concerned with the following type of problem. Given n cities, construct a network of roads of *minimum total length* so that a traveler can get from one city to any other. Roads may cross each other outside of the city limits and these points are called junction points. (Roads which cross within the cities, however, will not be referred to as junctions.) It is assumed that junction points add no extra cost to the construction of the network so that there may be as many as necessary to minimize the total length. Usually, the roads are straight-line connections and the distance between two points is the Euclidean distance. In this paper, however, the rectilinear distance is used. The rectilinear distance $d(p_1, p_2)$ between two points p_1 and p_2 is defined as

$$d(p_1, p_2) = |x_1 - x_2| + |y_1 - y_2|,$$

where (x_i, y_i) are the coordinates of p_i.

Rectilinear distance has application in printed circuit technology where n electrically common points must be connected with the shortest possible length of wire and the wires must run in the horizontal and vertical directions. The junction points of the wires are analogous to the above-mentioned road junctions.

Actually this is a well-known problem due to Steiner (cf. [2] and [6]) and is now formally stated.

Steiner's Problem. *Given n points in the plane find the shortest tree(s) whose vertices contain these n points.*

A tree with m vertices is a connected graph with $m - 1$ edges. (For graph-theoretic terminology see Berge [1].) Several necessary conditions about the solution of this problem are known when the distance between two points is taken to be the Euclidean distance. In this paper several necessary conditions are given for any n, using rectilinear distance. Some of these conditions are analogous to the problem with Euclidean distance, some hold only for rectilinear distance, and some are invariant with respect to the metric. Exact solutions are constructed for $n \leqq 5$.

Since rectilinear distance is not invariant with respect to rotations in the plane, the statement of Steiner's problem must be properly interpreted. Hence it is assumed throughout this paper that when n points in the plane

* Received by the editors March 15, 1965.

† Thomas J. Watson Research Center, International Business Machines Corporation, Yorktown Heights, New York.

are given, a Cartesian coordinate system is also given and the rectilinear distance is defined with respect to this coordinate system.

We now state two problems which are related to Steiner's problem and whose solutions we will have occasion to use in this paper. To distinguish these we refer to Steiner's problem as S_n and we now define P_n and T_n.

P_n : *Given n points* $(p_1, \cdots, p_n)$ *in the plane, find a point q such that the sum of the distances from q to* p_i, $i = 1, \cdots, n$, *is a minimum.*

T_n : *Given n points in the plane, find the shortest tree whose vertices are these n points.*

(We have departed slightly from the notation used by Melzak [6].) The P_n problem has been solved for both Euclidean distance (cf. [7]) and rectilinear distance (cf. [3]). The T_n problem has also been solved (cf. [5] and [8]) and the method of solution is independent of the metric used.

2. Steiner's problem with three points.

2a. Euclidean distance. Given three points in the plane, let T be the triangle whose vertices are these three points. If every angle of T is less than 120°, then the point q of P_3 lies inside T and the lines from p_i to q, $i = 1, 2, 3$, meet at 120° at q. If an angle of T is greater than or equal to 120°, then q coincides with that vertex. (See [4] for a proof and a construction of the solution.)[1] It is not difficult to see that P_3 yields the same solution as S_3. Also, if an angle of T is greater or equal to 120°, then the solution of S_3 is identical to the solution of T_3. See Fig. 1.

2b. Rectilinear distance. Using rectilinear distance, the solution to S_3 (or equivalently P_3) is simpler to construct than the corresponding problem using Euclidean distance. In place of the triangle T, we consider the *enclosing rectangle R* which we now define, in general, for n points.

DEFINITION 1. Given n points in the plane the enclosing rectangle R is the smallest rectangle whose sides are parallel to the x and y axes and which includes the n points either within or on its boundary.

We refer to the solution of the three-point problem throughout this paper and therefore state the result as a separate theorem.

THEOREM 1. *Let* (x_i, y_i) *be the coordinates of the given points* p_i, $i = 1, 2, 3$. *The q-point of* P_3 *is located at* (x_m, y_m) *where* x_m *and* y_m *are the medians of* $\{x_i\}$ *and* $\{y_i\}$, *respectively.*

As stated earlier, this special case of P_n is solved in [3]. The following theorem relates P_3, S_3 and T_3. Let d_{S_n}, d_{P_n}, and d_{T_n} be the total (rectilinear) distance in the solutions of S_n, P_n, and T_n, respectively.

---

[1] We would like to thank the referee for suggesting two other references to this problem: E. GOURSAT, *A Course in Mathematical Analysis*, vol. 1, Dover, New York, 1959, p. 130; H. S. M. COXETER, *Introduction to Geometry*, John Wiley, New York, 1961, p. 21.

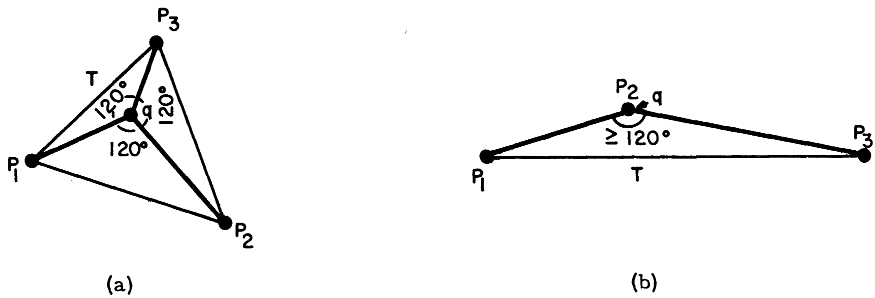

FIG. 1

THEOREM 2. *The solutions of S_3 and P_3 are identical; in fact,*

$$d_{S_3} = d_{P_3} = \tfrac{1}{2}P(R) \leqq d_{T_3}, \tag{1}$$

where $P(R)$ is the perimeter of the enclosing rectangle R. The equality sign holds in (1) *only if q is coincident with some p_i, i.e., $(x_m, y_m) = (x_i, y_i)$ for some $i = 1, 2, 3$.*

The proof of Theorem 2 is straightforward. It follows from Theorem 1 and the fact (which we prove later for S_n, in general) that the minimum tree solution to S_3 can have either zero or one additional vertex. See Fig. 2.

3. Necessary conditions on a solution to S_n. We use the following notation: p_i are the given n points and q_i, $i = 1, \cdots, k$, are the additional k vertices in the solution G of S_n. When we are referring to the vertices of G, we speak of p-vertices or q-vertices. When we are referring to the location of these vertices in the coordinate system we speak of p-points or q-points. We use the notation p_i (or q_i) interchangeably for a vertex of G or the location of that vertex. Its meaning should be clear from the context. We let P be the set of p-points or p-vertices and Q be the set of q-points or q-vertices. When we speak of a vertex a_i, $i = 1, \cdots, n + k$, we mean either p_i or q_i. We let $w(a_i)$ be the local degree of the vertex a_i, that is, the number of vertices adjacent to a_i and $C(a_i)$ be this set of vertices. (Two vertices are adjacent if they have an edge in common.) The following essentially sums up the present knowledge about the solution to S_n using Euclidean distance (cf. [2] and [6]).

(1) $w(q_i) = 3$, $1 \leqq i \leqq k$,
(2) $w(p_i) \leqq 3$, $1 \leqq i \leqq n$,
(3) $0 \leqq k \leqq n - 2$,
(4) each q_i, $1 \leqq i \leqq k$, is the q-point of $C(q_i)$.

These conditions are easy to prove. In fact (4) can be replaced by the stronger statement that every connected subtree of a solution G of S_n is a minimum tree of those $m \leqq n$ points.

FIG. 2

The analogous necessary conditions on a solution G of S_n using rectilinear distance are:

(1) $w(q_i) = 3$ or 4, $1 \leqq i \leqq k$,
(2) $1 \leqq w(p_i) \leqq 4$, $1 \leqq i \leqq n$,
(3) $0 \leqq k \leqq n - 2$.

Conditions (1) and (2) are almost obvious. In fact if $w(a_i) = 4$, then two pairs of vertices of $C(a_i)$ must be collinear and a_i is at the intersection of the straight lines connecting those pairs. See Fig. 3.

To prove the inequality on the right side of (3), assume that there are k q-vertices in G and find the least number of p-vertices possible. Assume the worst case, that is, $w(q_i) = 3$ and the q-vertices form a subtree with $k - 1$ edges. Since each edge counts twice in the total degree of the q-vertices,

$$n \geqq 3k - 2(k - 1) = k + 2,$$

or

$$k \leqq n - 2.$$

To show that zero is a true lower bound, we can easily construct an example where $k = 0$. Since $\frac{1}{2}P(R)$ is a lower bound for d_{S_n} and, in the example shown in Fig. 4, $d_{S_n} = \frac{1}{2}P(R)$, we have found a minimum tree with $k = 0$.

We state condition (4) as a separate lemma for future reference.

LEMMA 1. *Given n points in the plane, let G be a solution of S_n. If G' is a connected subgraph of G with m vertices, then G' is a solution to S_m.*

The following theorem has an analog in Euclidean geometry where the triangle T replaces the rectangle R. However, we have not seen it stated in the literature.

THEOREM 3. *If q is any q-vertex of G with degree three, then q can be the only vertex of G inside the enclosing rectangle R of $C(q)$.*

We note first that there may be vertices (including q itself) on the boundary of R. To prove the theorem, assume the contrary, that is, assume that there is another vertex a_m inside R. There must exist some path from

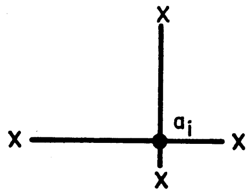

FIG. 3

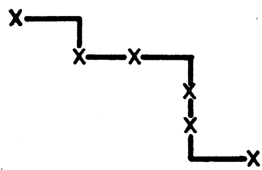

FIG. 4

a_m to one of the vertices $a_i \in C(q)$, say a_1. Now consider the problem S_3 consisting of the three points $\{a_m, a_2, a_3\}$ and let R_1 be the enclosing rectangle of these three points. Since a_m lies inside R, $\frac{1}{2}P(R_1) < \frac{1}{2}P(R)$, so that we can replace the subtree on the vertices (a_1, a_2, a_3, q) with the new subtree on (a_m, a_2, a_3, q_1), where q_1 is the new q-point. This subtree is connected to the rest of the graph by the path from a_m to a_1. Hence we have found a new graph G_1 with a smaller total distance than G, contradicting the hypothesis that G is a solution of S_n.

In general, a solution G to S_n is not unique, that is, there is more than one set Q which yields a minimum tree. Let $N(n, k)$ be the number of sets Q. The main result of Melzak [6] is that $N(n, k)$ is finite for all n and k in Euclidean geometry and there exists a finite sequence of Euclidean constructions yielding all minimizing trees of the problem S_n. In rectilinear geometry this is not true. (We will give an example in §4 where $N(4, 2)$ is infinite.) Hence we cannot guarantee that we can find (by construction) all solutions G to S_n. However, we now prove a theorem which does guarantee finding a finite subset of solutions. The theorem proves, in effect, that there always exists a solution G such that all the vertices in the set Q are located at a predetermined finite set of possible locations.

THEOREM 4. *Let $\{x_p\}$ and $\{y_p\}$ be the sets of x and y coordinates of the given n p-points. If (x_{q_j}, y_{q_j}) are the coordinates of any vertex $q_j \in Q$, then there exists a solution G to the problem S_n such that $x_{q_j} \in \{x_p\}$ and $y_{q_j} \in \{y_p\}$ for all $j = 1, 2, \cdots, k \leqq n - 2$.*

If straight lines are drawn parallel to the x and y axes through all the given points, a grid is imposed on the plane. Theorem 4 states that there exists a solution G such that all the q-vertices are on the intersections of these grid lines. Let I be the set of intersection points or, when referring to the vertices, the set of vertices of G which are located at these intersections. By definition, $P \subset I$. For any $q \in Q$, if $C(q)$ contains only p-vertices then $q \in I$. This last statement is an immediate consequence of Theorem 1 and Lemma 1 when $w(q) = 3$ and is obvious when $w(q) = 4$.

To facilitate the proof of Theorem 4, we first prove two lemmas.

LEMMA 2. *Let G be any solution of S_n and let q_j be any vertex in Q such that $C(q_j)$ contains two vertices in the set I, say i_1 and i_2. If $q_j \notin I$, then a solution G' can be obtained from G such that there is no vertex of G' located at the point q_j and if a vertex q_j' of G' is connected to both i_1 and i_2 then $q_j \in I'$.*

Lemma 2 states, in effect, that given a tree G, certain of the q-vertices can always be "moved" to new locations which are at the intersections of the grid. We first note that if $w(q_j) = 4$ then clearly $q_j \in I$, so we assume that $w(q_j) = 3$. Let i_1, i_2, and a be the three vertices of $C(q_j)$. If $a \in I$ then $q_j \in I$ (Theorem 2 and Lemma 1) so that $a \in Q$. Hence let us call this vertex q_1. The locations of these vertices with respect to q_j must be essentially as shown in Fig. 5. (We have drawn the connection from q_j to i_2 in the way shown for future use.)

By Lemma 1 and Theorem 1, at least one of the vertices i_1 or i_2 must be on the horizontal line through the point q_j. (Clearly the figure can be rotated through an angle of $m\pi/2$, $m = 1, 2, 3$. There is, of course, no loss of generality in assuming this configuration.) Assuming that i_1 is the vertex on this horizontal line, then i_2 can be anywhere in the quadrant $x > x_{q_j}$ and $y \geqq y_{q_j}$. Again, by Lemma 1 and Theorem 1, at least one vertex of $C(q_1)$, say a_1, must lie on the line $y = y_{q_1}$. Without loss of generality, assume that a_1 is to the right of q_1. There are now two possibilities which must be considered: (i) a_1 is to the right of i_2, and (ii) a_1 is in the interval between q_j and i_2.

Considering (i) first, the line joining q_j to q_1 can be moved parallel to itself to the line $x = x_{i_2}$ as indicated in Fig. 6. By making this move, the new graph G' is also a connected tree and its length is the same as G. Hence G' is also a solution to S_n. Clearly $q_j' \in I$ and there is no vertex of

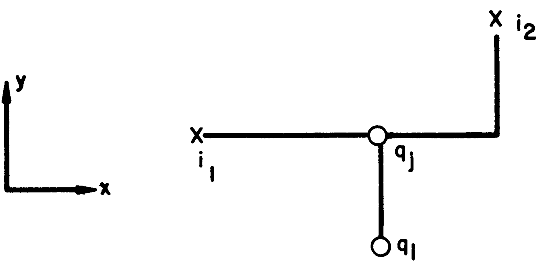

FIG. 5

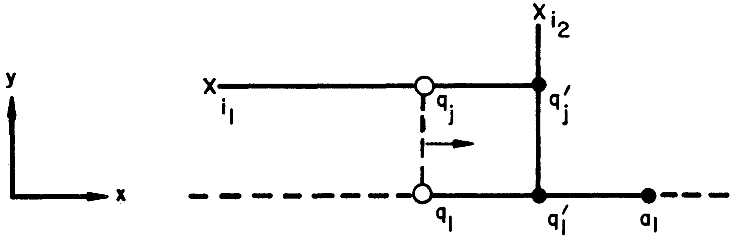

FIG. 6

G' located at the point q_j. The graph G' may have more, fewer, or the same number of vertices as the original graph G. For example, if i_2 were on the line $y = y_{q_j} = y_{i_1}$, then no q-vertex of G' would be generated at the point designated q_j' since this point would be occupied by i_2. Also if $w(q_i) = 3$ then G' has no vertex at q_1 and if $w(q_1) = 4$ then G' has vertices both at q_1 and q_1'. This completes the proof of Lemma 2 for the case (i).

We now examine case (ii). First, if $a_1 \in I$, we move the line joining q_j to q_1 parallel to itself to the line $x = x_{a_1}$ so that $q_j' \in I$ and there is no vertex of G' at the point q_j. Now assume $a_1 \in Q$ and $a_1 \notin I$. It is not difficult to see that no vertex in $C(a_1)$ can be in the region $y > y_{a_1}$. For, referring to Fig. 7, this implies that the subtree connecting this vertex to $\{q_j, q_1, a_1\}$ is not minimum, contradicting Lemma 1. Hence one vertex of $C(a_1)$, say a_2, must be on the line $y = y_{a_1}$ and to the right of a_1. We can now use the same arguments as above, with a_2 replacing a_1, to find another vertex a_3 on the line $y = y_{a_1} = y_{a_2}$ and to the right of a_2. Continuing this argument, the process must eventually end. Either $a_l \in P$ or a_l is to the right of i_2. In either case we can move the line joining q_j to q_1 such that a tree G' is generated with a vertex $q_j' \in I$ and no vertex at the point q_j. (Actually we can prove a stronger result, that is, $l \leqq 2$, but this is not essential to the proof of the Lemma.) This completes the proof of Lemma 2.

LEMMA 3. *If Q_1 is the set of vertices, which are not in I, of a minimum tree G then either Q_1 is empty or it contains at least one vertex adjacent to two vertices in I.*

Assume Q_1 is nonempty and let k_1 be the number of vertices in Q_1. To prove the lemma, assume the contrary, that is, assume all vertices in Q_1 are adjacent to at most one vertex in I. Let $E(Q_1)$ be the number of edges in the subgraph with the Q_1 vertices. Since $w(q_i) \geqq 3$ for all $q_i \in Q_1$,

$$E(Q_1) \geqq \sum_{i=1}^{k_1} \frac{w(q_1) - 1}{2} \geqq k_1 .$$

This implies that there exists a cycle in the subgraph with the Q_1 vertices, which is absurd.

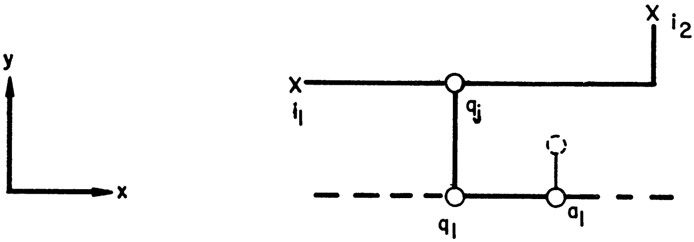

FIG. 7

The proof of Theorem 4 now follows immediately by successively applying Lemmas 2 and 3. Given a solution G_1, partition the vertices into two disjoint sets Q_1 and I_1. If Q_1 is not empty, then, by Lemma 3, at least one vertex in Q_1 has two adjacent vertices in I_1. By Lemma 2 this vertex can be moved to a new position which is in I_1. Partition the vertices of this new solution tree G_2, generated by this move, into two disjoint sets Q_2 and I_2. If Q_2 is empty, the theorem is proved. If Q_2 is nonempty, apply Lemma 3. By continuing this process, Q_l must be empty for some finite l since the set of points I is certainly finite.

4. The cases $n = 4$ and $n = 5$. In this section we concentrate on the solutions implied by Theorem 4, that is, those solutions which have all their vertices in I, although many of the statements made here are applicable to all solutions of S_4 and S_5.

We begin the study of S_4 by first solving the special case where the given four points are located on the corners of the enclosing rectangle R. By Theorem 4, there exists a solution G with no q-vertices since the four intersection points are occupied by p-vertices. Hence, in this case, there exists a solution to Steiner's problem which is the same as the solution to the minimum spanning tree problem. We state this as a separate lemma for future reference.

LEMMA 4. *If the four points of S_4 are located at the corners of the enclosing rectangle R, then*

$$d_{S_4} = d_{T_4} = l + 2w,$$

where l and w are the length and width[2] *of R.*

This is the simplest example where the number of sets Q is infinite. For, referring to Fig. 8, the line joining q_1 to q_2 can be moved parallel to itself anywhere in the interval $y_{p_1} \leqq y \leqq y_{p_2}$ which implies an infinite number of possible locations for q_1 and q_2.

[2] We assume throughout this paper that $w \leqq l$, i.e., we say the width, by definition, is the smaller of the two numbers.

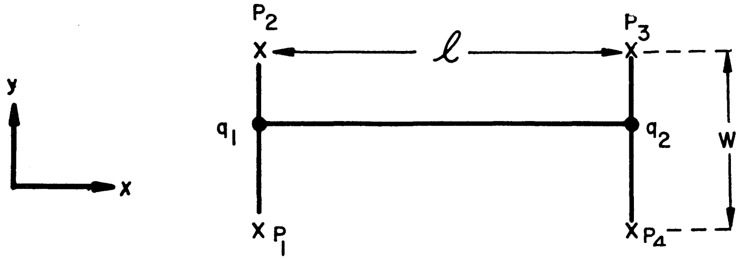

FIG. 8

We now show that when the four points of S_4 are located anywhere in the plane, the problem can always be reduced either to the above case or to a Steiner problem with less than four points. We first note that there cannot be any q-vertices on any side of R unless there are two p-points on this same side. (This can be deduced easily by Theorem 1 and Lemma 1.) Therefore, if there is a side s_1 of R with only one vertex on it, that vertex can be "moved" perpendicularly to s_1 to the closest intersection point. It is not difficult to see that, in doing this, we have reduced the general four-point Steiner problem to one where Lemma 4 is applicable (see Fig. 9a) or to a Steiner problem with less than four points (see Fig. 9b). In order to state these results more succinctly, some new terminology is introduced.

First order both (separately) the x- and y-coordinates of the given p-points in increasing order. (In doing this (x_i, y_i) no longer corresponds to the point p_i.) Then by drawing lines parallel to the y-axis through x_2 and x_3 and lines parallel to the x-axis through y_2 and y_3, this defines, in general, four points in I which we call $c_1, \cdots, c_4$. The rectangle which has these four points at its corners is called the *inner rectangle* R_1. Consider the four quadrants U_{c_i}, exterior to R_1, formed by the extended lines of R_1 and each of the c_i If there is a point p_j in a quadrant U_{c_i}, then we say that *p_j is transferred to the point c_i*. By construction, the p-vertex may, of course, be at the point c_i. The inner rectangle R_1 may degenerate to a straight line. There may or may not be vertices of G located at the points c_i.

A solution to S_4 can be found by applying Lemma 4 and the following theorem.

THEOREM 5. *Given four points in the plane, let l and w be the length and width of the enclosing rectangle R and let w_1 be the width of the inner rectangle R_1. If the p-vertices are transferred to four distinct points in $\{c_i\}$, then*

$$d_{S_4} = l + w + w_1,$$

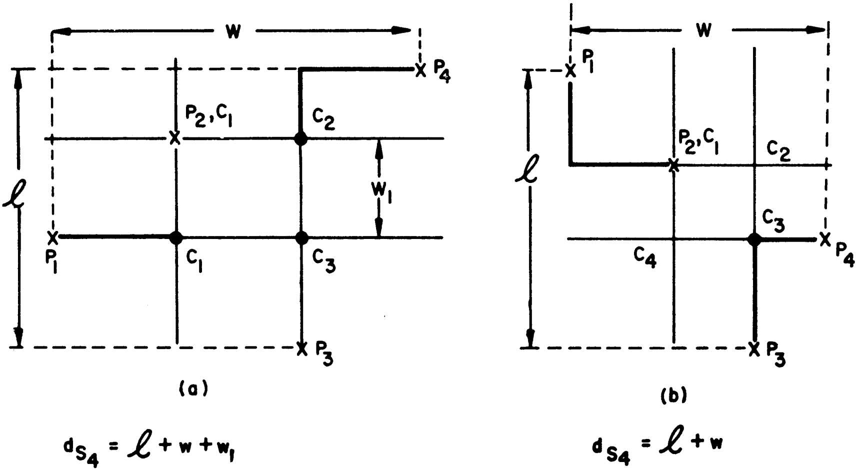

FIG. 9

and if they are transferred to less than four distinct points in $\{c_i\}$, *then*

$$d_{S_4} = l + w.$$

It is not difficult to see that, in general, d_{S_n} has a lower bound of $\frac{1}{2}$ the perimeter of the enclosing rectangle R, that is,

$$d_{S_n} \geqq \tfrac{1}{2}P(R) = l + w.$$

In fact, the following can easily be deduced from Theorem 5 and Lemma 4.

COROLLARY 1. *If* $d_{S_4} = l + w$ *then the solution to* S_4 *is unique; and if* $d_{S_4} > l + w$ *then there exists an infinity of solutions.*

The five-point Steiner problem can be treated in essentially the same way as the four-point problem. In this case, lines are drawn parallel to the y-axis through x_2, x_3, and x_4 and parallel to the x-axis through y_2, y_3, and y_4, so that there are in general nine c-points. The inner rectangle R_1 is the largest rectangle defined by these nine lines. The p-vertices are transferred to the c-points in a manner similar to the above, except that the concept of quadrants must be generalized to include the points c_i which are not at the corners of R_1. Hence, by transferring the p-points, we can always reduce the problem S_5 to the case where at least two p-vertices are on each side of the enclosing rectangle. The following theorem (the proof of which we omit) can then be used to find a solution to S_5 when the points are located anywhere in the plane.

THEOREM 6. *Given five points in the plane with at least two points on each side of the enclosing rectangle R. If four of the five points are at the corners of*

R, then

$$d_{S_5} = d_{S_4} = l + 2w \leqq d_{T_5},$$

where S_4 is the Steiner problem with these four corner points. If all five points are on the boundary of R, then

$$d_{S_5} = d_{T_5} \leqq l + 2w,$$

where l and w are the length and width of R.

5. General comments. An algorithm which incorporates several of the necessary conditions stated in §3 has been developed. It yields "good" approximate solutions to the n-point problem and exact solutions for $n \leqq 4$. The algorithm is easy and fast to do both by hand and on a computer.

The algorithm is rather elementary in concept and it is anticipated that a more sophisticated algorithm can be devised which incorporates almost all the results presented in this paper. For example, Theorem 4 states that there always exists a solution to Steiner's problem where the q-vertices are located at a predetermined finite set of points. Therefore, if the number of points n is not too large, we can find an exact solution by an exhaustive search procedure. These ideas will be investigated in the future.

Acknowledgment. The author wishes to thank P. H. Oden for many interesting discussions relating to this problem.

REFERENCES

[1] CLAUDE BERGE, *The Theory of Graphs*, John Wiley, New York, 1961.
[2] R. COURANT AND H. ROBBINS, *What is Mathematics?*, Oxford University Press, New York, 1941.
[3] RICHARD L. FRANCIS, *A note on the optimum location of new machines in existing plant layouts*, J. Indust. Engrg., 14 (1963), pp. 57–59.
[4] HUA LO-KENG ET AL., *Application of mathematical methods to wheat harvesting*, Chinese Math., 2 (1962), pp. 77–91.
[5] J. B. KRUSKAL, *On the shortest spanning subtree of a graph*, Proc. Amer. Math. Soc., 7 (1956), pp. 48–50.
[6] Z. A. MELZAK, *On the problem of Steiner*, Canad. Math. Bull., 4 (1961), pp. 143–148.
[7] F. P. PALERMO, *A network minimization problem*, IBM J. Res. Develop., 5 (1961), pp. 335–337.
[8] R. C. PRIM, *Shortest connecting networks*, Bell System Tech. J., 31 (1952), pp. 1398–1401.