ANNALES
POLONICI MATHEMATICI
XLIV (1984)

On the shortest curve which meets all the lines which meet a circle

by VANCE FABER (Denver, Colorado),
JAN MYCIELSKI and PAUL PEDERSEN (Boulder, Colorado)

Abstract. It is proved that the shortest path p: $[0, 1] \to \boldsymbol{R}^2$ which intersects every straight line which intersects the unit circle is one-to-one and has length $\pi + 2$ (see Figure 16).

1. Introduction. M. Magidor has related the following story to us. A telephone company, while repairing buried cable, has discovered that the cable is often not directly under the marker erected above it. To insure finding the cable, even when its direction is unknown, they have their repairmen dig in a circle of radius 2 meters about the marker. Magidor notes that this is not the most efficient way of finding the cable (see Figure 16), and asks what is the length of the shortest trench which will find the cable (assuming that the cable is straight and *does* pass within 2 meters of the marker).

The following mathematical problem arises: what is the shortest curve which meets all the lines which meet a given circle (or even more generally, any convex set)? Of course, the answer to this question may depend upon our definitions of the words "curve" and "length".

DEFINITION. Let S be a Borel subset of a metric space M. The length of S is the 1-dimensional Hausdorff measure of S, $\lambda(S) = \lambda_1(S)$, where the α-dimensional Hausdorff measure of S is defined by

$$\lambda_\alpha(S) = \lim_{\delta \to 0} \Big(\inf \Big\{ \sum_{i=1}^{\infty} (\operatorname{diam} E_i)^\alpha \Big| \bigcup_{i=1}^{\infty} E_i = S \text{ and } \operatorname{diam} E_i \leqslant \delta \text{ for all } i \Big\}\Big).$$

We shall mainly be concerned with the case $M = \boldsymbol{R}^n$.

DEFINITION. A *path* is a continuous function f: $I \to \boldsymbol{R}^n$, where $I = [0, 1]$. The *path length* of f is

$$a(f) = \sup \Big\{ \sum_{k=0}^{n-1} \|f((k+1)/n) - f(k/n)\| \ \Big| n = 1, 2, \ldots \Big\}.$$

DEFINITION. A path f is *closed* if $f(0) = f(1)$. A path f is *simple* if f is 1–1 on $(0, 1)$.

PROPOSITION. (i) *For every path f we have $\lambda(fI) \leqslant a(f)$.*

(ii) *If f is a simple path, then $\lambda(fI) = a(f)$.*

DEFINITION. An *n-path* is a sequence $C = (f_1, f_2, \ldots, f_n)$, where each f_i is a path. Its n-path length is $a(C) = \sum_{i=1}^{n} a(f_i)$.

For any sets $R, S \subseteq \mathbf{R}^2$ we write:

$P(R, S) \Leftrightarrow$(every line meeting R meets S).

con S = (the convex hull of S).

∂S = (the boundary of S).

int S = (the interior of S).

A path f is *convex* iff $fI \subseteq \partial\,\mathrm{con}\, fI$. $f|X$ denotes the restriction of f to X.

2. Discussion of results. Most of the paper is devoted to proving (Theorem 7) that a shortest path which meets (i.e., whose image meets) all the lines which meet a given compact convex set is one-to-one. We then show (Theorem 8) that the shortest path which meets all the lines which meet the unit circle C has path length $\pi+2$ (Figure 16). We have not succeeded in proving that this path yields the shortest (in terms of Hausdorff length) closed connected set S satisfying $P(C, S)$; see Conjectures 1 and 2. We do not even know if it is enough to consider a bounded set in the plane in order to find the shortest 2-path which meets all the lines which meet the unit circle (see Conjectures 3 and 4).

A similar problem is: Does there exist a closed set $S \subseteq \mathbf{R}^3$ of minimal area (that is, 2-dimensional Hausdorff measure) which intersects every line which intersects the unit sphere?

R. Laver found the following example: For every $\varepsilon > 0$ there exists a set S with the above property, the area of which is less than $2\pi+\frac{1}{2}\pi^2+\varepsilon$. His set S consists of the lower hemisphere (of area 2π) plus vertical rings standing on the surface of the sphere. The first ring stands on the equator, the second ring stands on the circle formed by the intersection of the plane of the top of the first ring and the sphere, etc. The last ring extends to the height of the north pole. (See Figure 0.) Since $\int_0^{\pi/2} 2\pi \cos \alpha d(\sin \alpha) = \frac{1}{2}\pi^2$, it is clear that if the consecutive rings are narrow enough, then their joint area is less than $\frac{1}{2}\pi^2+\varepsilon$. It is also clear that every line which intersects the sphere intersects S. We do not know if there exists a surface of area $2\pi+\frac{1}{2}\pi^2$ with the requisite property.

3. Computation of lower bounds. Consider the projection function π_θ which projects each point (x, y) in the plane onto the line L_θ through the origin and making an angle θ with the x-axis, that is,

$$\pi_\theta(x, y) = x \cos \theta + y \sin \theta.$$

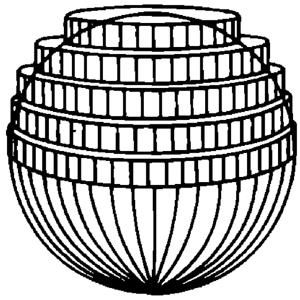


Fig. 0

THEOREM 1. *Let f be a path. Then*

$$\int_0^\pi a(\pi_\theta f)\, d\theta = 2a(f).$$

Proof. Simple computation shows that this formula holds if f is a line segment. Let $\{f_n\}$ be a sequence of polygonal paths with vertices on fI whose limit is f. The result follows from the facts:

$$\lim a(f_n) = a(f), \quad \lim \pi_\theta f_n = \pi_\theta f, \quad \lim a(\pi_\theta f_n) = a(\pi_\theta f)$$

and the monotone convergence theorem.

Remark 1. In [6], p. 124, this theorem is attributed to Cauchy. Also see [7] and [8].

COROLLARY. *Let S be a connected compact set and let f be a simple closed path such that $fI = \partial$ con S. Then f has smallest $a(f)$ among all closed paths f with $P(S, fI)$.*

Proof. Since f is closed,

$$a(\pi_\theta f) \geqslant 2\lambda(\pi_\theta f).$$

Since $P(S, fI)$, $\lambda(\pi_\theta f) \geqslant \lambda(\pi_\theta S)$. Thus $a(\pi_\theta f) \geqslant 2\lambda(\pi_\theta S) = 2\lambda(\pi_\theta \operatorname{con} S) = a(\pi_\theta \partial \operatorname{con} S)$. Hence

$$2a(f) = \int_0^\pi a(\pi_\theta f)\, d\theta \geqslant \int_0^\pi a(\pi_\theta \partial \cos S)\, d\theta = 2a(\partial \operatorname{con} S),$$

and so $a(f) \geqslant a(\partial \operatorname{con} S)$.

THEOREM 2. (See [8].) *Let S be a closed set. Then*

$$\int_0^\pi \lambda(\pi_\theta S)\, d\theta \leqslant 2\lambda(S).$$

Proof. If $\lambda(S)$ is infinite, there is nothing to prove. We suppose $\lambda(S)$ is finite. By Theorem 1, the inequality holds if S is the countable union of rectifiable simple paths. By [1], p. 304 and 324, and [2], p. 357, $S = S_1 \cup S_2 \cup S_3$ with $\lambda(\pi_\theta S_1) = 0$ for almost all θ, $\lambda(S_2) = 0$ and

$S_3 \subseteq \bigcup_{i<\omega} f_i I$ with $a(f_i) < \infty$ for paths f_i. For each $\varepsilon > 0$ and for each i, we can find simple paths f_{ij} with $\lambda(f_{ij} I) < \infty$ such that $f_i I \cap S \subseteq \bigcup_{j<\omega} f_{ij} I$ and $\lambda(\bigcup_{j<\omega} f_{ij} I - S) < \varepsilon/2^i$. Let $T = \bigcup_{i,j<\omega} f_{ij} I$. Then $S_3 \subseteq T$, $\lambda(\pi_\theta S_2) = 0$, $\lambda(T-S) < \varepsilon$ and

$$\int_0^\pi \lambda(\pi_\theta S)\, d\theta \leqslant \int_0^\pi \lambda(\pi_\theta T)\, d\theta \leqslant 2\lambda(T) \leqslant 2(\lambda(S)+\varepsilon).$$

Corollary. *Let $S = fI$, where f is a convex simple closed path, and let S_0 be a closed set. If $P(S, S_0)$, then*

$$\lambda(S_0) \geqslant \tfrac{1}{2}\lambda(S).$$

Proof. Since $P(S, S_0)$, $\lambda(\pi_\theta S_0) \geqslant \lambda(\pi_\theta S)$. Since f is a convex simple closed path, $a(\pi_\theta f) = 2\lambda(\pi_\theta S)$. Thus

$$2a(f) = \int_0^\pi a(\pi_\theta f)\, d\theta = \int_0^\pi 2\lambda(\pi_\theta S)\, d\theta \leqslant \int_0^\pi 2\lambda(\pi_\theta S_0)\, d\theta \leqslant 4\lambda(S_0),$$

and so $\frac{1}{2}\lambda(S) = \frac{1}{2} a(f) \leqslant \lambda(S_0)$.

Remark 2. For each $\varepsilon > 0$, there exists a convex simple closed path f and a path g such that $P(fI, gI)$ and $a(g)/a(f) \leqslant \frac{1}{2}+\varepsilon$. See Figure 1.

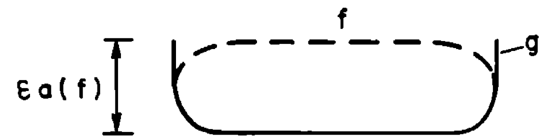


Fig. 1

4. Existence. In our proof of Theorem 5, we shall have to apply the following theorem of Gołąb [3] which depends on former work of Ważewski [11]. We first recall the definition of Hausdorff' metric in the space of compact subsets of a metric space M.

Definition. For each pair of non-empty compact subsets, $P, Q \subseteq M$, $d(P, Q) = \inf\{\eta \mid$ each of P and Q lies in the η-neighbourhood of the other$\}$.

Theorem 3. *If $K_1, K_2, \ldots$ is a sequence of continua (compact connected sets) in a metric space M which converges in Hausdorff' metric to a continuum K, then*

$$\lambda(K) \leqslant \liminf_{n\to\infty} \lambda(K_n).$$

This theorem has been generalized by Vituškin [10] (see also [4]), but this generalization is not needed here. Concerning the proof of the theorem, that of Gołąb and Ważewski is complicated and Vituškin's stronger result is still harder. However, before we knew the above references we found a simple proof of Theorem 3. Then Roy O. Davies found a gap in our proof but also showed us how to correct it (and gave us the above references). For convenience of the reader, this simple self-contained proof will be presented here.

First we need a theorem of Ważewski [11].

THEOREM 4. *If K is a continuum and $\lambda(K) < \infty$, then K is arcwise connected.*

Proof. A well-known theorem (see, for example, [5], § 45, II) asserts that every complete connected and locally connected metric space is arcwise connected. Hence it is enough to prove that K is locally connected. Suppose to the contrary that there exists a point $p \in K$ and an $r > 0$ such that the ball $B(p, r) = \{x \in K | \ d(x, p) \leqslant r\}$ does not contain any connected neighborhood of p. We can assume that $r \leqslant \frac{1}{2}\operatorname{diam} K$. Hence for each n there exists a component C_n of $B(p, r)$ such that $p \notin C_n$ and C_n intersects the ball $B(p, 1/n)$. Hence $\operatorname{diam} C_n \geqslant r - 1/n$ and the sequence $C_1, C_2, \ldots$ contains infinitely many disjoint sets. Hence $\lambda(\bigcup C_n) = \infty$, contrary to the assumption that $\lambda(K) < \infty$.

LEMMA 1. *If K is arcwise connected and $F \subseteq K$ is a finite set, then there exists a tree T (a finite union of arcs which is connected and does not contain any simple closed curve) such that $F \subseteq T \subseteq K$.*

Proof. This is an easy induction on the number of points in F.

LEMMA 2. *For every tree T with $\lambda(T) < \infty$ and every $\delta > 0$, T can be covered with continua (subtrees) $S^1, \ldots, S^m$ such that*

$$\Sigma \operatorname{diam} S^i \leqslant \Sigma\lambda(S^i) = \lambda(T); \tag{1}$$

$$\operatorname{diam} S^i \leqslant \delta \quad \textit{for } i = 1, \ldots, m; \tag{2}$$

$$m \leqslant 2\lambda(T)/\delta + 1. \tag{3}$$

Proof. If $\operatorname{diam} T \leqslant \delta$, then $m = 1$ and $S^1 = T$. Now suppose that $\operatorname{diam} T > \delta$. Fix a point $r \in T$ to be regarded as a root of T. For any two points $p, q \in T$, let $A(p, q)$ be the (unique) arc in T with the ends p, q. Choose $e \in T$ (an end) such that $\lambda(A(r, e))$ is maximum. Then $\lambda(A(r, e)) > \frac{1}{2}\delta$, because otherwise $\operatorname{diam} T \leqslant \delta$. Let $r' \in A(r, e)$ be the (unique) point such that $\lambda(A(r', e)) = \frac{1}{2}\delta$. The part of T "above" r' is a tree S^1 of diameter $\leqslant \delta$ and length $\geqslant \frac{1}{2}\delta$. Also $(T \backslash S^1) \cup \{r'\}$ is a tree T^1. Similarly in T^1 we choose S^2, and so on, until say S^{m-1} leaves a tree T^{m-1} of diameter $\leqslant \delta$ which we call S^m. Then properties (1) and (2) are obvious and (3) follows because $\lambda(S^i) \geqslant \frac{1}{2}\delta$ for $i = 1, \ldots, m-1$ and hence by (1), $(m-1)\frac{1}{2}\delta \leqslant \lambda(T)$.

Proof of Theorem 3. We can assume without loss of generality that $\lambda(K_n) < \infty$ for all n. We choose finite sets $F_n \subseteq K_n$ such that $F_n \to K$ in the Hausdorff metric. By Theorem 4 and Lemma 1, we choose a tree $T_n \subseteq K_n$

such that $F_n \subseteq T_n$. We can assume without loss of generality that $\lambda(T_n) \to t$ for some t. Of course, it will be enough to prove $\lambda(K) \leqslant t$ or, equivalently, that if $\lambda(K) > \alpha$, then $t > \alpha$. Choose $\delta > 0$ such that

$$\Sigma \text{ diam } E_i > \alpha, \tag{4}$$

for every covering $E_1, E_2, \dots$ of K with sets of diameters $\leqslant \delta$. Now perform the decomposition $S_n^1, \dots, S_n^{m(n)}$ of T_n given by Lemma 2. Since $\lambda(T_n) \to t$ and (3), we see that the numbers $m(n)$ are bounded. By choosing a subsequence we can assume that they are all equal, say m. By choosing subsequences again we can assume that the sequences $S_1^i, S_2^i, \dots$ converge to some S^i for $i = 1, 2, \dots, m$. Of course, $S^1 \cup \dots \cup S^m = K$ and diam $S^i \leqslant \delta$, by (2). Hence by (4) and (1)

$$\alpha < \Sigma \text{ diam } S_i = \lim_{n\to\infty} \Sigma \text{ diam } S_n^i \leqslant \lim_{n\to\infty} \lambda(T_n) = t,$$

so $\alpha < t$ as required.

THEOREM 5. *Let n be a positive integer and B a compact connected set in $\boldsymbol{R}^2$. If $S \subseteq B$ and S is compact, then there exists a compact set $S^{(n)}$ with at most n connected components such that $P(S, S^{(n)})$ and*

$$\lambda(S^{(n)}) = \inf\{\lambda(S_0) | \ S_0 \subseteq B,\ P(S, S_0),\ S_0 \textit{ compact with at most } n \textit{ connected components}\}.$$

Proof. The set $K = \{S_0 \subseteq B| \ P(S, S_0), S_0$ compact with at most n connected components$\}$ is compact in the space of compact subsets of $\boldsymbol{R}^2$ with Hausdorff' distance. Hence there exists a sequence $S_i \in K$ $(i = 1, 2, \dots)$ such that $\lim \lambda(S_i) = \inf\{\lambda(S_0)| \ S_0 \in K\}$. $\{S_i\}$ has a convergent subsequence $\{S_{n_i}\}$. Let $S^{(n)} = \lim_{i\to\infty} S_{n_i}$. Then $P(S, S^{(n)})$, $S^{(n)}$ has a most n components and by Theorem 3

$$\lambda(S^{(n)}) \leqslant \lim_{i\to\infty} \lambda(S_{n_i}) = \inf\{\lambda(S_0)| \ S_0 \in K\}.$$

Added in proof. We have recently proved a refinement of Theorem 5 which will appear in the American Mathematical Monthly.

COROLLARY. *There is a shortest connected closed set which meets all the lines which meet a given compact set S.*

Proof. Take B to be a sufficiently large disk with center in S. Then Theorem 5 easily yields the corollary.

THEOREM 6. *Let n be a positive integer and B a compact connected set in $\boldsymbol{R}^2$. If $S \subseteq B$ and S is compact, then there exists an n-path $S^{(n)}$ such that $P(S, S^{(n)})$ and*

$$a(S^{(n)}) = \inf\{a(S_0)| \ S_0 \subseteq B,\ P(S, S_0),\ S_0 \textit{ an } n\textit{-path}\}.$$

The proof is similar to the proof of Theorem 5.

COROLLARY. *There is a shortest path which meets all the lines which meet a given simple closed path.*

5. Computation of minimum path.

THEOREM 7. *Let f be a shortest path which meets all the lines which meet a given convex compact set S. Then f is one-to-one.*

We shall need several lemmas.

LEMMA 3. *For all connected sets S, $P(S, \operatorname{con} S)$ and $P(\operatorname{con} S, S)$.*

LEMMA 4. *If S and S_0 are connected sets, then $P(S, S_0)$ if and only if $S \subseteq \operatorname{con} S_0$.*

The proofs of these lemmas are easy and we omit them.

From now on, assume f and S satisfy the hypothesis of the theorem.

LEMMA 5. *If $f(t_1) = f(t_2)$ and $f|\,[t_1, t_2]$ is not a closed line segment, then $f(t_1) \in \partial \operatorname{con} fI$ and $f|\,[t_1, t_2]$ is a simple convex closed path. If $f|\,[t_1, t_2]$ is a maximal (in fI) closed line segment, again we have $f(t_1) \in \partial \operatorname{con} fI$.*

Proof. Suppose $f(t_1) = f(t_2)$ with $t_1 < t_2$. By continuity, for every $\varrho > 0$ there exists an $\varepsilon > 0$ such that $0 < t_1 - \varepsilon < t_1 + \varepsilon < t_2 - \varepsilon < 1$ and such that $f[t_1 - \varepsilon, t_1 + \varepsilon]$ and $f[t_2 - \varepsilon, t_2 + \varepsilon]$ are contained in $C_\varrho(f(t_1)) = C_\varrho(f(t_2))$, the disk of radius ϱ with center $f(t_1) = f(t_2)$. Suppose $f(t_1) \in \operatorname{int} \operatorname{con} fI$. Then there exists a ϱ such that $C_\varrho(f(t_1)) \subseteq \operatorname{int} \operatorname{con} fI$. Clearly f is a line segment on $[t_1 - \varepsilon, t_1 + \varepsilon]$ and $[t_2 - \varepsilon, t_2 + \varepsilon]$. As short a path as f with the same convex hull is formed by (see Figure 2)

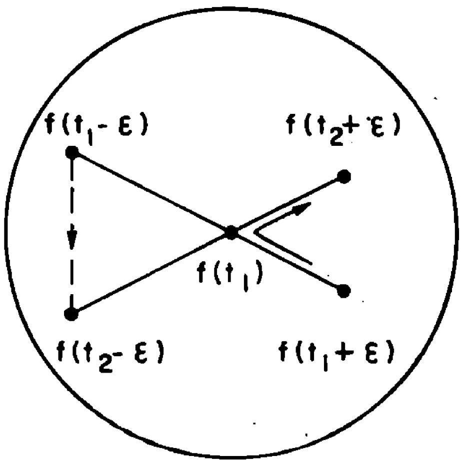


Fig. 2

$$f^*(t) = \begin{cases} f(t), & t \in [0, t_1 - \varepsilon], \\ \overline{f(t_1 - \varepsilon) f(t_2 - \varepsilon)}, & t \in [t_1 - \varepsilon, t_1], \\ \text{reverse of } f|\,[t_1, t_2 - \varepsilon], & t \in [t_1, t_2], \\ f(t), & t \in [t_2, 1]. \end{cases}$$

This path is shorter than f unless $f[t_1-\varepsilon, t_1] \cup [t_2-\varepsilon, t_2]$ is a line segment. Similarly $f[t_1, t_1+\varepsilon] \cup [t_2, t_2+\varepsilon]$ must be a line segment. Thus $f(t_1-\varepsilon) = f(t_2+\varepsilon)$. Now either $f(t_1-\varepsilon)$ is on the boundary of $\operatorname{con} fI$ or we can continue this process. In any case, we reach a point $f(s_1) = f(s_2)$ with $s_1 < t_1 < t_2 < s_2$ with $f(s_1)$ on the boundary of $\operatorname{con} fI$. Thus $f(s_1)$ is on the boundary of $\operatorname{con} f[s_1, s_2]$ and so $f| [s_1, s_2]$ must be the shortest closed curve meeting all the lines which meet $\operatorname{con} f[s_1, s_2]$. This contradicts the fact that f is not simple on $[s_1, s_2]$ unless $f| [s_1, s_2]$ is a closed line segment.

Suppose $f(t_1) \notin \partial \operatorname{con} fI$. Then the same proof shows that $f| [t_1, t_2]$ is either a simple convex closed path or a closed line segment.

LEMMA 6. *If $s \in (0, 1)$ there exists $t_1 < s < t_2$ such that either f is simple on $[t_1, t_2]$ or $f| [t_1, t_2]$ is a closed line segment one of whose is $f(s)$.*

Proof. If not, for every $\varepsilon > 0$, $f| [s-\varepsilon, s+\varepsilon]$ is not 1–1 nor is it a closed line segment. If $t_1 < s < t_2$ with $f(t_1) = f(t_2)$, then the theorem follows by Lemma 5. Thus for all $t_1 < s$ there does not exist $t_2 > s$ such that $f(t_2) = f(t_1)$. We may assume without loss of generality that, for every $\varepsilon > 0$, $f| [s-\varepsilon, s]$ is not 1–1 nor is it a line segment. Suppose for every n there exists $x_n \in (s-1/n, s)$ such that $f| (s-1/n, s)$ is not 1–1 at x_n. Then for all x_n there exists $y_n \in (s-\varepsilon_n, s)$ such that $f(x_n) = f(y_n)$ with $\lim \varepsilon_n = 0$ and (we may assume) $x_n < y_n < s$. (We may assume $f| [s-\varepsilon, s+\varepsilon]$ is 1–1 at s, for if $t > s$ with $f(t) = f(s)$, then by Lemma 5 either $f| [t, s]$ is a closed line segment or $f| [t, s]$ is a simple closed path and hence there cannot exist $t' \in (t, s)$ such that $f(t') = f(s)$.) By continuity and the fact that $f| [s-\varepsilon, s+\varepsilon]$ is 1–1, we may assume that $f[x_n, s]$ does not meet $f[x_{n-1}, y_{n-1}]$. Since the sequence $\{f(x_i)\}$ lies on the boundary of $\operatorname{con} fI$ by Lemma 5, so does $f(s)$. Note that for all i and $n \geq 2$, $\operatorname{con} f[x_i, y_{i+n}] \not\supseteq f[x_j, y_j]$ for all $i < j < i+n$, otherwise f could be shortened by the removal of $f| [x_j, y_j]$. Let L_i be a tangent to $\operatorname{con} fI$ at $f(x_i)$. If L is a tangent to $f[x_i, y_i]$ and $f[x_{i+n}, y_{i+n}]$ such that these two paths are contained in the figure bounded by L, L_i and L_{i+n}, it must be the case that L meets $f[x_j, y_j]$ at least twice for all $i < j < i+n$ (see Figure 3). Since f has finite length, $\lim_{i \to \infty} a(f| [x_i, y_i]) = 0$. We choose a subsequence

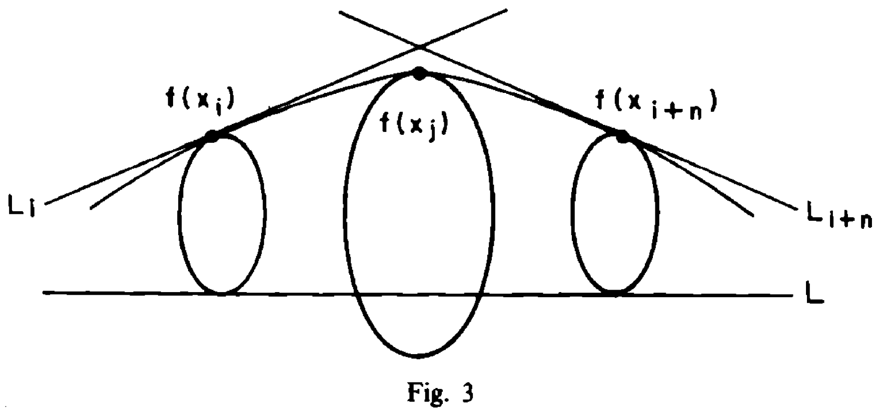


Fig. 3

$x_{i_1}, x_{i_2}, x_{i_3}$ of $\{x_i\}$ such that $a(f|\,[x_{i_{j+1}}, y_{i_{j+1}}])$ is less than the diameter of $f[x_{i_j}, y_{i_j}]$, $j = 1, 2$. Now without changing con fI we can delete $f[x_{i_2}, y_{i_2}]$ if we replace $f[x_{i_3}, y_{i_3}]$ by $f[x_{i_2}, y_{i_2}]$ (see Figure 4). This contradiction proves the lemma.

LEMMA 7. *The path f is at most two-to-one. In addition, excluding a finite number of intervals on which f is a closed line segment, f is not 1–1 at most finitely many points.*

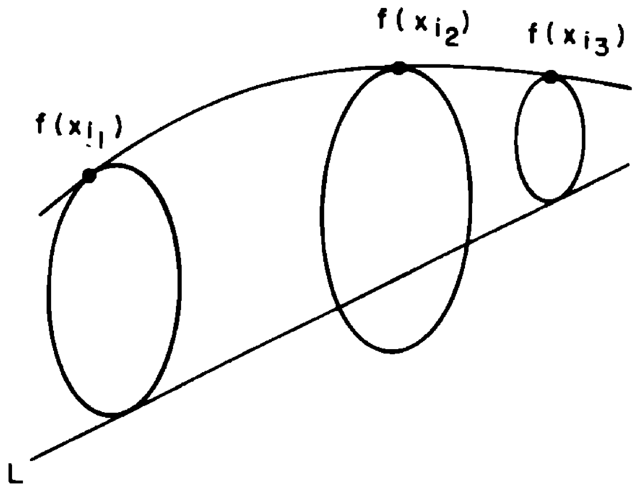


Fig. 4

Proof. If $s_1 < s_2 < s_3$ with $f(s_1) = f(s_2) = f(s_3)$, Lemma 5 shows that $f|\,[s_1, s_3]$ is a closed line segment. Thus f is at most two-to-one. Suppose there are infinitely many points, not elements of intervals on which f is a closed line segment, at which f is not 1–1. Then there exists an s and monotone sequences $\{s_i\}$ and $\{t_i\}$ of such points such that $s = \lim s_i$ and $f(s_i) = f(t_i)$. Of course, $t = \lim t_i$ exists and $f(s) = f(t)$. If $s = t$, Lemma 6 is violated. Repeated application of Lemma 5 shows that $f|\,[s, t]$ is a simple closed path and for all i, $[s_i, t_i] \not\subseteq [s, t]$ and $[s, t] \not\subseteq [s_i, t_i]$. Thus without loss of generality (by taking subsequences) we may assume that $s_i < s_{i+1} < s < t_i < t_{i+1} < t$ for all i. We now produce a path $\bar{f}$ with the same image as f as follows (see Figure 5):

$$\bar{f} = \begin{cases} f & \text{on } [0, s], \\ \text{reverse of } f & \text{on } [s, t], \\ f & \text{on } [t, 1]. \end{cases}$$

But now the t_i are transformed into t'_i with $\lim t'_i = s$ and $s < t'_{i+1} < t'_i$. Thus $\bar{f}$ is not 1–1 in a neighborhood of s, contradicting Lemma 6.

Suppose there are infinitely many maximal intervals on which f is a closed line segment. Then without loss of generality there exists a sequence of intervals $\{[s_i, t_i]\}$ such that f is a maximal closed line segment on $[s_i, t_i]$,

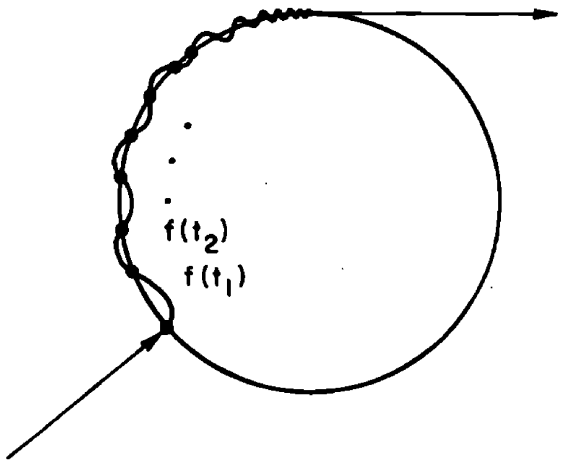


Fig. 5

$t_i < s_{i+1}$ for all i, and $s = \lim s_i$ and $t = \lim t_i$ exist. Clearly $s = t$ and thus Lemma 6 is violated.

LEMMA 8. *Let s_1 be the smallest such that there exists $s_2 \neq s_1$ with $f(s_1) = f(s_2)$. Then $s_1 \neq 0$ and $s_2 \neq 1$, and $t \notin [s_1, s_2]$ implies that $f(t) \notin f[s_1, s_2]$.*

Proof. Suppose $s_1 = 0$. One of the extreme tangents to $f[0, s_2]$ perpendicular to a tangent at $f(0)$ can be used to produce a shorter path (see Figure 6) unless they both coincide with the path $f|\ [0, s_2]$. In this case, $f|$ $[0, s_2]$ is a closed line segment and f can be shortened by eliminating an initial subpath. Hence $s_1 \neq 0$. Similarly, $s_2 \neq 1$.

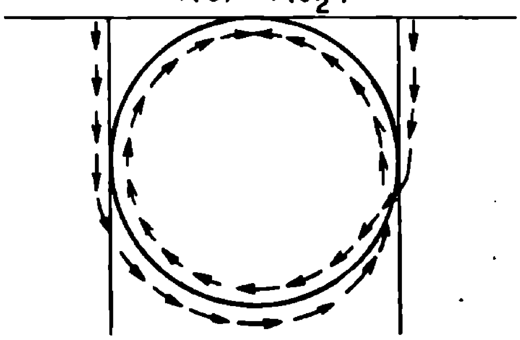


Fig. 6

Now by Lemma 7, there exists an $\varepsilon > 0$ such that for every $x \in [s_2, s_2+\varepsilon]$, $f(x) = f(x')$ implies that $x = x'$. If $f(s_2+\varepsilon) \in \operatorname{int} \operatorname{con} f[s_1, s_2]$, then $f|\ [s_2, s_2+\varepsilon]$ is a line segment. It is easy to shorten f as in Figure 7. Thus $f(s_2+\varepsilon) \notin \operatorname{int} \operatorname{con} f[s_1, s_2]$. Let P be a point on $f[s_1, s_2]$. Let A and B

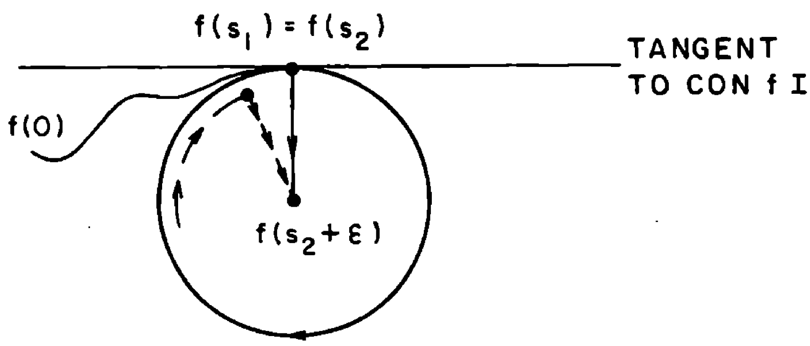


Fig. 7

be points on a tangent to $f[s_1, s_2]$ at $f(s_1)$. Since $f[0, s_1]$ lies outside $\operatorname{con} f[s_1, s_2]$, there exists $u < s_1$ such that $f[u, s_1]$ is contained wholely in either $\measuredangle Af(s_1)P$ or $\measuredangle Bf(s_1)P$. We assume the first. Then we must have $f(s_2+\varepsilon) \in \measuredangle Bf(s_1)P$ since all the other cases can be eliminated as in Figures 8 (a) and 8 (b).

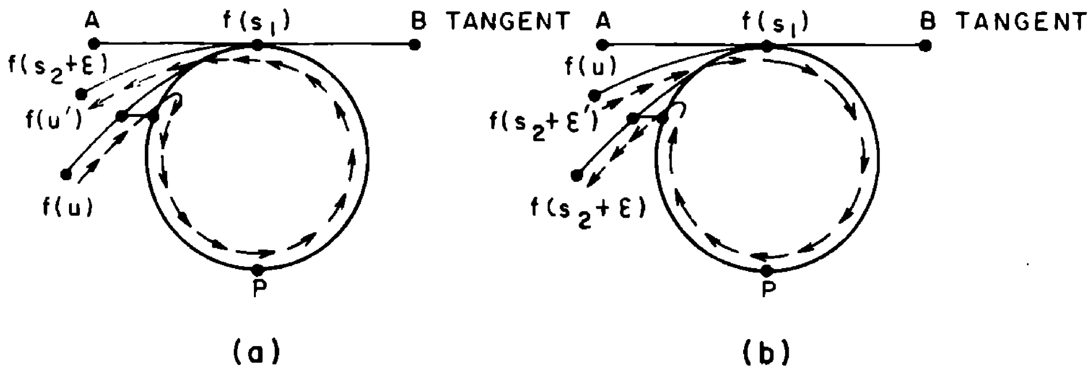


Fig. 8 (a) $\operatorname{Con} f[s_1, s_2+\varepsilon] \supseteq f[u', s_1]$; (b) $\operatorname{Con} f[u, s_2] \supseteq f[s_2, s_2+\varepsilon']$

Suppose there are t_1 and t_2 such that $s_1 < t_1 < s_2 < t_2$ and $f(t_1) = f(t_2)$. Then $f(t_1) \in \partial \operatorname{con} fI$ and the same proof as above shows that we have Figure 9 (a). Since $s_1 \neq 0$, then there is an $\varepsilon > 0$ such that $f[s_1, s_1+\varepsilon] \cup \cup f[s_2+\varepsilon, s_2] - \{f(s_1)\} \subseteq \operatorname{int} \operatorname{con} fI$, in which case the shorter path in Figure 9 (b) is used. The contradiction proves the lemma.

Proof of Theorem 7. Let s_1 and s_2 be as in Lemma 8. Let $J = [s_1, s_2]$ and let $g = f|J$. Let T be a tangent to $\operatorname{con} fI$ at $P = f(s_1)$. Let L be an extreme parallel to T with respect to $\operatorname{con} gJ$. Let $Q \in L \cap gJ$. Let A and B be points on T such that P is between them. We know that for some $\varepsilon > 0$, $f[s_1-\varepsilon, s_1)$ is on some side, say A, of $\overleftrightarrow{PQ}$ and $f(s_2, s_2+\varepsilon]$ is on the B side of $\overleftrightarrow{PQ}$. We now show that for all $t > s_2$, $f(t)$ is not on the A side of $\overleftrightarrow{PQ}$ and for all $t < s_1$, $f(t)$ is not on the B side of $\overleftrightarrow{PQ}$. Suppose to the contrary that $t > s_2$ such that $f(t)$ is on the A side of $\overleftrightarrow{PQ}$. Let $R = \overrightarrow{Pf(t)} \cap (gJ - \{P\})$ (see Figure 10). Since the part of g which goes

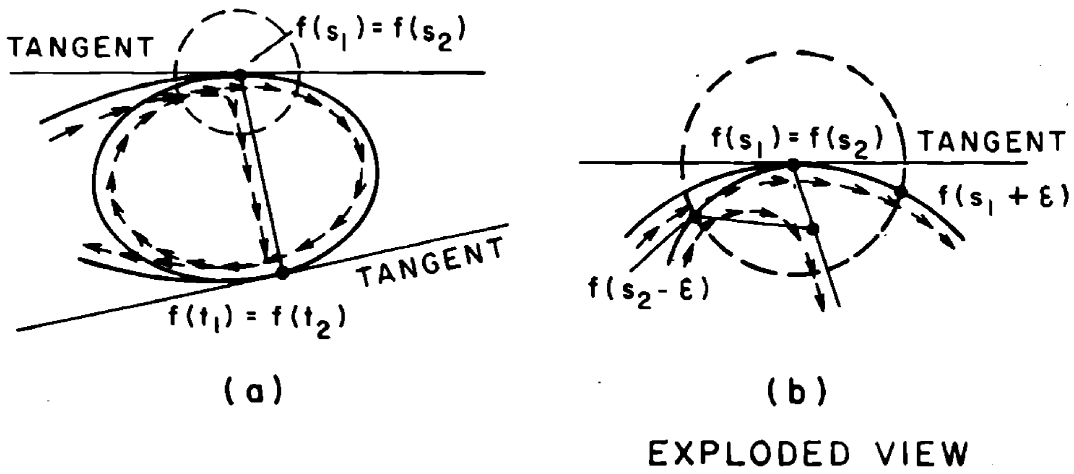


Fig. 9

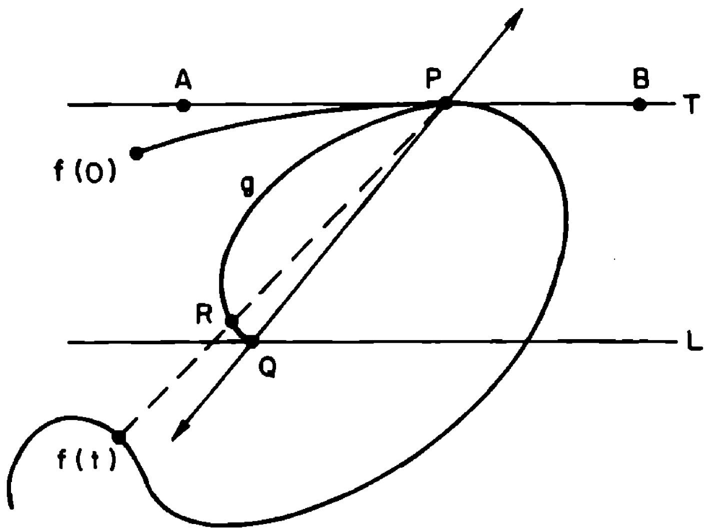


Fig. 10

from R to Q to P is interior to con fI, it must be a line segment, a contradiction.

Now consider parallels to $\overleftrightarrow{PQ}$ through $f(0)$ and $f(1)$. Let R and S be the points where these parallels meet T. We know we cannot have R and S on the same side of P. We eliminate g as follows: We add to $fI-gI$ certain pieces of gJ, $\overleftrightarrow{Rf}(0)$ and $\overleftrightarrow{Sf}(1)$ in order to insure that the new path meets L in passing from its beginning to s_1 and from s_2 to its end. Thus if $f(0)$ lies between L and T we add the pieces of the segment from $f(0)$ to L on $\overleftrightarrow{Rf}(0)$ which lie inside or on gJ and the parts of gJ which lie on the A side of PQ. Similarly, if $f(1)$ lies between L and T; otherwise we take just $fI-gI$ (see Figure 11). Clearly this new path f^* has the same convex hull as f, and f^* is shorter than f unless $f(0)$ and $f(1)$ are on T and g is a closed line segment. In this case (see Figure 12) there exists an ε such that $f[s_1-\varepsilon, s_1+\varepsilon] \subseteq \Delta f(0)f(1)Q$. Then substituting the line segment $\overline{f(s_1-\varepsilon)f(s_1+\varepsilon)}$ for $f[s_1-\varepsilon, s_1+\varepsilon]$ shortens f, a contradiction.

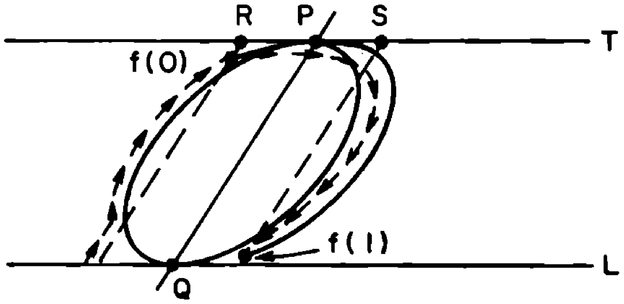


Fig. 11

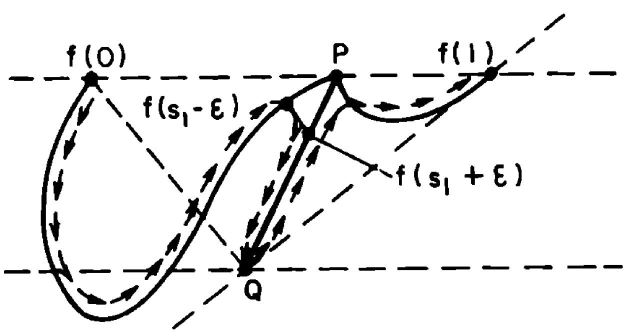


Fig. 12

THEOREM 8. *A shortest path which meets all the lines which meet the unit circle C has length $\pi+2$.*

Proof. Suppose f is a shortest path. Then $\operatorname{con} fI \supseteq C$. We know f is 1–1. Thus $fI \cap \partial \operatorname{con} fI$ consists of a collection of non-intersecting simple paths $\{C_\alpha\}$. Each simple path has a direction (inherited from f) and initial and terminal points (which may be equal when C_α is a single point). The remaining pieces of f, that is, $fI - \bigcup C_\alpha$, must be non-intersecting line segments. Suppose S is one of the segments in $fI - \bigcup C_\alpha$. If the ends of S are A and B, then the circle must lie between the tangents to $\operatorname{con} fI$ at A and B. If S does not cut the circle in two places, S must be contained in $\partial \operatorname{con} fI$, a contradiction. Thus S divides the circle into two parts and $\operatorname{con} fI$ into two parts. We consider a part of $\operatorname{con} fI$ which contains a smallest part of the circle. Let P be the endpoint of f in this part. We may assume without loss of generality that S has been chosen such that A is between P and B on fI and that the part of f which runs from P to A is one of the C_α, say C_0 (see Figure 13). Clearly $\overline{PB}$ must be tangent to the circle at some point R, otherwise C_0 can be shortened while retaining $C \subseteq \operatorname{con} fI$.

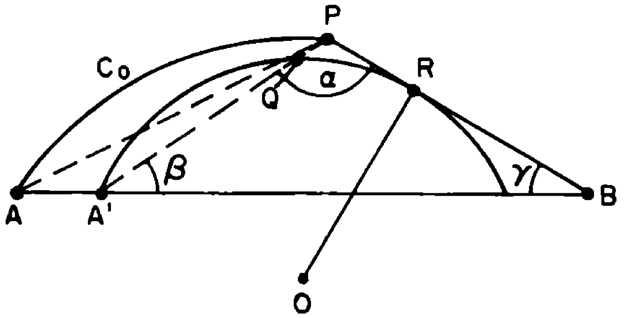


Fig. 13

Let O be the center of the circle, A' be the point on $\overline{AB}$ and C nearest to A, $Q = \overline{A'P} \cap C$, and $\alpha = \sphericalangle A'PB$, $\beta = \sphericalangle PA'B$ and $\gamma = \sphericalangle A'BP$. If $\overline{PB} \leqslant \overline{A'B}$, then we may replace $C_0 \cup \overline{AB}$ by the shorter path $C_0 \cup \overline{PB}$. Thus $\alpha < \beta$. Since $\gamma \leqslant \pi/2$, $\beta \geqslant \pi/4$. Now, by way of contradiction suppose $\sphericalangle QOR \leqslant \pi/2$. Fixing A' and moving Q toward A' along C increases β and decreases α, so the maximum β occurs when $\sphericalangle QOR = \pi/2$. Fix $\triangle A'PR$ so that $\sphericalangle QOR = \pi/2$ and increase β to the maximum by moving B along $\overleftrightarrow{PR}$. The maximum β is attained when O is on $\overline{A'B}$ since we are dealing with at most a semi-circle on the P side of $\overline{AB}$. In this case, $\beta = \alpha$ (see Figure 14) which is a contradiction. Thus $\sphericalangle QOR > \pi/2$.

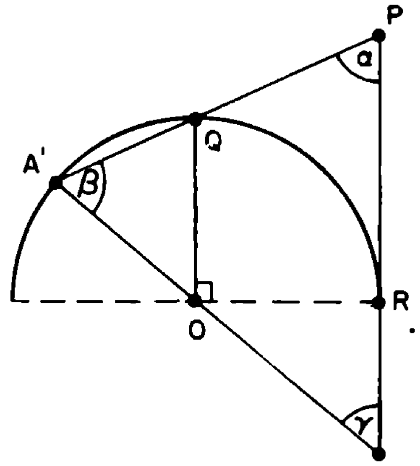


Fig. 14

Let $\overrightarrow{OX}$ be a ray perpendicular to $\overline{OR}$ meeting the circle at X and C_0 at T. Let $V = \overrightarrow{OA} \cap C_0$, let $Z = \overrightarrow{A'X} \cap \overline{PB}$, and let Y be the foot of the perpendicular from X to $\overline{PR}$ (see Figure 15). Now $\measuredangle ZA'B \leqslant \measuredangle XA'O = \measuredangle A'XO = \measuredangle A'ZB$, so $\overline{YB} \leqslant \overline{ZB} \leqslant \overline{A'B}$. We denote the shortest path from D to E along the path C by $C[D, E]$. It is easy to show from the corollary to Theorem 1 that $a(C_0[A, T]) \geqslant a(C[A', X])$. Since $a(C_0[T, P]) \geqslant \overline{XY}$, $a(C_0)+a(S) = \overline{AA'}+\overline{A'B}+a(C_0[A, T])+a(C_0[T, P]) \geqslant \overline{AA'}+\overline{YB}+ +a(C[A',X])+\overline{XY} = \overline{AA'}+2+\overline{RB}+a(C[A',X]) > \overline{AA'}+\pi/2+\overline{RB}+a(C[A',X]) = \overline{AA'}+\overline{RB}+a(C[A', R])$. It follows that f can be shortened by substituting $\overline{AA'} \cup C[A', R] \cup \overline{RB}$ for $C_0 \cup \overline{AB}$ while retaining $C \subseteq \operatorname{con} fI$. This contradiction shows that $fI - \bigcup C_\alpha = \emptyset$, and thus that f is convex. It is now easy to show that f must be a semi-circle with two radii attached (see Figure 16).

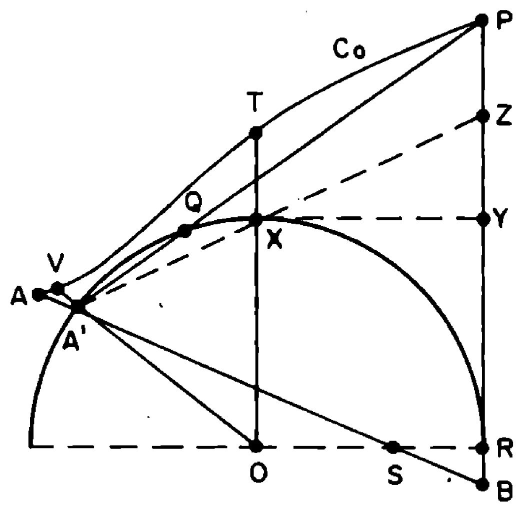


Fig. 15

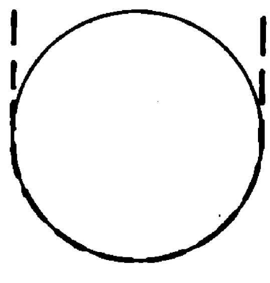


Fig. 16

CONJECTURES. 1° The shortest connected compact set meeting all the lines which meet a segment of a circle is a path;

2° The shortest connected compact set meeting all the lines which meet a triangle or square is the shortest connected compact set meeting all the vertices.

The problem of finding the shortest connected compact set meeting all the vertices of a convex polygon is surveyed in [9].

6. The disconnected case.

CONJECTURES. 3° (see Section 4). Let S be a compact set in $\mathbf{R}^2$. There exists a bounded set B containing S such that for each compact set S_0 such that $P(S, S_0)$ there exists a compact set $S_0^* \subseteq B$ such that $P(S, S_0^*)$ and $\lambda(S_0^*) \leqslant \lambda(S_0)$.

4° The shortest closed set S meeting all the lines which meet the unit circle C has two components and has length

$$\lambda(S) = 4.818924\ldots$$

Remark. Conjecture 4° is derived as follows. Four points P_1, P_2, P_3, P_4 are placed on the unit circle with center O. Let L_i be the tangent to C at P_i. Let $\sphericalangle P_1OP_4 = 2\alpha$, $\sphericalangle P_2OP_3 = 2\beta$ and $\sphericalangle P_1OP_2 = 2\gamma$. We suppose $2\gamma < \pi$ and that L_1 and L_2 meet at P, $L_1 \cap L_4 = \{Q\}$ and $L_2 \cap L_3 = \{R\}$ (see Figure 17), but note that P_1, O and P_3 do not necessarily lie on a straight line. We let $S(\alpha, \beta) =$ (part of C from P_3 to P_4 not passing through P_2) $\cup\, \overline{QP_4} \cup \overline{RP_3} \cup$ (altitude of $\triangle PQR$). Now

$$\lambda(S(\alpha, \beta, \gamma)) = 2\pi - (2\alpha - \tan\alpha) - (2\beta - \tan\beta) + \\ + \frac{(\tan\alpha + \tan\gamma)(\tan\beta + \tan\gamma)\sin 2\gamma}{[(\tan\alpha + \tan\gamma)^2 + (\tan\beta + \tan\gamma)^2 + 2(\tan\alpha + \tan\gamma)(\tan\beta + \tan\gamma)\cos 2\gamma]^{1/2}}.$$

We conjecture that the absolute minimum of $\lambda(S)$ is reached when $\alpha = \beta$ but we have been unable to verify this. Assuming that $\alpha = \beta$, then $\partial\lambda/\partial\gamma = 0 = \partial\lambda/\partial\alpha$ yields

$$(*) \qquad 2\cos 2\alpha = \sin\gamma, \quad \cos\gamma + \tan\alpha + \sin\gamma(\sec^2\gamma + 1) = 2,$$

whose simultaneous solution is $\alpha_0 \cong 36.855833°$ and $\gamma_0 \cong 34.121111°$. We have chosen $\lambda(\alpha_0, \alpha_0, \gamma_0)$ in Conjecture 4. Added evidence that this is indeed

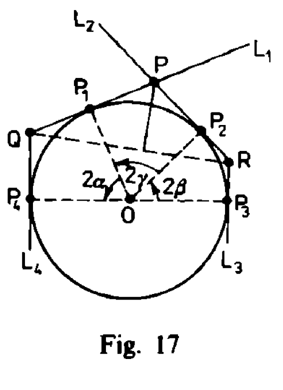


Fig. 17

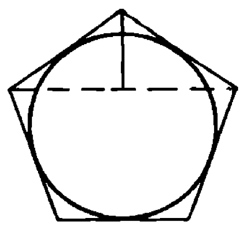
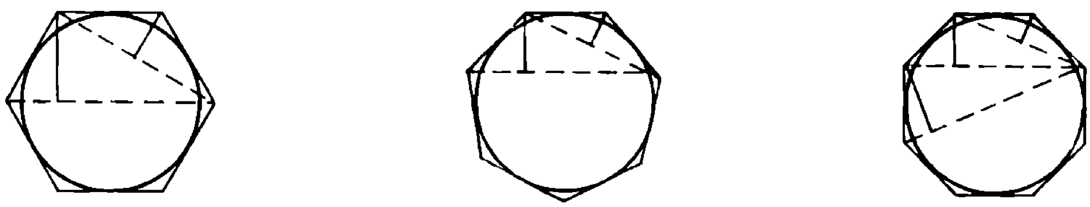


Fig. 18

the minimum λ is given by the fact that $d\lambda(\alpha_0, \beta, \gamma_0)/d\beta = 0$ when $\beta = \alpha_0$. We note, however, that if α_1 and γ_1 do not satisfy (*), then

$$d\lambda(\alpha_1, \beta, \gamma_1)/d\beta \neq 0 \quad \text{at } \beta = \alpha_1.$$

The method used to derive Conjecture 4 is an off-shoot of an earlier conjecture we had about the minimum $\lambda(S)$. At that time we thought (Platonic inspiration) that the best possible S was the closed set derived from the regular circumscribed pentagon as illustrated in Figure 18 ($\alpha = \beta = \gamma = 36°$) but it has length $\lambda = \frac{4}{5}\pi + 2\tan\frac{1}{5}\pi(1+\sin\frac{1}{5}\pi) \cong 4.82046115\ldots$ In fact, the minimum of $\lambda(\alpha, \alpha, \alpha)$ is $\lambda = 4.820427351\ldots$ at $\sim 35.8585677°$. Other short closed sets S derived from circumscribed regular polygons are illustrated in Figure 19. The generalization of our method, suggested by Figure 19, to n points on a circle ($n \geqslant 5$) looks extremely hard to handle and we have made no attempt.



(a) (b) (c)

Fig. 19 (a) $\frac{2}{3}\pi + \sqrt{3} + 1 \approx 4.82644591$;
(b) $\frac{6}{7}\pi + 2\tan\frac{1}{7}\pi(1 + \sin\frac{1}{7}\pi + \sin\frac{2}{7}\pi) \approx 4.82685813$;
(c) $\frac{3}{4}\pi + 2\tan\frac{1}{8}\pi(1 + \sin\frac{1}{8}\pi + \sin\frac{1}{4}\pi + \sin\frac{3}{8}\pi) \approx 4.85280026$.

Acknowledgement. The authors are indebted to Roy O. Davies who improved our earlier proof of Theorem 3 and to R. Laver whose example is given in Section 2.

References

[1] A. S. Besicovitch, *On the fundamental geometrical properties of linearly measurable plane sets of points* (II), Math. Ann. 115 (1938), 296–329.

[2] —, *On the fundamental geometrical properties of linearly measurable plane sets of points* (III), ibidem 116 (1939), 349–357.
[3] S. Gołąb, *Sur quelques points de la théorie de la longuer*, Ann. Soc. Polon. Math. 7 (1929), 227–241.
[4] L. D. Ivanov, *Variations of sets and functions*, "Nauka", Moscow 1975 (Russian).
[5] K. Kuratowski, *Topologie*, Vol. II, Warszawa 1958.
[6] L. Nachbin, *The Haar integral*, Van Nostrand, Princeton 1965.
[7] L. A. Santaló, *Integral geometry and geometric probability*, Encyclopedia of Mathematics and its Applications, Vol. I. Addison–Wesley Pub. Co., 1976.
[8] S. Sherman, *A comparison of linear measures in the plane*, Duke Math. J. 9 (1942), 1–9.
[9] R. Spira, *The isogonic and Erdös–Mordell points of a simplex*, Amer Math. Monthly 78 (1971), 856–864.
[10] A. G. Vituškin, *A proof of upper semicontinuity of a set variation*, Dokl. Akad. Nauk. SSSR, 166 (1966), 1022–1025 (Engl. transl. Soviet Math. Dokl. 7 (1966), 206–209).
[11] T. Ważewski, *Kontinua prostowalne w związku z funkcjami i odwzorowaniami absolutnie ciągłymi*, Ann. Soc. Polon. Math. 3, Suppl. (1927), 9–49 (in Polish).

UNIVERSITY OF COLORADO, DENVER, COLORADO
UNIVERSITY OF COLORADO, BOULDER, COLORADO

Reçu par la Rédaction le 19. 07. 1978