ON THE STEINER PROBLEM

by

ERNEST COCKAYNE
M.A., (Oxon) 1965
M.Sc., McGill University, 1963.

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the Department
of
Mathematics

We accept this thesis as conforming to the
required standard.

THE UNIVERSITY OF BRITISH COLUMBIA

June, 1967

Thesis Supervisor: Dr. Z. A. Melzak.

ABSTRACT

The classical Steiner Problem may be stated: Given n points $a_1,\ldots,a_n$ in the Euclidean plane, to construct the shortest tree(s) (i.e. undirected, connected, circuit free graph(s)) whose vertices include $a_1,\ldots a_n$.

The problem is generalised by considering sets in a metric space rather than points in E^2 and also by minimising a more general graph function than length, thus yielding a large class of network minimisation problems which have a wide variety of practical applications.

The thesis is concerned with the following aspects of these problems.

1. Existence and uniqueness or multiplicity of solutions.
2. The structure of solutions and demonstration that minimising trees of various problems share common properties.
3. Solvability of problems by Euclidean constructions or by other geometrical methods.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I wish to express my warmest thanks to Dr. Z. A. Melzak both for his suggestion of the topic and his generous encouragement and assistance during the preparation of this thesis.

I am also indebted to Dr. Virginia Walsh of the University of Victoria for her constructive criticism of the first draft of the material and to Mrs. Mae Peters for her excellent typing of the manuscript.

The financial support of the Canadian National Research Council and the University of British Columbia is gratefully acknowledged.

I INTRODUCTION

Our starting point is the well known elementary problem:

(S_3) Given 3 distinct points a_1, a_2, a_3 in E^2 to find the point p which minimises the sum of distances $pa_1 + pa_2 + pa_3$. If triangle $a_1a_2a_3$ has an angle $\geq 120^\circ$, then p is its vertex, otherwise p is the unique point at which the sides of the triangle subtend angles of 120° and is called the Steiner point of the triangle. (S_3) may be generalised in many ways. For example

(P_n) Given n distinct points $a_1, \ldots, a_n$ in E^2 to find the point p which minimises the function $\sum_{i=1}^{n} pa_i$.

(P_3) and (S_3) are identical.

Let $b_1, \ldots, b_N$ be any set of distinct points in the plane. By a tree U on the vertices $b_1, \ldots b_N$ we mean any set consisting of some of the $\binom{N}{2}$ closed straight segments b_ib_j with the property that any two vertices can be joined by a sequence of segments belonging to U in one and only one way. A segment b_ib_j in U is called a branch of U, the length L(U) of U is the sum of the lengths of its branches and $\{b_i\}$ is the set of all vertices sending branches to the vertex b_i. The valency of b_i, written $w(b_i)$, is the number of vertices in $\{b_i\}$. We can now formulate further generalisations of (S_3):

(S_n) Given n distinct points $a_1, \ldots, a_n$ in the plane ($n \geq 3$), to construct the shortest tree(s) whose vertices include $a_1, \ldots, a_n$ and any set of k additional plane points $s_1, \ldots, s_k$ ($k \geq 0$).

($S_{n\alpha\beta\gamma}$) Given three non-negative real numbers α, β, γ and n distinct points $a_1, \ldots, a_n$ in the plane, to find an integer k and k additional points $s_1, \ldots, s_k$, and to construct the tree(s) U on the vertices $a_1, \ldots, a_n$, $s_1, \ldots, s_k$ so as to minimise the sum

$$T = L(U) + \alpha \sum_{i=1}^{n} w(a_i) + \beta \sum_{i=1}^{k} w(s_i) + \gamma k.$$

If $\alpha = \beta = \gamma = 0$, $(S_{n\alpha\beta\gamma})$ reduces to (S_n). Suppose now that $\beta = 0$ and $\alpha > \gamma$ where γ is sufficiently large. T will then be smallest when each $w(a_i)$ has its minimum value 1 so that as few extra vertices as possible are adjoined. However, since a tree is connected, $w(a_i) = 1$ for each i implies that $k \geq 1$. It follows, therefore, that when $\beta = 0$ and for suitable α, γ, the minimising trees of $(S_{n\alpha\beta\gamma})$ will be precisely the minimum length trees among those having $w(a_i) = 1$ for each i and $k = 1$ i.e. $(S_{n\alpha\beta\gamma})$ reduces to (P_n). A similar argument shows that if $\max(\beta, \gamma) \gg 1$ and $\alpha = 0$, then $(S_{n\alpha\beta\gamma})$ reduces to

(C_n) To connect n distinct given points in the plane by the shortest trees(s) whose vertices are these n points.

(C_n) is not a generalisation of (S_3) and has the important property of being discrete i.e. the length is to be minimised over a finite set of trees, while (S_n), (P_n) and $(S_{n\alpha\beta\gamma})$ are not discrete, since the co-ordinates of the extra vertices $s_1, \ldots, s_k$ are continuously varying unknowns.

$(S_{n\alpha\beta\gamma})$ and its special cases may be extended still further by replacing the n given points $a_1, \ldots, a_n$ with n disjoint plane, compact, connected sets $A_1, \ldots, A_n$. The definition of a tree given above is still valid with the following minor modifications. A vertex is a set and by the "segment" B_iB_j we mean a line of shortest distance joining the sets B_i and B_j. Such extremals certainly exist by a standard continuity and compactness argument.

The final generalisation is to change the metric space in the formulation. In the definition of a tree "segment" is replaced by "geodesic". For example we may consider identical problems in E^m, on the surface of a sphere in E^3 or in Minkowski metric spaces M^m. We formulate our most general problem:

$(S_{n\alpha\beta\gamma})$ Let M be a metric space with metric ρ which has the following properties:

1. M is finitely compact.
2. There exists a geodesic in M joining each two points of M.
3. For all a, b $\in$ M, ρ (a,b) is equal to the length of a geodesic joining a and b.

Given three non negative real numbers α, β, γ and n disjoint, compact, connected sets $A_1, \ldots, A_n$ in M, to find an integer k and k additional points $s_1, \ldots, s_k \in M$, and to construct the tree U on the vertices $A_1, \ldots, A_n, s_1, \ldots, s_k$ so as to minimise the sum

$$T = L(U) + \alpha \sum_{i=1}^{n} w(A_i) + \beta \sum_{i=1}^{k} w(s_i) + \gamma k.$$

Conditions (2) and (3) give meaning to the idea of a minimum length tree in the space M while (1) will be used to demonstrate the existence of such a tree.

This class of problems offers a wide variety of practical applications. The problems of joining geographical points, metropolitan areas, a set of lakes or sets of electrical terminals by minimum length systems of roads, railways, canals or connecting wire respectively are all examples of (S_n) for points or sets in some surface in E^3 or in E^3 itself. The particular Minkowski metric space which has distance function $d(z_1, z_2) = |x_1 - x_2| + |y_1 - y_2|$ is called the Manhattan metric. If there are n stores in a network of city blocks to be supplied by separate trips in rotation from a central supply depot, the optimal position for the depot is a solution of (P_n) in this Manhattan metric space. This has further application in some printed circuit designs where terminals may be joined only by wires running in two perpendicular directions. Finally suppose we wish to minimise the cost of a communications network joining areas $A_1, \ldots A_n$ in which there is a cost per unit length and also costs per terminal depending on the number of connections at the terminal, then the minimum cost networks will be solutions of $(S_{n\alpha\beta\gamma})$ in some metric space and for some α, β, γ.

The problem (S_3) dates back to Fermat and the generalisation (S_n) in E^2 is called the Steiner problem, and appears in the collected works of Steiner. In [1] there is a summary of the knowledge of (S_n) in 1941 and some interesting solutions, found by stretching soap films between pegs and glass plates, are exhibited. Recently due to the diverse applications, there has been renewed interest in these problems. Principally Z.A. Melzak [2] showed that (S_n) in E^2 is solvable by a finite number of Euclidean constructions (i.e. ruler-compass constructions in the classical sense) and posed $(S_{n\alpha\beta\gamma})$. R.C. Prim [3] has given an algorithm for solving $[C_n]$. Other results include a solution of (S_n) for n points in Manhattan metric space by Hanan [4], and a uniqueness theorem for (P_n) by Palermo [5]. Other minor references are [6] and [7].

The Thesis is concerned with the following aspects of these problems:

(a) Multiplicity or uniqueness of solution (b) feasibility of constructing all solutions by Euclidean constructions or possibly by wider geometrical algorithms, (c) the structure of minimising trees and demonstration that minimising trees of various problems share common structures.

The first section proves the existence of all minimum points and minimising trees mentioned in the thesis. We then consider the problem (S_n) for n points in E^2 giving another proof of the theorem that it is solvable by a finite number of Euclidean constructions. The proof clearly explains the algorithm involved and exhibits the structure of minimising trees. The methods used here are then generalised to solve (S_n) for n sets in E^2. We continue with a demonstration that solutions of (S_n) in other spaces namely E^n, a general surface in E^3 and Minkowski metric spaces share common properties. The final section discusses the problem $(S_{n\alpha\beta\gamma})$, showing that in Minkowski space it has a finite number of solutions and that in E^2 it is not in general solvable by Euclidean constructions.

II - EXISTENCE THEOREM

In this section we prove that minimising trees (points) exist for all problems mentioned in the thesis. A preliminary result shows that $(S_{n\alpha\beta\gamma})$ is equivalent to a finite number of minimum length problems.

Let $A_1,\dots,A_n$ be n disjoint, compact, connected sets in a metric space M having the three properties listed in section 1.

Definition. U is a μ-tree on $A = \{A_1,\dots,A_n\}$ if U is a tree whose vertices are $A_1,\dots,A_n$ and the points $s_1,\dots,s_k$ which satisfies:

(i) $w(s_i) \geq 3$, $i = 1,\dots,k$.

(ii) $0 \leq k \leq n-2$.

Lemma. If a solution of $(S_{n\alpha\beta\gamma})$ exists, it is a μ-tree on A.

Proof. We shall call the extra vertices $s_1,\dots,s_k$, "s-points". $w(s_i) \geq 2$ for each i, otherwise we could decrease T by deleting s_i. Now suppose $w(s_i) = 2$ for some i. Then the tree formed by replacing s_ix, s_iy, the two branches joining s_i, by the single branch xy, has a smaller value of T. Hence (i). The number of branches of U leading to s-points is $\geq 3k/2$ from (i), the connectivity of U assures us that the number of branches from the A-sets is $\geq n/2$ and a tree with n+k vertices has n+k-1 branches. Therefore $(n+3k)/2 \leq n+k-1$, from which we deduce (ii).

Definition. By the association of a μ-tree U on A, we mean the integer k and the sets $\{A_i\}$ and $\{s_j\}$ $(i = 1,\dots,n$, $j = 1,\dots,k)$.

Theorem 1. The problem $(S_{n\alpha\beta\gamma})$ is reducible to a finite number of minimum length problems.

Proof. The relation "has the same association as" on the set of all μ-trees on A is an equivalence relation. We show that the number of equivalence classes is finite. Suppose there are k additional vertices. Then we have n+k vertices on which to construct a tree. i.e. n+k-1 pairs of vertices to be joined, must be selected from the possible $\binom{n+k}{2}$ pairs.

Thus the number of associations with k extra vertices is not greater than

$$g(n,k) = \left(\begin{array}{c} \binom{n+k}{2} \\ \\ n+k-1 \end{array} \right)$$

and the total number of associations is not greater than $\sum_{k=0}^{n-2} g(n,k)$ and hence is finite.

Let the equivalence classes be $C_1,\ldots,C_N$ and let C be any one of these classes. The tree(s) which minimise T in C are precisely the tree(s) of minimum length of C (if one exists), since the association common to all the tree(s) of C fixes the other three terms of T i.e. for all $U \in C$

$$\alpha \sum_{i=1}^{n} w(A_i) + \beta \sum_{j=1}^{k} w(s_j) + \gamma k \text{ is constant.}$$

The minimising trees of $(S_{n\alpha\beta\gamma})$ are a subset of the trees belonging to $C_1,\ldots,C_N$ by the Lemma. Therefore, each minimising tree of $(S_{n\alpha\beta\gamma})$ is a solution to one of the following N minimum length problems:

(1) For $i = 1,\ldots,N$ to construct the tree(s) $U \in C_i$ which minimise $L(U)$.

<u>Theorem 2.</u> Let $C \in \{C_1,\ldots,C_N\}$. There exists a tree of minimum length in C.

<u>Proof.</u> The association of trees in C stipulates which of the pairs s_iA_j , s_is_j , A_iA_j will be joined by geodesics as branches of trees in C. We exclude the case $k = 0$ for which the theorem is obvious. Let

$\{A_{i1}, A_{i2},\ldots,A_{i\lambda_i}\} = \{A_t : A_t \in A \text{ and } s_iA_t \text{ is a branch of trees in C}\}$

and let R_1, R_2, sets of unordered pairs of integers be defined as follows:

$$R_1 = \{(i,j) : s_is_j \text{ is a branch of trees in C}\}.$$

$$R_2 = \{(i,j) : a_ia_j \text{ is a branch of trees in C}\}.$$

Then the length of a tree in C is

$$f(s_1,\dots,s_k) = \sum_{i=1}^{k} \sum_{j=1}^{\lambda_i} \rho(s_i,A_{ij}) + \sum_{(i,j)\in R_1} \rho(s_i,s_j)$$

$$+ \sum_{(i,j)\in R_2} \rho(A_i,A_j) \ .$$

Suppose L is the length of the shortest tree with vertices $A_1,\dots,A_n$ only. Let

$$Z = \{z : z \in M \text{ and } \min_i \rho(z,A_i) \leq L\} \ .$$

Then every s-point of a tree of shortest length in C (if one exists) is in Z, for otherwise the length of the tree would necessarily be greater than L. Thus if $\{s_1,\dots,s_k\}$ is a set of s-points of a minimum length tree in C, then $\{s_1,\dots,s_k\}$ is an element of the cartesian product Z^k. Now since Z is closed and bounded and M is finitely compact, Z is compact implying by the Tychonoff Theorem that Z^k is compact. But $f(s_1,\dots,s_k)$ is continuous on Z^k and so has a minimum value on Z^k. Hence the theorem.

Corollary. There exist minimising trees of $(S_{n\alpha\beta\gamma})$ in M. For $(S_{n\alpha\beta\gamma})$ is equivalent to the finite set (1) of minimum length problems (Theorem 1.) each of which by the theorem has a solution.

III - AN EFFECTIVE ALGORITHM FOR (S_n) IN E^2.

We begin by stating elementary constructions of the solution of (S_3) which was given in the introduction. These constructions are of fundamental importance in the development of the algorithm for (S_n) both for points in the plane and for n sets in the plane which will be discussed in the next chapter.

Let a_1, a_2, a_3 be the vertices of a plane triangle T with no angle $\geq 120^\circ$ and $a_{12} a_{23} a_{31}$ be the three third vertices of the equilateral triangles built outward on the sides of T (with obvious notation). Then p, the minimum point of (S_3), lies at the unique point of intersection inside T of the circumscribing circles of the equilateral triangles. Again, if p' is the intersection of the line $a_{12}a_3$ with the circle through $a_1 a_2 a_{12}$, each side of T subtends 120° at p' which therefore coincides with p. It follows that each of the lines $a_{12}a_3$, $a_{23}a_1$, $a_{31}a_2$ passes through p giving a third method of construction.

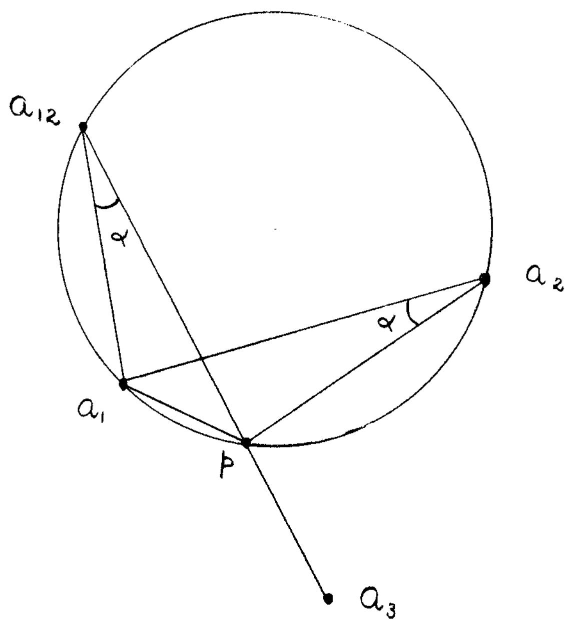


Fig. 1.

Using the notation as in fig. 1. by the sine law

$$\frac{a_1p}{\sin\alpha} = \frac{a_2p}{\sin(60-\alpha)} = \frac{a_{12}p}{\sin(60+\alpha)} .$$

$$\therefore \quad a_1p + a_2p = a_{12}p\left\{\frac{\sin\alpha + \sin(60-\alpha)}{\sin(60+\alpha)}\right\} = a_{12}p .$$

The addition of a_3p to each side gives

$$a_{12}a_3 = a_1p + a_2p + a_3p .$$

Thus each of the three lengths $a_{12}a_3$, $a_{31}a_2$, $a_{23}a_1$ equals the minimum value of the function $a_1x + a_2x + a_3x$. We remark that if T has an angle $\geq 120^\circ$ none of the above constructions will produce a point inside T. Suppose $A = \{a_1,\ldots,a_n\}$ is a set of n distinct points in the plane and that U, a minimising tree of (S_n) for the set A, has extra vertices $s_1,\ldots,s_k$. Then

P1. U is non-self intersecting. i.e. two branches of U do not intersect except at an end point.

P2. $w(s_i) = 3$, $i = 1,\ldots,k$.

P3. $w(a_j) \leq 3$, $j = 1,\ldots,n$.

P4. $0 \leq k \leq n-2$.

P5. Each s_i is the Steiner point of the triangle formed by $\{s_i\}$.

These properties are given in [1]. We indicate the proofs for completeness. Suppose that two branches x_1x_2 , x_3x_4 of U intersect at x. Then one of the angles at x (say $\measuredangle x_1xx_3$) is less than 120° and U may be shortened by replacing the segments x_1x, x_3x by x_1p, x_3p, xp where p is the solution of (S_3) for the triangle x_1x_3x. Hence P1. An identical argument proves that $w(x) \leq 3$ for all vertices x of U. Therefore U has the property P3 and $w(s_i) \leq 3$ for each i. However $w(s_i) \geq 3$ for otherwise there is no gain in introducing the additional vertex s_i. Thus P2 is established and from this P5 is immediate. P4 is proved as in the Lemma on Page 5.

<u>Definitions.</u> A tree with vertices $\{a_1,\ldots,a_n\}$ and $\{s_1,\ldots,s_k\}$ (from this point these will be termed a-points and s-points respectively) has the property $\overline{P4}$ if $k = n-2$.

V is a subtree of a tree U if and only if (i) V is a tree and (ii) the set of geodesics of V is contained in the set of geodesics of U.

A tree U is an S-tree on A if it has properties P1, P2, P3, P4, P5.

A tree U is an $\overline{S}$-tree on A if it has properties P1, P2, P3, $\overline{P4}$, P5.

A tree U is an S^*-tree on A if it has properties P2, P3, $\overline{P4}$, P5.

The finite set $\{A_1,A_2,\ldots,A_t,R\}$, $(t \geq 0)$ is a Division of the set A if and only if

1. Each $A_i \subseteq A$, $R \subseteq A$.
2. $R \cap A_i = \emptyset$, $i = 1,\ldots,t$.
3. No a_j can be an element of more than 3 of the sets A_i.
4. Each A_i has 3 or more elements.
5. $A_1 \cup A_2 \cup \ldots \cup A_t \cup R = A$.

<u>Lemma 1.</u> If U is any S-tree on A then for some division $\mathfrak{A} = \{A_1,A_2,\ldots,A_t,R\}$ of A there exist $\overline{S}$-subtrees of U on A_i for $i = 1,\ldots,t$.

<u>Proof.</u> If U contains no s-point, the required division is $\{A\}$. If S, the set of s-points of U is non-empty, we define the relation "o" on S as follows.

$s_i \text{ o } s_j$ iff the sequence of segments of U joining s_i to s_j contains no a-point of U. The relation is an equivalence relation and therefore partitions S into mutually exclusive and exhaustive sets $S_1,\ldots,S_t (t > 0)$. Define $A_i = \{a_j : a_j \in \{s_k\}$ for some $s_k \in S_i\}$, $i = 1,\ldots,t$ and $R = A - \bigcup_{i=1}^{t} A_i$. The set $\{A_1,A_2,\ldots,A_t,R\}$ is a division of A. It remains to show that there is an $\overline{S}$-subtree of U on each A_i. Let U_i be the subtree of U whose vertex set is $A_i \cup S_i$, $i = 1,\ldots, t$. (This is certainly a subtree of U by construction). U_i has the properties P1,P2,P3 and P5. We prove $\overline{P4}$. Let A_i contain p points, S_i q points and further suppose that S_i contains n_1,n_2,n_3 s-points which directly join 1,2 and 3 other s-points respectively. Then

(2) $$n_1 + n_2 + n_3 = q \ .$$

The number of branches of U_i connecting s-points is $(n_1 + 2n_2 + 3n_3)/2$. But by the defining property of S_i this number is q-1. Hence

(3) $$n_1 + 2n_2 + 3n_3 = 2(q-1).$$

The number of branches of U_i connecting an a-point to an s-point is $2n_1 + n_2$ since the valency of each s-point is 3. Hence

(4) $$2n_1 + n_2 = p \ .$$

From equations (2) (3) and (4) we deduce $q = p-2$. Therefore U_i satisfies $\overline{P}4$ and is an $\overline{S}$-subtree of U. Hence the Lemma.

Incidentally one can also prove $n_1 - n_3 = 2$ from (2) (3) and (4) which implies that $n_1 \geq 2$. We note that the non-selfintersection property is not involved in the establishment of these equations and state that an S^*-tree on a set A has at least two s-points which directly join exactly one other s-point and two a-points. This fact will be used in the next Lemma.

We call the subtrees U_i $(i = 1,\ldots,t)$ the Components of U and suggest that the components of a minimising tree U may be considered as stability sets for U in the following sense. If one a-point a_q is perturbed by a sufficiently small amount, there is a minimising tree U' for the new set of n a-points which is identical to U except for a small perturbation of the components to which a_q belongs. If p,q are points in the plane we shall denote by (pq) and (qp) the third vertices of the equilateral triangle on pq as base, (pq) being the point to the left of p looking from p along pq.

The construction we now describe, a direct consequence of the solution of (S_3), is crucial to the proof. Let U be an $\overline{S}$-tree on $A = \{a_1,\ldots,a_n\}$ with s-points $s_1,\ldots,s_k$ then there exists (see note following Lemma 1.) an s-point say s_1 which is directly connected to two a-points say a_1 and a_2 and a third vertex x. In fact a portion of U appears as in fig. 2.

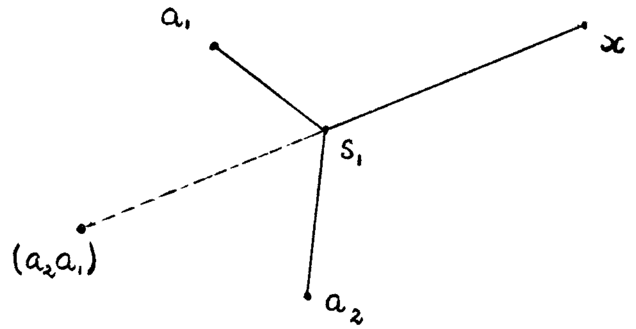


Fig. 2.

Since s_1 is the Steiner point of $\{s_1\}$, the line xs_1 produced passes through (a_2a_1) and the tree U' on $A' = \{(a_2a_1), a_3, \ldots, a_n\}$ with s-points $s_2, \ldots, s_k$ formed from U by replacing the branches a_1s_1 , a_2s_1 , s_1x by the single branch $(a_2a_1)x$ is an S^*-tree on A' (the non-selfintersection property may have been contradicted). Further $a_1s_1 + a_2s_1 + s_1x = (a_2a_1)x$ therefore U,U' have equal lengths. The important point is that (a_2a_1) is constructed from the original a-points only. We call the above the "Equilateral construction."

We next define the term "Association" of an $\bar{S}$-tree on a set A. From U, we form a tree U' and set A' as above. The construction is repeated forming a new tree U" and set A", U"' and A"' etc. until the set $A^{(r)}$ contains only two points. (Actually r = the number of s-points in the original tree U). The two points of $A^{(r)}$ can be expressed in terms of the original a-points of U and the equilateral triangle bracketing notation defined above. This representation of $A^{(r)}$ we call an "Association" of the tree U and the line joining the two points of $A^{(r)}$ we call an "axis" of U. We give a simple example below. We note the following:

(i) The process is always possible since at every stage the tree $U^{(k)}$ is an S^*-tree on $A^{(k)}$ and hence has an s-point which directly joins two a-points, (in fact at least 2 such s-points by the note following Lemma 1).

(ii) It follows from (i) that every $\bar{S}$-tree on A has an association and an axis (certainly not a unique association.)

(iii) At each stage length is preserved. i.e. The length $L(U)$ = length of an axis of U.

(iv) No two $\bar{S}$-trees on A have a common association.

<u>Example</u>. In Fig. 3, U is an $\bar{S}$-tree on $A = \{1,2,3,4,5\}$ with s-points s_1, s_2, s_3.

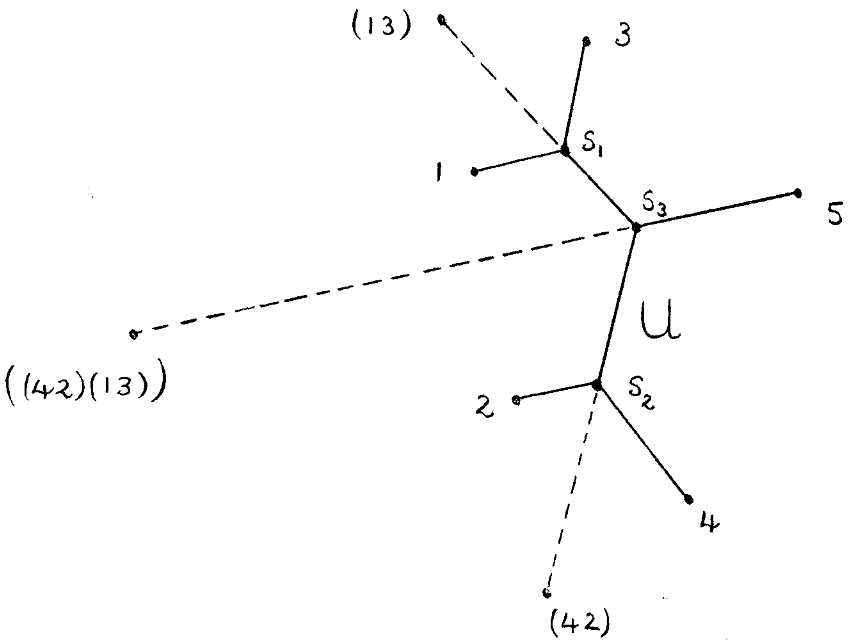


Fig. 3.

We "pair" the points 1 and 3 giving

$A' = \{(13), 2, 4, 5\}$ and U' with branches $(13)s_3, s_3 5, s_2 s_3, 2s_2, 4s_2$.

Next we pair 4 and 2

$A'' = \{(13),(42),5\}$, U'' has branches $(42)s_3, (13)s_3, s_3 5$.

Finally we pair (13) and (42)

$A''' = \{\underline{((42)(13))\ ,\ 5}\}$, U''' has branch $((42)(13))5$.

The underlined portion i.e. A''' without the set parentheses is an association of U. The length of U is the length of the branch of U'''.

<u>Lemma 2.</u> If it is known that U, an $\bar{S}$-tree on A has a certain association α, we can construct U by a finite number of Euclidean constructions.

<u>Proof</u> The Lemma is true for $n = 3$. Assume it is true for $n = N$ and let $A_{N+1} = \{a_1,\ldots,a_{N+1}\}$ be any plane set of N+1 points on which U is an $\bar{S}$-tree with association α. Suppose the labelling of points in A_{N+1} is such that a_1, a_2 in this order in α have no brackets or comma separating them. We now consider the set $A_N = \{(a_1a_2), a_3,\ldots,a_{N+1}\}$. From the equilateral construction there exists U', an $\bar{S}$-tree on A_N which has association α except that (a_1a_2) is now regarded as a single point. By the inductive hypothesis we can construct U' by a finite number of Euclidean constructions. Let $(a_1a_2)x$, the branch of U' connecting (a_1a_2), be replaced by the branches a_1s, a_2s, sx where s is the point of intersection of the circle through (a_1a_2), a_1, a_2 with the line $(a_1a_2)x$. The resulting tree U, by the equilateral construction, is the (unique by note (iv) Page 13) $\bar{S}$-tree on A_{N+1} with association α. Hence the Lemma by induction.

<u>Lemma 3.</u> The set of all minimum length $\bar{S}$-trees on $A = \{a_1,\ldots,a_n\}$ is finite and may be constructed by a finite number of Euclidean constructions.

<u>Proof.</u> Any two points formed by combining the elements of A by the above equilateral point bracketing notation we call an association of A. Then the set of all assocations of all $\bar{S}$-trees on A is a subset of the finite

set $\mathcal{B}$ of all associations of A (in fact a proper subset for $n > 3$). If for each $b \in \mathcal{B}$ we perform the finite number of Euclidean constructions (Lemma 2) that constructs the $\overline{S}$-tree on A with association b (if such a tree exists) we will construct all the $\overline{S}$-trees on A. Hence the Lemma.

For the main theorem of this chapter we shall need to refer to Prim's efficient algorithm for

(C_n) : Given n compact, connected, disjoint sets $W_1,\ldots,W_n$, to connect them together by the shortest tree(s) whose vertices are exactly these sets.

The method is as follows:

(1) Join W_1 to its nearest neighbor, say W_2,

(2) Replace W_1 and W_2 by their union.

(3) Repeat the same procedure for the new class of (n-1) sets and keep on repeating until only one set remains.

Actually, Prim's algorithm was originally intended for the case of points; however, it works equally well for sets. Moreover if the nearest neighbor of some W_i can be connected to it by several segments of the same minimal length, or if W_i has several nearest neighbors, we perform the connection in all possible ways and get then the set of all connecting trees of the same minimal length.

Theorem 3. For every n, there exists a finite number of Euclidean constructions yielding all the minimising trees of the problem (S_n). The minimising trees of (S_n) are precisely the minimum length S-trees on A, each of which has minimum length components (Lemma 1) on some division of A. The following method, therefore yields all solutions of (S_n).

1. From the finite set $\mathfrak{A} = \{\mathfrak{A}_1,\ldots,\mathfrak{A}_N\}$ of all divisions of A.

2. For each $\mathfrak{A}_i = \{A_{i1}, A_{i2},\ldots,A_{it_i}, R_i\}$, we construct by a finite sequence of Euclidean constructions, the finite set C_{ij} of all minimum length $\overline{S}$-trees on A_{ij} $(j = 1,\ldots,t_i)$. (Lemma 3.)

3. If $C_{ij} = \phi$ for any $j = 1,\ldots, t_i$, (there may not exist an $\overline{S}$-tree on an arbitrary set of points), we reject the division $\mathfrak{A}_i$. If each $C_{ij} \neq \phi$, we call $\mathfrak{A}_i$ "admissible".

4. For each admissible division $\mathfrak{A}_i$ we now form the finite set

$$\Gamma_i = \{(U_{i1},U_{i2},\ldots,U_{it_i}) : U_{ij} \in C_{ij}\} .$$

5. For each element of Γ_i we connect the U_{ij} to each other and the residual set R_i in optimal way(s) so that the resulting tree(s) on A have minimum length. ($\bar{S}$-trees on non-disjoint sets A_{ip},A_{iq} are automatically joined).

Prim's algorithm may be used to effect the joining. We note that the application of this algorithm must not contradict any of the properties P1-P5 on the resulting tree could not have minimum length e.g. every connection must be a segment joining two a-points. The number of optimal joinings is certainly finite. We thus obtain a finite set of trees on A from which we select the set V_i with minimum length. ($V_i = \phi$ if $\mathfrak{A}_i$ is not admissible). Then the solutions of (S_n) are precisely the minimum length trees of the set $\bigcup_{i=1}^{N} V_i$ which has been constructed by a finite sequence of Euclidean constructions. This completes the proof of the theorem.

We conclude this chapter with a diagram of an S-tree on $A = \{a_1,\ldots,a_{11}\}$ which has two component $\bar{S}$-trees U_1,U_2 on sets $A_1 = \{a_1,a_2,a_3,a_4\}$ and $A_2 = \{a_2,a_5,a_6,a_7,a_8,a_9\}$ respectively. The residual set R is $\{a_{10},a_{11}\}$ and the corresponding division of A is $\mathfrak{A} = \{A_1,A_2,R\}$.

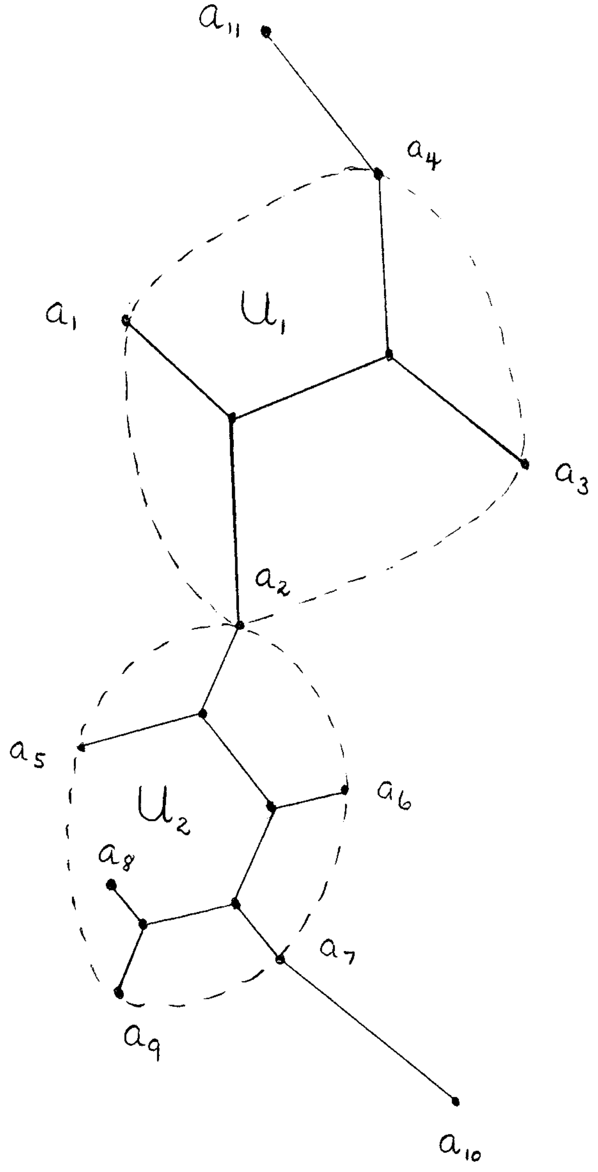


Fig. 4.

IV - (S_n) for sets in E^2.

We now discuss the extension of the techniques of Chapter III to the problem (S_n) for n compact, connected, disjoint sets $A_1,\dots,A_n$ in the plane. Let U be a minimising tree of (S_n) which has additional vertices $s_1,\dots,s_k$. An end point of a branch of U is either a point in $B(A_i)$ or a vertex s_i. The following properties of U are simply deducible by the same methods as their counterparts in Chapter III.

Q1. Two branches may not intersect except at an end point.

Q2. If two branches share an end point, they subtend there an an angle $\geq 120^\circ$.

Q3. For each s_i $(i = 1,\dots,k)$, $w(s_i) = 3$.

Q4. Each s_i is the Steiner point of $\{s_i\}$, where $\{s_i\}$ is the set of three end points of branches joining s_i.

Q5. $0 \leq k \leq n-2$.

Definition. Let A,B be compact, connected sets in E^2 and let (ab) be defined as in Chapter III. We define the equilateral sum (AB) of A and B by

$$(AB) = \{(ab) : a \in A, b \in B\}.$$

(AB) is compact and connected. (AB) $\neq$ (BA). If A is a point, (AB) and B are congruent under a rotation of 60° about A. If c is a boundary point of (AB) then c = (ab) where a,b are boundary points of A,B respectively. Distributive laws hold. If $A = C \cup D$ then

$$(AB) = (CB) \cup (DB)$$

and similarly for $B = E \cup F$. The following properties hold for (AB) whenever A and B have them: convexity, arcwise-connectedness, being the smooth boundary of a region, being a simple polygon. Let d(X,Y) denote the distance between two compact, disjoint sets X,Y.

Lemma. (Generalised equilateral construction).
Let A_1,A_2,A_3 be three compact, connected, disjoint sets in E^2. Suppose that a minimum length tree U connecting these sets consists of three

straight segments a_1v, a_2v, a_3v , $(a_i \in B(A_i))$, meeting at an additional vertex v and let the rotation $a_1 \to a_2 \to a_3 \to a_1$ be counter-clockwise about v.
Then (i) $a_1(a_2a_3)$, $a_2(a_3a_1)$, $a_3(a_1a_2)$ intersect in v.

(ii) $a_1(a_2a_3) = d(A_1, (A_2A_3))$

$= a_2(a_3a_1) = d(A_2, (A_3A_1))$

$= a_3(a_1a_2) = d(A_3, (A_1A_2)) = L(U)$.

i.e. a_1,a_2,a_3 are selected in A_1,A_2,A_3 so that $a_1(a_2a_3)$ is a shortest segment connecting A_1 and (A_2A_3) etc.

Proof. v is the Steiner point of a_1,a_2,a_3 and therefore by the equilateral triangle solution of (S_3) in Chapter III, $a_1(a_2a_3)$, $a_2(a_3a_1)$, $a_3(a_1a_2)$ intersect at v and the length of each of these segments is equal to L(U). Hence U is a minimising tree if and only if $a_1(a_2a_3)$ (or similarly $a_2(a_3a_1)$, $a_3(a_1a_2)$) attains its minimum value i.e. $a_1(a_2a_3) = d(A_1,(A_2A_3))$ as required.

The problem (S_n) for a class A of n sets $\{A_1,\ldots,A_n\}$ can now be solved from the properties Q1-Q5 and the generalised equilateral construction by exactly the same sequence of steps which solved (S_n) for points using P1-P5 and the equilateral construction. We form the finite set of all divisions of the class A. A division of A has the form $\{\alpha_1,\alpha_2,\ldots,\alpha_6\ r\}$ where each α_i and r is a subclass of A. We find the minimum length $\overline{S}$-trees on each α_i of each division and join together these components optimally using Prim's method. The components are constructed using the association and axis technique as before. The higher equilateral sums e.g. $((A_2A_1)(A_6A_5))$, are unambiguously defined, and by an axis e.g. $[((A_2A_1)(A_6A_5)), (A_4A_3)]$ we understand the straight segment joining points in $X = ((A_2A_1)\ (A_6A_5))$ and in $Y = (A_4A_3)$ for which the minimum d(X,Y) is attained. We note the following:

Theorem 4. Let $A = \{A_1,\ldots,A_n\}$ be a class of n simple polygons. Then minimum length trees on A can be found using a finite sequence of Euclidean constructions.

To prove this it suffices to observe that:

(a) the equilateral sum of two polygons is a polygon and hence constructible by Euclidean means.

(b) the closest distance between two polygons can be found by Euclidean means.

Fig. 5. shows that there may be an infinite number of minimum length trees connecting a family of polygons. $A = \{A_1, A_2, A_3\}$ where A_1, A_2 are equilateral triangles and A_3 is a single point. It is easily verified, using the generalised equilateral construction, that if a_1, a_2 are any pair of points of X_1Y_1, X_2Y_2 which are symmetrically placed with respect to A_3 then the network U has minimum length.

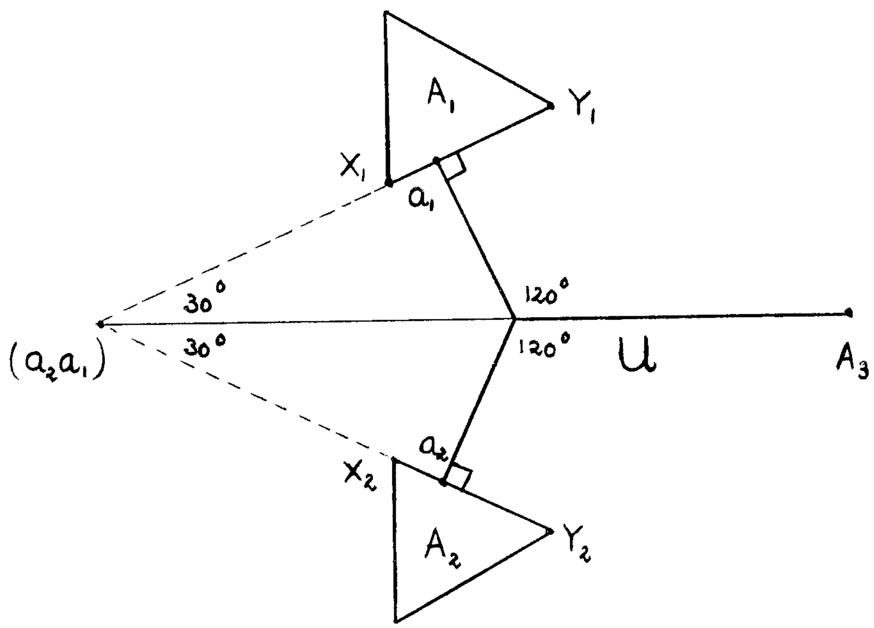


Fig. 5

Theorem 4 leads immediately to

Theorem 5. Let $A = \{A_1, \ldots, A_n\}$ be n sets and suppose that each A_i is arbitrarily well approximable by simple polygons. Then minimum length trees on A can be found by a finite sequence of Euclidean constructions to within arbitrary accuracy.

For if A_1 and A_2 are approximated by the polygons P_1 and P_2, then (A_1A_2) is approximated by (P_1P_2).

V - (S_n) IN OTHER METRIC SPACES

1. Euclidean m-Space.

Theorem 6. The minimising trees of (S_n) for n points in E^m $(m \geq 3)$ have the properties P1-P5 listed in Chapter III.

Proof. We show that each vertex of a minimising tree U which has a-points $a_1,\ldots,a_n$ and s-points $s_1,\ldots,s_k$ has valency ≤ 3. For suppose U has a vertex x and branches xx_i along the directions of the unit vectors u_i, $(i=1,\ldots,4)$. Then one of the angles at x (say $\measuredangle\, x_1xx_2$) is less than 120° since $u_i \cdot u_j \leq -\frac{1}{2}$ for all $i \neq j$ implies that

$$w^2 = \left[\sum_{i=1}^{4} u_i\right]^2 = \sum_{i=1}^{4} u_i^2 + 2 \sum_{i \neq j} u_i \cdot u_j < 4 + 2.6.(-\tfrac{1}{2}) = -2,$$

which is impossible. It follows that xx_1, xx_2 is not the minimum length network connecting xx_1x_2 contrary to assumption. The rest of the proof is identical to that given for the properties P1-P5 in E^2.

2. A Surface in E^3.

Let D be a surface in E^3 free from singularities of any kind. It will be shown that minimising trees of (S_n) in D have properties identical to those for E^m. We first prove two results which show that the 120° property of additional vertices holds in D.

Suppose A, B, C are distinct points in D and $P \notin \{A,B,C\}$ minimises the sum $\rho(P,A) + \rho(P,B) + \rho(P,C)$, we prove that the angles at P between the geodesics PA, PB, PC are each 120°. Let $\rho(P,A) = a$, $\rho(P,B) = b$ and $\rho(P,C) = c$. Consider the geodesic ellipse E (= the locus of points Z such that $\rho(Z,A) + \rho(Z,B) = a + b$) and the geodesic circle $\rho(Z,C) = c$. These closed curves touch at P, for otherwise there would be a point Y interior to both curves such that $\rho(Y,A) + \rho(Y,B) < a + b$ and $\rho(Y,C) < c$ contradicting the minimum property of P. Since geodesic PC meets $\rho(Z,C) = c$ orthogonally, geodesic PC meets E orthogonally. By a result of classical Differential Geometry [8] page 120, E bisects the angle between the geodesic parallels $\rho(Z,A) = a$ and $\rho(Z,B) = b$ and therefore, since the geodesics AP, BP meet these circles orthogonally, the angles α and β of Fig. 6. are equal.

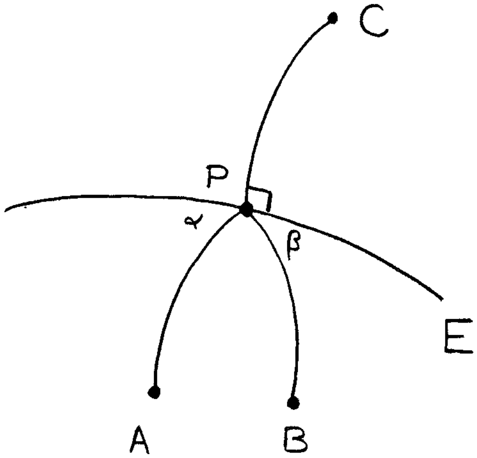


Fig. 6.

$\therefore$ (5) $\measuredangle APC = \measuredangle BPC$.

Similarly by considering the geodesic ellipse $\rho(Z,A) + \rho(Z,C) = a + c$ we prove $\measuredangle APB = \measuredangle BPC$ and this together with (5) proves the result.

Secondly we show that if ABC is a geodesic triangle on D with the angle at A less than 120°, then A does not minimise the sum $\rho(Z,A) + \rho(Z,B) + \rho(Z,C)$. Let V be an ϵ-neighborhood of A sufficiently small so that for all r, s $\in$ V there is only one geodesic joining them. Let B(V) intersect the geodesics AB, AC in X and Y. Consider the following 1-1 mapping of the geodesic triangle AXY onto the tangent plane at A. For Q in the geodesic triangle AXY with $\rho(A,Q) = q$, the corresponding point Q' is the point on the tangent line to the geodesic AQ at A such that AQ' = q. Since the angle A of the plane triangle AX'Y' is less than 120°, there exists P' in the tangent plane such that $AP' + X'P' + Y'P' < AX' + AY'$ and furthermore the difference is proportional to ϵ i.e. $AX' + AY' - (AP' + X'P' + Y'P') = \epsilon k_1$ for some k_1.

If ϵ is sufficiently small, for all $Q_1, Q_2 \in V$, $|Q_1' Q_2' - \rho(Q_1, Q_2)| < L \epsilon^2$ where L is constant. Hence if $P \in V$ corresponds to P' in the tangent plane, $|(AP' + X'P' + Y'P') - \{\rho(A,P) + \rho(X,P) + \rho(Y,P)\}| < k_2\epsilon^2$ for some k_2.

$$
\begin{aligned}
\therefore\quad & \{\rho(A,X) + \rho(A,Y)\} - \{\rho(A,P) + \rho(X,P) + \rho(Y,P)\} \\
= \quad & (AX' + AY') - (AP' + X'P' + Y'P') \\
+ \quad & (AP' + X'P' + Y'P') - \{\rho(A,P) + \rho(X,P) + \rho(Y,P)\} \\
> \quad & k_1\epsilon - k_2\epsilon^2 > 0 \text{ if } \epsilon \text{ is sufficiently small.} \\
\therefore\quad & \rho(A,X) + \rho(A,Y) > \rho(A,P) + \rho(X,P) + \rho(Y,P).
\end{aligned}
$$

If we now add $\rho(X,B) + \rho(X,C)$ to each side and apply the triangle inequality on the right we obtain

$$\rho(A,B) + \rho(A,C) > \rho(A,P) + \rho(B,P) + \rho(C,P)$$

showing that A does not minimise $\rho(Z,A) + \rho(Z,B) + \rho(Z,C)$ as required.

Using these 120° properties and proofs identical to those for P1-P5 (Chapter III), we deduce that U, a minimising tree of S_n in D with extra vertices $s_1,\ldots,s_k$, has the properties P1,P2,P3,P4 and the following analog of P5: P'5: For each $i = 1,\ldots k$ if $\{s_i\}$ contains points p_i, q_i, r_i then each of the angles at s_i between the geodesics p_is_i, q_is_i, r_is_i is 120°.

3. Plane Minkowski Metric Space.

Let Σ be a centrally symmetric convex surface in E^m with centre 0. The m-dimensional Minkowski metric space M^m associated with Σ is obtained by defining the distance $\rho(x,y)$ for $x,y \in E^m$ as follows. If $x = y$, $\rho(x,y) = 0$. If $x \neq y$ let the ray with initial point 0 which is parallel to xy meet Σ at P. Then $\rho(x,y) = xy/OP$, where xy and OP are usual Euclidean distances. M^m is a metric space satisfying the three conditions of the introduction and having the following properties ([9] page 21), x and y are considered as m-dimensional vectors. For all $x,y \in M^m$

(i) $\rho(x,y) = \rho(0,x-y)$ and more generally, any translation is an isometry.

(ii) The triangle inequality is strict provided that Σ is strictly convex and the three points involved are non-collinear.

(iii) For Σ strictly convex

$$\rho(0, x+y) \leq \rho(0,x) + \rho(0,y)$$

and this inequality is strict unless $0,x,y$ are collinear with x,y lying on the same side of 0.

Small results in the necessary preliminary theory of this section will be termed propositions and the principal results are Theorems 7 and 8. For the rest of the section M^2 will mean a plane Minkowski metric space, the defining curve of which is strictly convex.

Proposition 1. Given n distinct non-collinear points $a_1,\ldots,a_n$ in M^2. There exists a unique point z minimising the function

$$f(z) = \sum_{i=1}^{n} \rho(z,a_i).$$

Proof. The existence of a minimum point was proved in Chapter II. Suppose, contrary to the proposition, that $f(z)$ has minima λ at z_1 and z_2. Then

$$f\left(\frac{z_1 + z_2}{2}\right) = \sum_{i=1}^{n} \rho\left(\frac{z_1 + z_2}{2}, a_i\right)$$

$$= \sum_{i=1}^{n} \rho\left(\frac{z_1 + z_2}{2} - a_i, 0\right)$$

$$= \sum_{i=1}^{n} \rho\left(\tfrac{1}{2}(z_1-a_i) + \tfrac{1}{2}(z_2-a_i), 0\right)$$

$$\leq \frac{1}{2}\sum_{i=1}^{n} \rho(z_1-a_i, 0) + \frac{1}{2}\sum_{i=1}^{n} \rho(z_2-a_i,0) \quad \text{(triangle law)} \tag{6}$$

z_1,z_2 and some a_i are non-collinear since we are given that not all the a_i are collinear. Hence the inequality (6) is strict proving that

$$(7) \qquad f\left(\frac{z_1 + z_2}{2}\right) < \frac{1}{2} f(z_1) + \frac{1}{2} f(z_2) = \lambda/2 + \lambda/2 = \lambda.$$

(7) contradicts the assumption that λ is the minimum and also shows that $f(z)$ is strictly convex.

The next few results concern the unique point P which minimises $\rho(P,A) + \rho(P,B) + \rho(P,C)$ where A,B,C are distinct points of M^2. We shall use the following abbreviations:

$P = \min \rho\{ABC\}$ for the above point, $Q\{ABC\}$ for the sum $\rho(Q,A) + \rho(Q,B) + \rho(Q,C)$ and AB for $\rho(A,B)$.

We omit the proof of the following

Proposition 2. Let $P = \min \rho\{ABC\}$ then P lies within or on the triangle ABC.

Proposition 3. Let $P = \min \rho\{ABC\}$ and suppose $I \notin \{A,B,C\}$. Then if A' is any point of the line PA between P and A, $P = \min \rho\{A'BC\}$.

Proof. Suppose the contrary and $\min \rho\{A'BC\} = Q \neq P$. Then

$$A'Q + BQ + CQ < A'P + BP + CP.$$

$$\therefore \quad (AA' + A'Q) + BQ + CQ < A'P + BP + CP + AA' = AP + BP + CP.$$

Applying the triangle inequality on the left we obtain $Q\{ABC\} < P\{ABC\}$, contradicting the hypothesis $P = \min \rho\{ABC\}$.

Proposition 4. Let $P = \min \rho\{APB\}$. Then for all A',B' on PA, PB on the same side of P as A,B respectively $P = \min \rho\{A'PB'\}$.

Proof. Case (i) Let A',B' be between A and P, B and P respectively and suppose the contrary. Then there exists Q such that

$$A'Q + B'Q + PQ < A'P + B'P.$$

$$\therefore \quad (A'Q + AA') + (B'Q + BB') + PQ < (A'P + AA') + (B'P + BB').$$

Application of the triangle inequality to the terms bracketed on the left gives $Q\{ABP\} < P\{ABP\}$ contradicting the minimum property of P.

Case (ii) If conditions of Case (i) are not satisfied, draw A"B" parallel to A'B' with A",B" satisfying these conditions (see Fig. 7). Then by Case (i) P = min ρ {A"PB"} and by similar triangles P = min ρ {A'B'P} as required.

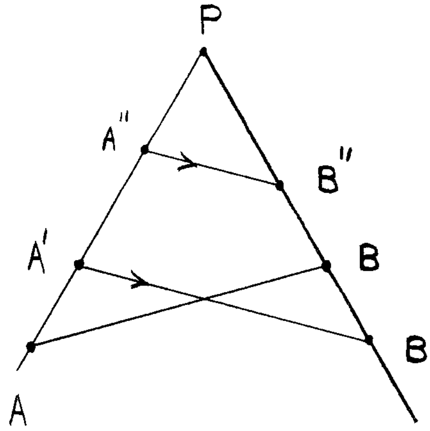


Fig. 7.

<u>Proposition 5.</u> Let min ρ {ABC} = B and A' be any point on the line AC but not in the closed segment AC. Then B = min ρ {A'BC}.

<u>Proof</u> Take points R,S,T on BA', BA, BC respectively such that BR = BS = BT. By hypothesis and Proposition 4, min ρ {BST} = B. Suppose the contrary of the Proposition i.e. min ρ {A'BC} $\neq$ B. This implies by Proposition 4 that

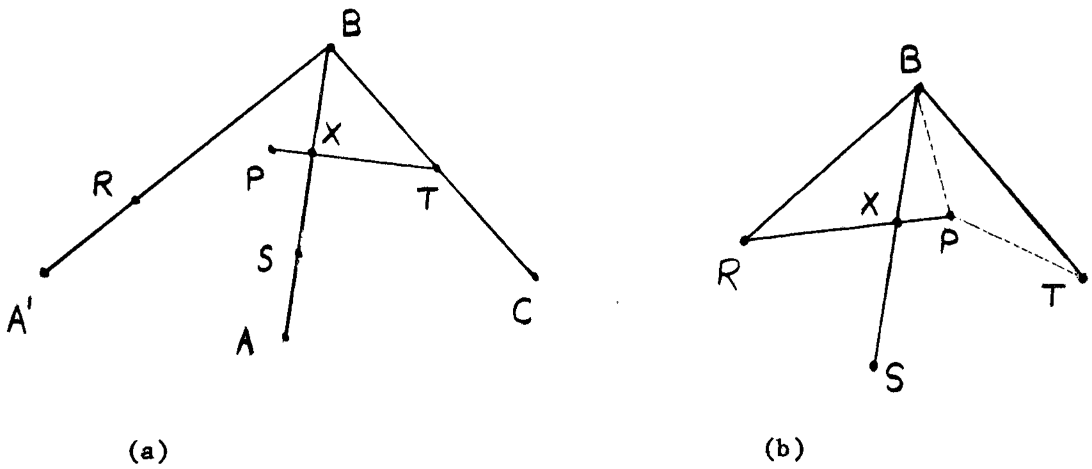


Fig. 8.

$$(8) \qquad \min \rho\{BRT\} = P \neq B.$$

There are two cases to consider:

Case (i) P lies on or to the left of BS (see fig. 8a). Let PT meet BS at X, (internally since the unit ball Σ of the metric is strictly convex). Then

$$\begin{aligned} BX + SX + TX = BS + TX = BR + TX &\leq BR + TP \\ &\leq BP + RP + TP \quad \text{by the triangle inequality.} \\ &< BR + BT \quad \text{by Equation (8).} \\ &= BS + BT \;. \end{aligned}$$

i.e. $X\{BST\} < B\{BST\}$. Contradiction since $B = \min \rho\{BST\}$.

Case (ii) P lies to the right of BS (fig. 8b). Let PR meet BS at X. Now $P = \min \rho\{BRT\}$ implies by Proposition 3 that $P = \min \rho\{BXT\}$. Using the uniqueness property, $B \neq \min \rho\{BXT\}$ and therefore by Proposition 4. $B \neq \min \rho\{BST\}$. Contradiction. Cases (i) and (ii) together prove the proposition.

<u>Definition.</u> Let PA_1, PA_2, PB_1, PB_2 be four lines from P meeting a line ST not through P in A_1, A_2, B_1, B_2 respectively such that B_1 and B_2 are contained in the closed segment A_1A_2. We say that $\measuredangle A_1PA_2$ contains $\measuredangle B_1PB_2$ and properly if B_1 or B_2 or both is in the open segment A_1A_2.

<u>Proposition 6.</u> (i) If $P = \min \rho\{B_1PB_2\}$ and $\measuredangle A_1PA_2$ contains $\measuredangle B_1PB_2$ then $P = \min \rho\{A_1PA_2\}$

(ii) If $B = \min \rho\{ABC\}$ and P is any point within or on triangle ABC except on AC, then $P = \min \rho\{APC\}$.

<u>Proof</u> (i) is immediate from the above definition and Proposition 5. To prove (ii) draw parallels to AB, BC through P and let these meet AC in X and Y. Since $\measuredangle XPY$ is a translation of $\measuredangle ABC$, $P = \min \rho\{PXY\}$. Then by part (i), $P = \min \rho\{APC\}$ as asserted.

<u>Proposition 7.</u> If $C = \min \rho\{ABC\}$ there exists A' between A and B on the line AB such that for all X on this line between A' and B, $\min \rho\{XBC\} \neq C$.

<u>Proof</u> Let the metric ball centre B through C cut AB at A'. A' is between A and B since AB is the "longest" (in the sense of the metric) side. Then for all X between A' and B, $XB < A'B$.

$\therefore \quad XC + XB < XC + A'B = XC + BC.$ (by construction $A'B = BC$).

i.e. $X\{XBC\} < C\{XBC\}$ hence $C \neq \min \rho\{XBC\}$.

<u>Definition</u> $\measuredangle$ ABC is a critical angle if and only if

(i) $\min \rho\{ABC\} = B$

and (ii) $\min \rho\{A'BC'\} \neq B$ for any A', C' such that $\measuredangle$ A'BC' is properly contained in $\measuredangle$ABC.

Propositions 4 and 7 show that this definition is meaningful and critical angles exist. For the Euclidean metric, critical angles are 120° angles. In Minkowski spaces critical angles will vary in their Euclidean magnitude.

<u>Proposition 8.</u> (i) $C = \min \rho\{ABC\}$ if and only if $\measuredangle$ACB contains a critical angle.

(ii) In any triangle exactly one angle or no angle contains a critical angle.

(iii) The angle vertically opposite a critical angle is itself critical.

<u>Proof</u> (i) follows from the definition, (ii) is immediate from the definition and the uniqueness of the minimum point. (iii) is proved using similar triangles. We now prove our principal result on the minimising point of a triangle:

<u>Theorem 7.</u> For any triangle ABC in M^2 exactly one of the following occurs:

Either (i) Exactly one angle (say $\measuredangle$ BAC) contains a critical angle and $A = \min \rho\{ABC\}$

or (ii) There exists a unique point P at which the sides of the triangle subtend critical angles and $P = \min \rho\{ABC\}$.

Proof (i) Suppose $\measuredangle$BAC contains a critical angle. Then no other angle of ABC contains a critical angle and A = min ρ\{ABC\}. (Proposition 8)

(ii) Suppose no angle of ABC contains a critical angle, then P = min ρ\{ABC\} is not at a vertex. Now assume that $\measuredangle$APB does not contain a critical angle. Then

$$\text{(9)} \qquad \min \rho\{APB\} = X \neq P.$$

$$CX \leq PX + PC. \qquad \text{(triangle inequality)}$$

$$\begin{aligned} \therefore \quad AX + BX + CX &\leq AX + BX + PX + PC \\ &< AP + BP \qquad + PC . \qquad \text{by (9).} \end{aligned}$$

i.e. X \{ABC\} < P\{ABC\} contradicting the minimum property of P. Therefore $\measuredangle$APB (and similarly $\measuredangle$APC, $\measuredangle$BPC) contains a critical angle.

We continue the proof by showing that there exists only one point at which each angle subtended by a side of the triangle contains a critical angle. Suppose the contrary, then there exist two such points namely Q and P = min ρ\{ABC\}. We consider two separate cases:

Case 1. Assume P is strictly inside triangle BQC and AP meets QC in X. Then since Q = min ρ\{QBC\}, Proposition 6 implies that X = min ρ\{XBC\}. But P = min ρ\{ABC\}. Therefore, by Proposition 3, P = min ρ\{XBC\} which contradicts the uniqueness of min ρ\{XBC\}. Proofs for P strictly inside AQC or AQB are similar.

Case 2. Assume P lies in the open segments QA. $\measuredangle$AQB contains a critical angle, therefore $\measuredangle$APB properly contains a critical angle and there exists a point R on the open segment CP such that $\measuredangle$ARB contains a critical angle. $\measuredangle$s ARC, BRC also contain critical angles (Proposition 6), hence each angle subtended at R by a side of $\triangle$ ABC contains a critical angle. We now apply Case 1 using R instead of Q and obtain a similar contradiction.

To complete the proof of Theorem 7 we have only to show that if the angles subtended at a point P by the sides of triangle ABC, contain critical angles then these angles are exactly critical angles. Suppose the contrary and $\measuredangle$APB properly contains a critical angle. Then by

Propositions 5 and 7 there is a point D on AB such that for all A' in the closed segment AD, the angles subtended at P by A'B, BC, A'C each contain a critical angle and hence for all such A', $P = \min \rho\{A'BC\}$ which is impossible with a strictly convex unit ball. Hence the Theorem.

<u>Proposition 9.</u> If of the three angles subtended by AB, AC, CA at P, two are critical, then the third is critical.

<u>Proof.</u> Suppose $\measuredangle$ APC, $\measuredangle$ BPC are critical and $\measuredangle$ APB does not contain a critical angle. Then $\min \rho\{APB\} = X \neq P$ and by Theorem 7 the angles PXB, PXA, AXB are critical. Draw parallels to AX, BX through P and produce XP. We see that one of the critical angles APC, BPC (in Fig. 9 it is $\measuredangle$ APC) properly contains a critical angle which is impossible. Therefore $\measuredangle$ APB contains a critical angle and by Theorem 7 it is a critical angle.

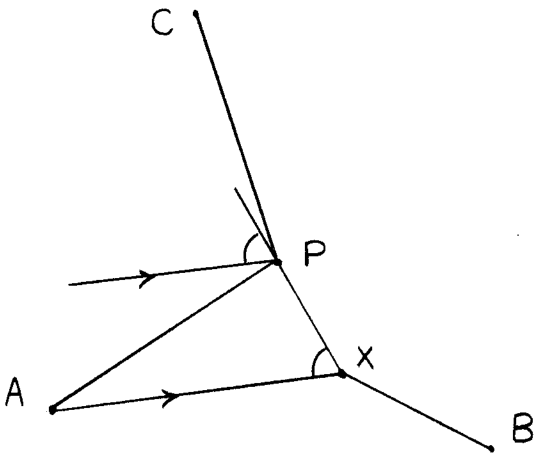


Fig 9.

Corollary 1. If of 3 angles at a point, 2 contain critical angles, one of them properly, then the third angle does not contain a critical angle.

Corollary 2. Supplementary angles cannot both contain critical angles.

Proof Suppose the contrary and let $\measuredangle$ AOB, $\measuredangle$ BOC be supplementary angles each containing a critical angle (Fig. 10). Take any point A' within $\measuredangle$ AOB' as shown. Then $\measuredangle$ A'OB properly contains the critical angle AOB and $\measuredangle$ BOC contains a critical angle by hypothesis. Hence by Corollary 1, $\measuredangle$ A'OC does not contain a critical angle. But the critical angle B'OC (Proposition 8 iii) is contained in $\measuredangle$ A'OC. Contradiction.

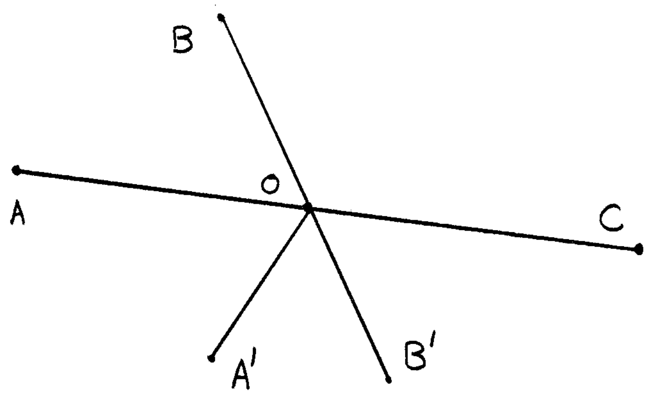


Fig. 10.

We digress and state the following fact which will be used in the next section. Suppose A,B,C are distinct points in an m-dimensional Minkowski Metric Space with strictly convex defining surface S and let the plane π defined by A,B,C meet S in the curve Σ . Then the above theory holds in the plane Minkowski space defined on π by Σ i.e. we can apply Propositions 1-9 and Theorem 7 to three points A,B,C in an m-dimensional Minkowski Metric Space.

<u>Theorem 8.</u> Let U be a minimising tree of (S_n) in M^2 with additional vertices $s_1,\ldots,s_k$. Then U has the properties P1-P4 of section 3 and the following analog of P5:

P"5. For each $i = 1,\ldots,k$, $s_i = \min \rho\{xyz\}$ where x,y,z are the points of $\{s_i\}$ and each angle xs_iy, ys_iz, zs_ix is a critical angle.

<u>Proof</u> Suppose branches x_1y_1, x_2y_2 intersected at p (not a vertex of U) then some angle at P say x_1px_2 does not contain a critical angle. (Proposition 9, Corollary 2). Therefore $\min \rho\{x_1px_2\} = z \neq p$ and a replacement of x_1p, x_2p by the three lines x_1z, x_2z, pz shortens the assumed minimising tree which proves P1. Similarly, no vertex x of U has $w(x) > 3$. Thus U satisfies P3 and $w(s_i) \leq 3$ for all $i = 1,\ldots,k$. There is no gain in introducing additional vertices with valency < 3. Hence P2. The proof of P4 is identical to that given for (S_n) in E^2 and P"5 is the result of theorem 7.

It is shown in [4] that for (S_n) in the Manhattan metric, where the defining curve is not strictly convex, the property $w(x) \leq 3$ for each vertex x of a minimising tree does not hold.

The equilateral construction of the Steiner point of a triangle in E^2 enabled us to solve the problem (S_n). Accordingly one is lead to search for a generalisation of this construction for $P = \min \rho\{ABC\}$ in M^2 when no angle of $\triangle ABC$ contains a critical angle. The following two conjectures were made:

(i) Let (AB) be the third vertex of the triangle built outward on AB whose exterior angles are critical angles. Then the line (AB)C passes through P.

(ii) A triangle XYZ in M^2 is ρ- equilaterial if $\rho(X,Y) = \rho(Y,Z) = \rho(Z,X)$. Let (AB) be the third vertex of the ρ- equilateral triangle built outward on AB. Then (AB)C passes through P.

We note that in E^2 (i) and (ii) are equivalent and may be used to construct the Steiner point. The counter examples below show that in M^2

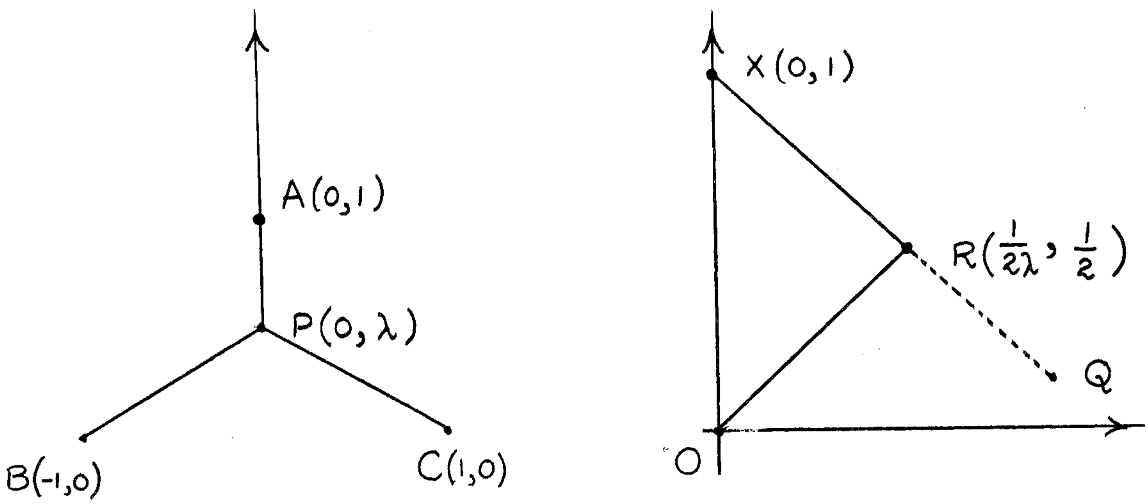
the conjectures are not equivalent and that neither is true in general. The examples will use the metric with unit ball $|x|^p + |y|^p = 1$, which is strictly convex if $p > 1$ Then $\rho\{(x_1,y_1), (x_2,y_2)\} = \{|x_2-x_1|^p + |y_2-y_1|^p\}^{1/p}$. Let A = (0,1) , B = (1,0) , C = (-1,0). Symmetry and uniqueness insist that P = min $\rho\{ABC\}$ lies on the y-axis and by elementary methods, the sum $\rho(A,P) + \rho(B,P) + \rho(C,P)$ takes its minimum value at the point $P(0,\lambda)$ where

$$\lambda = \left(\frac{1}{2^{p/p-1}-1} \right)^{1/p} .$$

Recall that if A is moved to any point A' on PA on the same side of P then P = min $\rho\{A'BC\}$.

Example 1. To show that a triangle whose exterior angles are critical, is not necessarily ρ - equilateral.

Using the special case given above, the three angles at P in fig. 11(a) are critical angles since P = min $\rho\{ABC\}$. AP, BP, CP have slopes ∞, λ, $-\lambda$. The sides of triangle ORX shown in fig. 11(b) have identical slopes and hence the exterior angles of this triangle are critical angles e.g. $\measuredangle$ ORQ is a translation of the critical angle BPC. We show that triangle ORX is not ρ - equilateral.



(a) (b)

Fig. 11

$$\rho\,(PX) = \left\{\left(\frac{1}{2}\right)^p + \left(\frac{1}{2\lambda}\right)^p\right\}^{1/p} = \frac{1}{2}\left(1 + \frac{1}{\lambda^p}\right)^{1/p}$$

$$= \frac{1}{2}\left(1 + \left(2^{p/p-1} - 1\right)\right)^{1/p} = 2^{(2-p)/(p-1)}$$

But $\rho\,(OX) = 1$. Hence triangle is ρ-equilateral if and only if $2^{(2-p)/(p-1)} = 1$.

i.e. if and only if $p = 2$ and the metric is Euclidean.

<u>Example 2.</u> To show conjecture (i) is false.

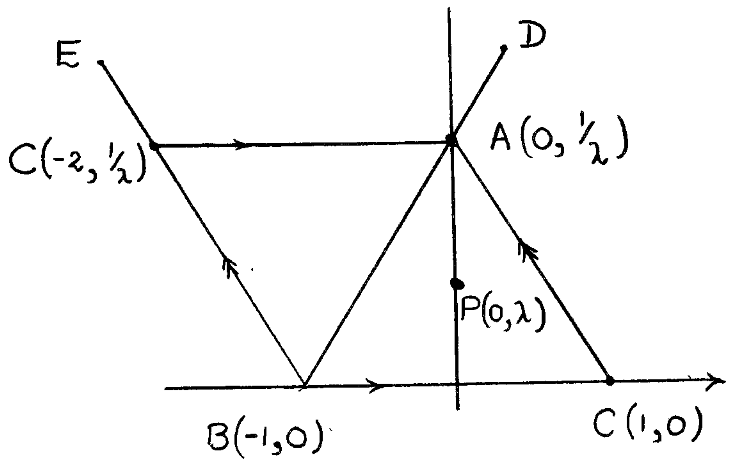


Fig. 12

In Fig. 12 $P = \min \rho\,\{ABC\}$; the exterior angles DAC', EC'A of triangle ABC' are translations followed by a rotation through 90° of the angles APC, APB. Since each such transformation is an isometry for this metric, the angles DAC', EC'A are critical angles and hence triangle ABC' has its exterior angles critical. We show that CC' does not pass through P.

CC' has equation $3\lambda y = 1 - x$ and passes through $(0,\lambda)$ if and only if $3\lambda^2 = 1$ or $\lambda = \frac{1}{\sqrt{3}}$ i.e. CC' passes through P if and only if $p = 2$ and the metric is Euclidean.

Example 3. To show conjecture (ii) is false.

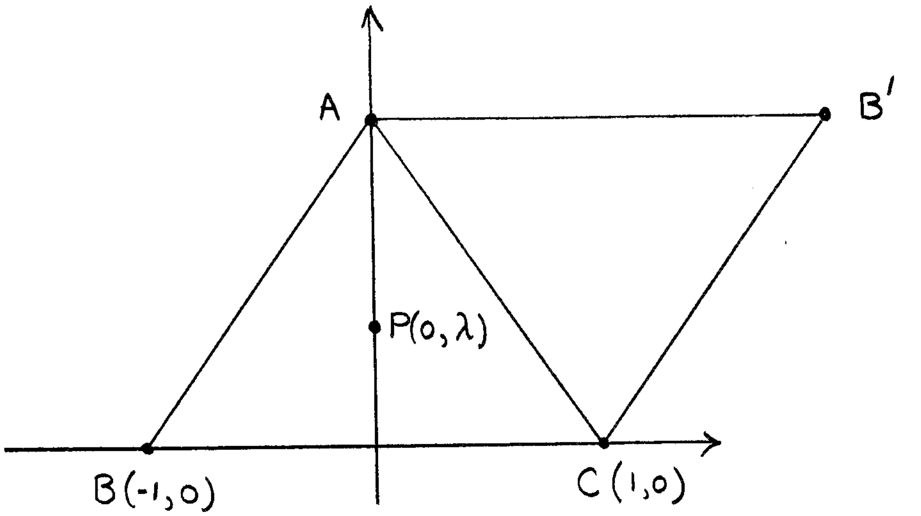


Fig. 13

In Fig. 13 A,B' have co-ordinates $(0,(2^p-1)^{1/p})$, $(2,(2^p-1)^{1/p})$. Then it is easily verified that triangle AB'C is ρ - equilateral with sides of length 2. We show that BB' does not in general pass through P = min ρ {ABC}. BB' has equation $y = \frac{1}{3}(2^p-1)^{1/p}(x+1)$ hence has y-intercept $(2^p-1)^{1/p}/3$. $\therefore$ BB' passes through P if and only if

$$(2^p-1)^{1/p}/3 = (2^{p/p-1}-1)^{-1/p}$$

or

$$(2^p-1)(2^{p/p-1}-1) = 3^p .$$

We see that this is satisfied for the Euclidean case $p = 2$ but is certainly false for any integer $p > 2$.

VI - MORE PROPERTIES OF $(S_{n\alpha\beta\gamma})$

In this chapter we prove two results for $(S_{n\alpha\beta\gamma})$, concerning finiteness of solution in Minkowski Space and non-constructibility of solutions in E^2 by Euclidean constructions.

The following threorem uses the notation of Theorem 2. Chapter II.

Theorem 9. If M is a Minkowski Metric space M_m for which the defining surface Σ is strictly convex, then there is a unique tree of minimum length in $C \in \{C_1,\ldots,C_N\}$.

Proof Suppose the contrary and f has minima, value ℓ, at $\{s_1,\ldots,s_k\}$ and $\{t_1,\ldots,t_k\}$ where $t_i \neq s_i$ for some i. Consider the set $\left\{\frac{s_1+t_1}{2},\ldots,\frac{s_k+t_k}{2}\right\}$.

Using Properties (i) - (iii) of Minkowski Spaces Chapter V.

$$\begin{aligned}\rho\left(\frac{s_i+t_i}{2}, a_{ij}\right) &= \rho\left(0, \frac{s_i-a_{ij}}{2}+\frac{t_i-a_{ij}}{2}\right)\\ &\leq \rho\left(0, \frac{s_i-a_{ij}}{2}\right)+\rho\left(0, \frac{t_i-a_{ij}}{2}\right)\\ &= \frac{1}{2}\rho(0, s_i-a_{ij})+\frac{1}{2}\rho(0, t_i-a_{ij})\\ &= \frac{1}{2}\rho(s_i,a_{ij})+\frac{1}{2}\rho(t_i,a_{ij}).\end{aligned}$$

This inequality is strict unless a_{ij}, s_i, t_i are collinear with s_i, t_i on the same side of a_{ij}.

$$\therefore \quad \sum_{i=1}^{k}\sum_{j=1}^{\lambda_i}\rho\left(\frac{s_i+t_i}{2}, a_{ij}\right) \leq \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{\lambda_i}\left\{\rho(s_i,a_{ij})+\rho(t_i,a_{ij})\right\}$$

and the inequality is strict unless for each $i = 1,\ldots,k$, $a_{i1}, a_{i2},\ldots, a_{i\lambda_i}$, s_i, t_i are collinear with s_i, t_i occupying suitable positions on the line. Such a situation cannot occur in a minimum length tree of C. Assuming $n > 2$ (otherwise the problem is trivial), there exists i for which $\lambda_i \geq 2$ and s_i joins only one other s-point. If $\lambda_i > 2$ a simple application of the triangle inequality proves that the assumed tree could not be minimum length in C and the case $\lambda_i = 2$ is disposed of using Proposition 9, Corollary 2. Thus we can conclude

$$(10) \qquad \sum_{i=1}^{k} \sum_{j=1}^{\lambda_i} \rho\left(\frac{s_i+t_i}{2}, a_{ij}\right) < \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{\lambda_i} \rho(s_i, a_{ij}) + \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{\lambda_i} \rho(t_i, a_{ij}).$$

By a similar use of the properties of M^m we can show

$$\sum_{(i,j)\in R_1} \rho\left(\frac{s_i+t_i}{2}, \frac{s_j+t_j}{2}\right) \leq \frac{1}{2} \sum_{(i,j)\in R_1} \rho(s_i, s_j) + \frac{1}{2} \sum_{(i,j)\in R_1} \rho(t_i, t_j).$$

Adding this to (10) and $\sum_{(i,j)\in R_2} \rho(a_i, a_j)$ to both sides we obtain

$$f\left(\frac{s_1+t_1}{2}, \ldots, \frac{s_k+t_k}{2}\right) < \frac{1}{2} f(s_1, \ldots, s_k) + \frac{1}{2} f(t_1, \ldots, t_k) = \ell \text{ which}$$

contradicts the minimum property of ℓ.

Corollary In M^m, $(S_{n\alpha\beta\gamma})$ and (S_n) have a finite number of minimising trees.

Proof Immediate from Theorems 1,2 and 9.

Similar proofs to those given here may be used to establish identical results when the function to be minimised is

$$F\left(L(U), \sum_{i=1}^{n} w(a_i), \sum_{j=1}^{k} w(s_j), k\right)$$

where F is any positive function which is strictly increasing in each of its four variables.

It was demonstrated in Chapter 1 that $(S_{n\alpha\beta\gamma})$ reduces to (P_n) for suitable values of the constants α, β, γ. We show now that in general (P_n) in E^2 (and hence $(S_{n\alpha\beta\gamma})$) is not solvable by Euclidean constructions.
We use n = 5 for our example since (P_3) is solvable by Euclidean constructions (by our equilateral construction). The solution of (P_4) is the intersection of the diagonals if the configuration is convex and the vertex interior to

the convex hull otherwise. We take 5 points A_i ($i = 1,\ldots,5$) symmetrically placed with respect to the x-axis as shown in fig. 14.

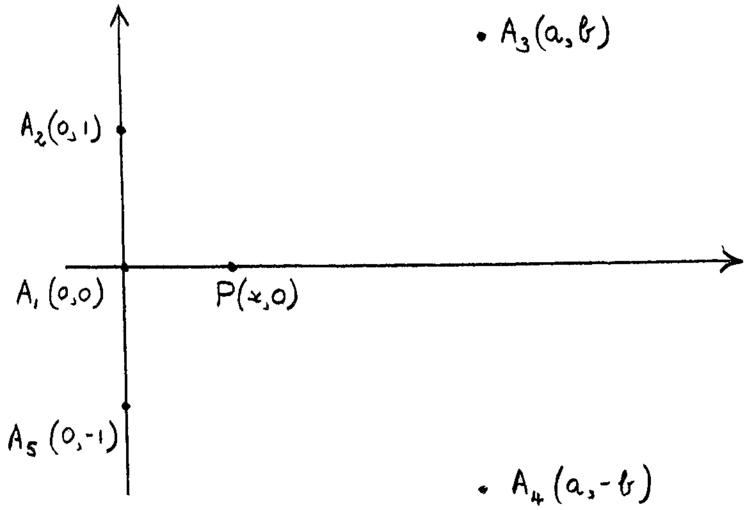


Fig. 14.

The minimum point P lies on the x-axis with its co-ordinate in [0,a]

$$\sum_{i=1}^{5} PA_i = x + 2\sqrt{1+x^2} + 2\sqrt{b^2+(a-x)^2} .$$

Minimising this function by elementary methods, we find that the co-ordinate x of P satisfies an eighth degree polynomial equation $f(x) = 0$ whose coefficients are polynomials in a and b.

We show that for suitable integers a,b, $f(x)$ is irreducible over the rationals and $f(x) = 0$ has Galois group over the rationals which does not have order 2^k where k is a positive integer. Therefore x is not an element belonging to an extension field of the rationals of degree 2^k and hence the segment OP is not constructible by Euclidean constructions (See [10] page 185). i.e. for suitable choices of the five points (P_5) is not solvable by Euclidean constructions.

The leading coefficient of $f(x)$ is 15. In order to use the theory of [10] page 190-191, we need to work with a monic polynomial and therefore make the transformation $x = y/15$ and multiply the equation through by 15^7 thus obtaining equation $g(y) = 0$ where $g(y)$ is monic. We note that such a transformation does not affect reducibility over the rationals or the Galois group of the equation. The coefficients of $g(y)$ are:

$$
\begin{array}{lcl}
y^8 & : & 1 \\
y^7 & : & -60a \\
y^6 & : & 15\ (90a^2 + 22b^2 + 22) \\
y^5 & : & -15^2\ (88a + 60a^3 + 44ab^2) \\
y^4 & : & 15^3\ (42 + 15a^4 + 154a^2) \\
y^3 & : & -15^4\ (88a^3 + 120ab^2 - 36a) \\
y^2 & : & 15^5\ (22a^4 + 60a^2b^2 + 6b^4 + 6b^2 - 54a^2) \\
y & : & 15^6 \cdot 12a \cdot (3a^2 - b^2) \\
1 & : & -15^7\ (b^2 - 3a^2)^2 .
\end{array}
$$

We take $a = 10, b = 3$ and notice that the coefficients of y^8 and the constant are odd while the rest of the coefficients are even, so that Eisenstein's irreducibility criterion using the prime 2 shows that $g(y)$ is irreducible over the rationals and hence $g(y) = 0$ has no multiple root. Let $I/(7)$ be the field of residue classes of integers modulo 7. Over this field $g(y) = 0$ reduces to the equation

$$g_7(y) = y^8 + 2y^7 + y^6 + y^5 + 4y^4 + y^3 + 4y^2 + 4y + 5 = 0 .$$

The greatest common divisor of $g_7(y)$ and its derivative over $I/(7)$ in 1, hence $g_7(y) = 0$ has no multiple root. $g_7(y)$ has the following factorisation mod 7:

$$(y^3 + 5y^2 + 4y + 2)(y^5 + 4y^4 + 5y^3 + 4y + 6) \tag{11}$$

and the cubic factor is irreducible mod 7. Therefore the Galois group of the equation $g(y) = 0$ contains a permutation α whose representation as a product of disjoint cycles contains a cycle of order 3. α does not have order 2^k for any positive integer k, hence the Galois group does not have order 2^k and the proof is complete. We state the result formally:

<u>Theorem 10.</u> (P_n) and $(S_{n\alpha\beta\gamma})$ in E^2 are not, in general, solvable by Euclidean constructions.

Finally, we show how the above example was found. 7 is the smallest prime which can be used. (It is clear from the coefficients that g(y) has multiple roots modulo 2, 3 and 5). One therefore searches for values of a and b such that $g_7(y) = 0$ has no multiple root and an irreducible cubic, quintic or sextic factor mod 7. The arithmetic being mod 7, it suffices to consider a,b = 0,...,6. In fact since b only occurs in the coefficients as b^2, we need only take b = 0,1,2,3 as $b^2 \equiv (7-b)^2$ mod 7. The 28 polynomials were tested for an irreducible cubic factor using a digital computer. The method used was to divide each polynomial by the finite number of irreducible cubics over I/(7) and investigate the remainder. The values a = 3, b = 3 yielded the factorisation (11). However Eistenstein's criterion works conveniently for g(y) with prime 2 if a is even and b is odd so we take a = 10, b = 3. The polynomial obtained has the same reduction (11) mod 7.

This method of reduction mod p in conjunction with a computing machine can provide much information on the structure of Galois groups of equations.

BIBLIOGRAPHY

1. R. Courant and H. Robbins: What is Mathematics? (New York 1941).

2. Z. A. Melzak: On the Problem of Steiner. Canadian Math. Bulletin, Vol. 4 No. 2. May 1961.

3. R. C. Prim: Shortest Connecting Networks. Bell System Tech. Journal 31, pp 1398 - 1401.

4. M. Hanan: On the Steiner Problem with Rectilinear distance. S.I.A.M. Journal of Applied Mathematics, Vol. 14 No. 2. March 1966.

5. F. P. Palermo: A Network Minimisation Problem. I.B.M. J. of Research and Development 5 (1961) pp 335 - 337.

6. R. L. Francis: A note on Optimum Location of New Machines in Existing Plant Layout. J. Indus. Eng. 14 (1963) pp 57 - 59.

7. Hua-lo-keng et al: Application of Math. Methods to Wheat Harvesting. Chinese Math. 2 (1962) pp 77 - 91.

8. C. E. Wetherburn: Diff. Geometry of 3 Dimensions. (Cambridge 1947)

9. L. M. Blumenthal: Theory and Applications of Distance Geometry. (Oxford 1953)

10. B. L. Van der Waerden: Modern Algebra Vol. 1. (New York 1949)