On the Length of Optimal TSP Circuits in Sets of Bounded Diameter

SHLOMO MORAN

Computer Science Department, Technion, Israel Institute of Technology, Haifa, Israel

Communicated by the Managing Editors

Received January 19, 1983

Let $V$ be a set of $n$ points in $\mathbb{R}^k$. Let $d(V)$ denote the diameter of $V$, and $l(V)$ denote the length of the shortest circuit which passes through all the points of $V$. (Such a circuit is an "optimal TSP circuit"). $l^k(n)$ are the extremal values of $l(V)$ defined by $l^k(n) = \max \{ l(V) \mid V \in \mathcal{V}_n^k \}$, where $\mathcal{V}_n^k = \{ V \mid V \subseteq \mathbb{R}^k, |V| = n, d(V) = 1 \}$.

A set $V \in \mathcal{V}_n^k$ is "longest" if $l(V) = l^k(n)$. In this paper, first some geometrical properties of longest sets in $\mathbb{R}^2$ are studied which are used to obtain $l^2(n)$ for small $n$'s, and then asymptotic bounds on $l^k(n)$ are derived. Let $\delta(V)$ denote the minimal distance between a pair of points in $V$, and let: $\delta^k(n) = \max \{ \delta(V) \mid V \in \mathcal{V}_n^k \}$. It is easily observed that $\delta^k(n) = O(n^{-1/k})$. Hence, $c_k = \lim_{n \to \infty} \delta^k(n) n^{1/k}$ exists. It is shown that for all $n$, $c_k n^{-1/k} \leq \delta^k(n)$, and hence, for all $n$, $l^k(n) \geq c_k n^{1-1/k}$. For $k = 2$, this implies that $l^2(n) \geq (\pi^2/12)^{1/2} n^{1/2}$, which generalizes an observation of Fejes-Toth that $\lim_{n \to \infty} l^2(n) n^{-1/2} = (\pi^2/12)^{1/2}$. It is also shown that $l^k(n) \leq [(3 - \sqrt{3})k/(k-1)] \delta^k(n) + o(n^{1-1/k}) \leq [(3 - \sqrt{3})k/(k-1)] n^{1-1/k} + o(n^{1-1/k})$.

The above upper bound is used to improve related results on longest sets in $k$-dimensional unit cubes obtained by Few (Mathematika 2 (1955), 141–144) for almost all $k$'s. For $k = 2$, Few's technique is used to show that $l^2(n) \leq (\pi n/2)^{1/2} + O(1)$.

1. INTRODUCTION

Let $\mathbb{R}$ denote the set of the real numbers. The Euclidean traveling salesman problem (TSP) in $\mathbb{R}^k$ is the following: Given $n$ points $x_1, ..., x_n$ in $\mathbb{R}^k$, find the shortest circuit (i.e., closed curve) which passess through them. Such a circuit is an "optimal TSP circuit." It is easily verified that an optimal TSP circuit is a polygonal line through $x_1, ..., x_n$. In some applications it is required that the distance between any 2 points in the given set is bounded by some constant $D$ (e.g., when the points represent nodes in a communication network. $D$ represents the maximal distance at which 2 nodes can communicate. An optimal TSP circuit in this case corresponds to a most efficient communication protocol [6].)
The problem addressed in this paper is the following: Given $n$, $k$, and $D$, what is the maximal length of an optimal TSP circuit through $n$ points in $R^k$, the distance between any pair of which is at most $D$. Denote this length by $l^k(n, D)$. It is not hard to verify that $l^k(n, D) = D \cdot l^k(n, 1)$. Hence, we can restrict ourselves to the case where $D = 1$. For brevity, we denote $l^k(n, 1)$ by $l^k(n)$. We shall be interested in both the values $l^k(n)$ and the properties of points sets of “maximal length” which realize these values.

It had been noted (see [5]) that this problem is closely related to the following problem of “optimal packing”: Allocate $n$ points in $R^k$ such that the distance between any pair of them is 1, and the minimal distance between any pair of them is maximized. Denote this “maximal minimal distance” by $\delta^k(n)$. It is easily observed that $l^k(n) \geq n\delta^k(n)$. Thue and others (see [4, pp. 160–166; 7]) had shown that $\delta^2(n)$ is asymptotically equal to $(\pi^2/12)^{1/4}n^{1/2}$. The exact value of $\delta^2(n)$ for $n \leq 7$ (and the geometrical properties of the corresponding configuration) are given in [1]. For $k > 2$, even the asymptotic values of $\delta^k(n)$ are not known (see, e.g., [2, pp. 405–411]).

The paper has 5 sections. The rest of this section includes the necessary definitions and notations. In Section 2 some geometrical properties of sets of maximal length in $R^2$ are proved. These properties are then used to compute $l^2(4)$ and to give some results concerning $l^2(5)$. In Section 3 we give lower bounds on $l^k(n)$, which generalize the observation mentioned above about the connection between $\delta^k(n)$ and $l^k(n)$. In that section we also give a result on $\delta^k(n)$ which seems to be of independent interest (Theorems 3.2 and 3.2'). In Section 4 upper bounds on $l^k(n)$ are given: first we give an upper bound for arbitrary $k$, which improves a result on longest sets in unit cubes obtained by Few in [5], and then we use the technique of Few to give a better bound on $l^2(n)$. In Section 5 two related results are discussed.

Notations and Definitions

Let $V = \{x_1, ..., x_n\}$ be a set of $n$ points in $R^k$ (for some $k$). A path in $V$ is a sequence $P = (x_{i_1} - x_{i_2} - ... - x_{i_m})$ of points of $V$. For $j = 1, ..., k - 1$, $(x_{i_j} - x_{i_{j+1}})$ is an arc of $P$. An arc $(x - y)$ will be identified with the straight line segment connecting $x$ and $y$. The length of a path $P$ is defined by

$$l(P) = l(x_{i_1} - ... - x_{i_m}) = \sum_{j=1}^{m-1} \delta(x_{i_j}, x_{i_{j+1}}),$$

where $\delta(x, y)$ is the Euclidean distance between $x$ and $y$.

A Hamiltonian circuit or a TSP circuit in $V$ is a path $H = (x_1 - x_{i_2} - ... - x_{i_n} - x_1)$ in which $i_j \neq i_k$ for $j \neq k$. We shall identify 2 Hamiltonian circuits if one is obtained from the other by reversing the order of the points. Thus, for $n \geq 3$, there are $(n - 1)!/2$ distinct Hamiltonian circuits on sets of $n$ points.
DEFINITION 1.1. Let \( V = \{x_1, \ldots, x_n\} \). Then the length of \( V \), \( l(V) \) is defined by
\[
l(V) = \min\{l(H) \mid H \text{ is a Hamiltonian circuit in } V\}.
\]
For a set \( V \) and a point \( x \), \( d(x, V) = \max\{\delta(x, y) \mid y \in V\} \). The diameter of \( V \), \( d(V) \), is defined by
\[
d(V) = \max\{d(x, V) \mid x \in V\}\]
For positive integers \( n \) and \( k \), let
\[
V^k_n = \{V \mid V \subseteq \mathbb{R}^k, d(V) = 1, |V| = n\}.
\]
DEFINITION 1.2. For a positive integer \( n \)
\[
l^k(n) = \max\{l(V) \mid V \in V^k_n\}\]
A set \( V^*_n \) is a “longest set” if
(i) \( V^*_n \in V^k_n \);
(ii) \( l(V^*_n) = l^k(n) \).
DEFINITION 1.3. Let \( V \subseteq \mathbb{R}^k \) for some \( k \). Then
\[
\delta(V) = \min\{\delta(x, y) \mid x \neq y, x, y \in V\}.
\]
For positive integers \( n \) and \( k \)
\[
\delta^k(n) = \max\{\delta(V) \mid V \in V^k_n\}.
\]
The numbers \( \delta^k(n) \) are sometimes called “packing constants” [3].
Most of the proofs in the paper are given for the case \( k = 2 \), and it will be clear from the text when they generalize to arbitrary \( k \). \( V^2_n \) will be denoted by \( V_n \), and \( V_n \) will denote a set in \( \mathbb{V}_n \). Similarly, \( l(n) \) and \( \delta(n) \) will denote \( l^2(n) \) and \( \delta^2(n) \), respectively.

2. SOME PROPERTIES OF LONGEST SETS IN \( \mathbb{R}^2 \)

In this section we prove lemmas which provide some insight on the structure of planar longest sets. We then use these lemmas to find a longest

---

1 The diameter of \( V \) is sometimes denoted as “the maximal chord length of \( V \).”
2 The use of the term “\( \max \)” (and not “\( \sup \)” in the definition of \( l^k(n) \)) is justified by the fact that \( V^k_n \) is homeomorphic to a compact subset of \( \mathbb{R}^{kn} \) and that \( l(V) \) is a continuous function. Similar remarks apply to a few other definitions in the paper.
set \( V_4^* \) of cardinality 4, (and thus to compute \( l(4) \)), and to give some results concerning \( V_5^* \) and \( l(5) \). Note that trivially \( l(1) = 0, l(2) = 2, \) and \( l(3) = 3 \). \( V_n^* \) for \( n = 1, 2, 3 \) are given in Fig. 2.1.

**Lemma 2.1.** Let \( V_n = \{x_1, \ldots, x_n\} \) be a set of \( n \) points in the plane \( (n \geq 4) \), not all of them on the same line. Then an optimal TSP circuit in \( V_n \) is a simple curve (that is: a curve which does not intersect itself).

**Proof.** Let \( H = (x_1 - x_i, \ldots, x_i - x_1) \) be a TSP circuit in \( V_n \). We shall show that if \( H \) intersects itself, then \( H \) is not optimal. For simplicity, assume that \((i_2, \ldots, i_n) = (2, \ldots, n)\).

Suppose that for some \( i \) and \( j \) \((|i - j| > 1)\), the arcs \((x_i - x_{i+1})\) and \((x_j - x_{j+1})\) intersect (see Fig. 2.2).

Assume first that \( x_i, x_{i+1}, x_j, x_{j+1} \) are not collinear. Then by replacing \((x_i - x_{i+1})\) and \((x_j - x_{j+1})\) by \((x_i - x_j)\) and \((x_{i+1} - x_{j+1})\) we obtain a TSP circuit which is shorter than \( H \) (due to the triangle inequality). The proof for the case where \( x_i, x_{i+1}, x_j, x_{j+1} \) are collinear is also not hard and is omitted. \( \square \)

**Definition 2.1.** Let \( V \) be a set of points. Then \( \text{CON}(V) \) is the boundary of the convex hull of \( V \) (i.e., the boundary of the smallest convex figure which contains \( V \)).
FIGURE 2.3

**Lemma 2.2.** Let \( V^*_n \subset V \). If \( V^*_n \) is longest, then for each \( x \) in \( V^*_n \),
\[
d(x, V^*_n) = 1 \text{ iff } x \in \text{CON}(V^*_n).
\]

*Proof.* Since \( d(V^*_n) = 1 \), if \( x \in V^*_n \) and \( d(x, V^*_n) = 1 \), then \( x \) must in \( \text{CON}(V^*_n) \). Hence, it suffices to show that for every \( x \) in \( \text{CON}(V^*_n) \),
\[
d(x, V^*_n) = 1.
\]
For contradiction, assume that for some \( x \in \text{CON}(V^*_n) \),
\[
d(x, V^*_n) < 1.
\]
We shall show that there is a \( V' \subsetneq V \) such that \( l(V') > l(V^*_n) \). Let \( L \) be a supporting line of \( V^*_n \) through \( x \) (i.e., a line tangent to \( \text{CON}(V^*_n) \) at \( x \), see Fig. 2.3). \( V' \) is obtained by removing \( x \) a small distance \( h \) in a direction perpendicular to \( L \), as shown in Fig. 2.3. By doing this, \( d(x, y) \) is increased for all \( y \in V^*_n \), and hence \( l(V') > l(V^*_n) \). On the other hand, if \( h \) is small enough, \( d(x, V'_n) \) is less than 1 (since \( d(x, V^*_n) < 1 \)) and hence \( d(V'_n) = 1 \). Thus, \( V' \) is in \( V_n \). This completes the proof of the lemma. \( \blacksquare \)

**Definition 2.2** Let \( x, y \in V \). Then the arc \((x - y)\) is essential if it participates in every optimal TSP circuit in \( V \). \((x - y)\) is redundant if it participates in no optimal TSP circuit in \( V \). \((x - y)\) is nonessential (nonredundant) if it is not essential (redundant). A point \( x \) of \( V \) is internal in \( V \) if it is not in \( \text{CON}(V) \).

**Lemma 2.3.** Let \( V^*_n \) be a longest set, and let \( x \) be an internal point in \( V^*_n \). Then

(a) For all \( y \in V^*_n \), \((x - y)\) is nonessential.

(b) Let \( L \) be any line through \( x \). Then there are \( y, z \in V^*_n \) such that \( L \) separates \( y \) and \( z \) and both \((x - y)\) and \((x - z)\) are nonredundant.

*Proof.* (a) Assume that for some \( y \in V^*_n \), \((x - y)\) is essential. We derive a contradiction by showing that \( V^*_n \) is not longest. Suppose first that there is no \( z \in V^*_n \) such that \( x \) lies on the arc \((y - z)\) as in Fig. 2.4. Then, by

FIGURE 2.4
removing $x$ a distance $h$ away from $y$ along the line containing $(x - y)$, $\delta(x, y)$ increases by $h$, while for any $u \in V_n^*$, if $\delta(x, u)$ decreases, it decreases by less than $h$, and if $x \notin \{u, v\}$, $\delta(u, v)$ remains unchanged. It follows that the length of any TSP circuit which contains $(x - y)$ (and hence of any optimal TSP circuit) increases by some positive value. By making $h$ small enough, $d(x, V_n^*)$ remains smaller than 1 and the lengths of the nonoptimal TSP circuits remain larger than the length of the previously optimal circuits and hence $l(V_n^*)$ is increased, in contradiction with the assumption that $V_n^*$ is longest.

The argument above does not work if there is a point $z$ such that $x$ lies in $(y - z)$ as in Fig. 2.4, because then removing $x$ as before does not increase the length of the TSP circuits which use the path $(y - x - z)$. In this case, $x$ can be removed away from $y$ in a direction which forms a small but positive angle $\alpha$ with $(x - z)$, and a similar argument does apply.

(b) For contradiction, assume that there is a line $L$ through $x$ as in Fig. 2.5, such that for all nodes $y$ on the left side of $L$, $(x - y)$ is redundant. Then, by removing $x$ in a direction perpendicular to $L$ as shown in Fig. 2.5, $\delta(x, z)$ increases for all $z$ which are not on the left side of $L$, and hence, for all $z$ such that $(x - z)$ is nonredundant. Hence, similarly to the proof of (a), one can increase $\delta(V_n^*)$ by removing $x$ a small distance $h$ in that direction.

**Corollary.** Let $V_n^*$ be a longest set, and let $x$ be an internal point of $V_n^*$. Then there are at least 3 points $y_1, y_2, \text{ and } y_3$ in $V_n^*$ such that $(x - y_i)$ is nonredundant $(i = 1, 2, 3)$.

**Proof.** Since every TSP circuit must pass through $x$, there are $y_1, y_2$ in $V_n^*$ such that the path $(y_1 - x - y_2)$ is in an optimal TSP circuit. Hence, $(y_1 - x)$ and $(y_2 - x)$ are nonredundant. If there is no $y_3$ such that $(y_3 - x)$ is nonredundant, then both $(y_1 - x)$ and $(y_2 - x)$ occur in every optimal circuit, which means that $(y_1 - x)$ and $(y_2 - x)$ are essential, in contradiction to Lemma 2.3(a).

**Definition 2.3.** Let $C$ be a closed curve in the plane and let $D$ be a real number. Then $C$ is a curve of constant width $D$ if for each $x \in C$, $d(x, C) = D$. 
A "figure of constant width" is a convex figure whose boundary is a curve of constant width. Examples of curves of constant width $D$ are a circle of diameter $D$ and a Reuleaux triangle of side length $D$ (see Fig. 2.6). For more about curves of constant width see [9]. We shall use the following lemma concerning these curves.

**Lemma 2.4.** (a) Every curve of constant width $D$ has a perimeter $\pi D$.

(b) The area of a figure of constant width $D$ is at most $\pi D^2/4$.

**Proof.** Part (a) is Barbier's theorem; (b) follows from (a) by the isoperimetric theorem [9, pp. 51–58].

Since every set of points of diameter 1 can be embedded in a figure of constant width 1, Lemma 2.4 implies

**Lemma 2.5.** Let $V_n$ be in $V_n$. Then

(a) The perimeter of $\text{CON}(V_n)$ is less than $\pi$.

(b) There is a convex figure whose area is less than $\pi/4$ which contains $V_n$.

**Definition 2.4.** Let $k, n$ be given, $2 < k \leq n$. Then

$$V_{n,k} = \{V_n \mid V_n \in V_n, |V_n \cap \text{CON}(V_n)| = k\};$$

$$l(n, k) = \sup \{l(V_n) \mid V_n \in V_{n,k}\}.$$

**Lemma 2.6.** $l(4) = 2(1 + \sqrt{3}/3) = 3.1547... .$

**Proof.** Let $V^*_4$ be a longest set of 4 points. Then $V^*_4$ is either in $V_{4,3}$ or in $V_{4,4}$. Hence, $l(4) = \max\{l(4, 3), l(4, 4)\}$. We shall prove first that $l(4, 4) \leq \pi$, and then that $l(4, 3) = 2(1 + \sqrt{3}/3) > \pi$.

Let $V_4$ be in $V_{4,4}$. Then $l(V_4)$ is the perimeter of $V_4$, which by Lemma 2.5 is less than $\pi$. Hence, $l(4, 4) \leq \pi$.

Let $V'_4 = \{x, y, z, u\}$, where $x, y, \text{ and } z$ are the vertices of an equilateral triangle of side length 1, and $u$ is the center of this triangle (see Fig. 2.7).

\[\text{In fact, one can show that } l(4, 4) = 2 + 1/\cos 15^\circ = 3.0353^{(c)}... .\]
Then, as one can easily verify
\[ l(V_4) = l(u - y - x - z - u) = 2(1 + \sqrt{3}/3) > \pi. \]
Hence, \( l(4) \geq l(V_4') > l(4, 4) \). Therefore, a longest set in \( V_4 \) must be in \( V_{4,3} \).

It remains to show that \( V_4' \) is, in fact, a longest set: Let \( V_4'^* = \{x', y', z', u'\} \) be a longest set, and let \( u' \) be the internal point of \( V_4'^* \). Using the same technique that was used in the proof of Lemma 2.2, one can show that \( \{x', y', z'\} \) are the vertices of an equilateral triangle of side length 1. Also, by Lemma 2.3 and its corollary, each of the arcs \((x' - y'), (x' - z'), (x' - u')\) is nonredundant. Hence, all of the possible 3 TSP circuits have the same length, which implies that \( \delta(u', y') = \delta(u', z') = \delta(u', x') \) and hence that \( u' \) is the center of the triangle. The lemma follows.

An interesting corollary to the last two lemmas is the following: Let \( n \geq 4 \), and let \( V_n^* \) be a longest set in \( V_n \). Then \( V_n^* \) is not in \( V_{n,n} \).

Deriving \( l(n) \) and \( V_n^* \) for \( n \geq 5 \) seems to be hard. Using Lemmas 2.1–2.5, we have been able to prove some results concerning \( l(5) \) and \( V_5^* \). These results are stated below.

**Lemma 2.7.** \( l(5, 3) \approx 3.2175(-) \ldots \). Moreover, \( l(5, 3) = l(V_{5,3}^*) \), where \( V_{5,3}^* \) is defined by (see Fig. 2.8):

1. \( \{x, y, z\} \) are the vertices of an equilateral triangle of side length 1.
2. \( u \) is the center of the triangle.
3. \( v \) lies on the height from \( x \) to \( (y - z) \) and \( \delta(v, z) - \delta(v, u) = 1 - \sqrt{3}/3 = 0.4226^{(+)} \ldots \). 

The proof that the set \( V_{5,3}^* \) defined in Lemma 2.7 above is indeed longest
Let \( a = \xi vzy \).

Then one can check that, by (3):

\[
\delta(v, z) - \delta(v, u) = \frac{1}{2 \cos a} - \left[ \frac{\sqrt{3}}{6} - \frac{\tan a}{2} \right] = \frac{(1/2 \cos a + \tan a - \pi/3)/2}{1 - \sqrt{3}/3}.
\]

This implies that a longest set in this case is obtained when \( a \approx 19.79^\circ \). The value for \( l(V_{5,3}) \) follows by computing the length of one of the optimal routes— \((x - u - v - y - z - x)\), say:

\[
l(x - u - v - y - z - x) \approx \frac{\sqrt{3}}{2} - \frac{\tan 19.79^\circ}{2} + \frac{1}{2 \cos 10.79^\circ} + 2 \\
\approx 3.2175.
\]

**Lemma 2.8.** Let \( V^*_{s} \) be a longest set in \( V_s \). Then \( V^*_{s} \in V_{5,4} \).

**Proof.** In view of Lemma 2.7, we have the show that there is a \( V^*_{s} \in V_{5,4} \) such that \( l(V^*_{s}) > l(5, 3) \approx 3.2175 \). Such a \( V^*_{s} \) is given in Fig. 2.9: \( x, y, \) and \( z \) form an equilateral triangle of side length 1, \( u \) and \( v \) lie on the bisector of \( \xi zxy \), \( \delta(x, u) = 1 \), and

\[
\delta(u, v) = \delta(y, u)/2 = 1/(4 \cos 15^\circ) = 1/(2(3^{1/2} + 2)^{1/2}) = (2 - \sqrt{3})^{1/2}/2.
\]

In \( V' \), all the 4 circuits which do not intersect themselves, are optimal. (Note that, by Lemma 2.1, these are the only candidates for optimal circuits.)

\[
l(V^*_{s}) = \delta(x, y) + \delta(y, u) + \delta(u, v) + \delta(v, z) + \delta(z, x) \\
= 1 + 1/(2 \cos 15^\circ) + 1/(4 \cos 15^\circ) + d(v, z) + 1 \\
\approx 2 + 3/(4 \cos 15^\circ) + 0.5153505 \\
\approx 3.291807 > l(5, 3).
\]

We conjecture that \( V^*_{s} \) is a longest set, though we do not yet have a formal proof for this.
In this section we derive a lower bound on $l^k(n)$. The results are stated and proved for $k = 2$, but are easily generalized to arbitrary $k$.

Let $V_n \subseteq V_n$ be such that for all $x, y$ in $V_n$, $\delta(x, y) \geq r$. Then clearly, $l(V_n) \geq nr$. Taking $r$ to be $\delta(n)$, we have that $l(n) \geq n\delta(n)$. Let

$$c_2 = \limsup_{n \to \infty} \delta(n) \sqrt{n}.$$

Then, by the discussion above

$$c_2 \leq \limsup_{n \to \infty} l(n)/\sqrt{n}.$$

We shall generalize this observation to the following stronger result:

**Theorem 3.1.** For all $n$, $l(n) \geq c_2 \sqrt{n}$.

Theorem 3.1 follows easily from

**Theorem 3.2.** For all $n$, $\delta(n)\sqrt{n} \geq c_2$.

Theorem 3.2 seems to be of independent interest, since it implies not only that $c_2 = \lim_{n \to \infty} \delta(n) \sqrt{n}$, but also that $c_2$ is a lower bound of $\delta(n) \sqrt{n}$. The $k$-dimensional version of Theorem 3.2 is

**Theorem 3.2'.** Let $c_k = \limsup_{n \to \infty} \delta^k(n) n^{1/k}$. Then for all $n$, $\delta^k(n) n^{1/k} \geq c_k$.

The key lemma for the above theorems is Lemma 3.2, which uses a relation between the "packing constants" and densities of "sparse sets," as described below:

Let $R$ be a planar figure of area $A$, and let $S$ be a finite set of points contained in $R$. Then the density of $S$ in $R$ is $|S|/A$. The set $S$ is "sparse" if for each pair of points $x, y$ in $S$, $\delta(x, y) \geq 1$. The packing constants $\delta(n)$ correspond to the density of sparse sets of points in certain planar figures in the following way: Suppose that for some $n$ and $c$, $\delta(n) \geq c/\sqrt{n}$. Then, by using appropriate scaling, one can obtain a planar figure $R$ of diameter $\sqrt{n}/c$ which contains a sparse set of $n$ points. Using the fact that, by Barbier's theorem, the area of $R$ cannot exceed $\pi n/(4c^2)$, we have that the density of $S$ in $R$ is at least $4c^2/\pi$. On the other hand, if we can embed a sparse set $S$ in a circle $C$ of diameter $\sqrt{n}/c$ such that the density of $S$ in $C \geq 4c^2/\pi$, then, since the area of $C$ is $\pi n/(4c^2)$, $S$ contains at least $n$ points. This implies (again by scaling), that $\delta(n) \geq c/\sqrt{n}$. The next lemma summarizes the above.
Lemma 3.1. (a) If $\delta(n) \geq c/\sqrt{n}$, then there is a sparse set $S$ of $n$ points contained in a figure $R$ of diameter $\leq \sqrt{n}/c$, and (hence) the density of $S$ in $R$ is at least $4c^2/\pi$.

(b) If a sparse set $S$ is contained in a circle $C$ of diameter $\sqrt{n}/c$ such that the density of $S$ in $C$ is at least $4c^2/\pi$, then $\delta(n) \geq c/\sqrt{n}$.

Due to the fact that for all $k$, the $k$-dimensional set of given diameter and maximal volume is a $k$-dimensional sphere, Lemma 3.1 has a simple generalization to the $k$-dimensional case. In this generalization, $\sqrt{n}/c$ is replaced by $n^{1/k}/c$, and $4c^2/\pi$ is replaced by $c^k/w_k$, where $w_k$ is the volume of the $k$-dimensional sphere of diameter 1.

In view of Lemma 3.1(b) above, Theorem 3.2 will follow if we show that for each $r$, there is a circle $C$ of radius $r$ which contains a sparse set $S$, such that the density of $S$ in $C$ is at least $4c^2/\pi$. (Recall that $c_2 = \lim \sup_{n \to \infty} \delta(b)\sqrt{n}$.) This will follow from Lemma 3.2, for which we need the following definitions:

Definition 3.1. For each positive real number $t$, let $R_t$ be a set of diameter $t$ and area $A_t$, and let $S_t$ be a finite set of points contained in $R_t$. For a given $r > 0$, $R_{t,r}$ is the set obtained by deleting from $R_t$ all the points whose distance from the boundary of $R_t$ is less than $r$. Let $A_{t,r}$ be the area of $R_{t,r}$, and let $S_{t,r} = S_t \cap R_{t,r}$. We say that the family $\{(R_t, S_t)\}$ is balanced if for each fixed $r$, the following hold:

$$\lim_{t \to \infty} \frac{A_{t,r}}{A_t} = 1;$$

(3.1.1)

$$\lim_{t \to \infty} \frac{|S_{t,r}|}{|S_t|} = 1.$$  

(3.1.2)

Definition 3.2. A family $\{(R_t, S_t)\}$ as above has density $e$ if it is balanced and

$$\lim_{t \to \infty} \frac{|S_t|}{A_t} = e.$$  

(3.2)

Example 1. Let $m$ be a positive integer, and let $G_m$ be the lattice

$$G_m = \left\{ \left( \frac{p}{m}, \frac{q}{m} \right) \mid p, q \text{ are integers} \right\}.$$  

For each $t$, let $R_t$ be a set of constant width $t$, and let $G_{m,t} = G_m \cap R_t$. Then $\{R_t, G_{m,t}\}$ has density $m^2$.

* The author is indebted to M. Perles for bringing this fact to his attention.
EXAMPLE 2. For each $t$, let $R_t$ be as in Example 1, and let $S_t$ be a sparse set of maximum possible cardinality contained in $R_t$. Then $\{(R_t, S_t)\}$ has a density $\geq 4c_t^2/\pi$ (this follows from Lemma 3.1(a) and the definition of $c_t$).

DEFINITION 3.3. Let $G_m$ be as in Example 1, and let $B$ be a bounded region of the plane. Then

$$N(B, m) = |G_m \cap B|.$$ 

The proofs of the following propositions are easy and omitted.

PROPOSITION 3.1. Let $B$ be a convex figure of area $F > 0$. Then

$$\lim_{m \to \infty} \frac{N(B, m)}{m^2} = F,$$

and the convergence is uniform (i.e., it does not depend on the specific location of $B$ in the plane).

PROPOSITION 3.2. Let $\{(R_t, S_t)\}$ be a family of density $e$ and let $r$ be a fixed number. Then

$$\limsup_{t \to \infty} \frac{|S_{t,r}|}{A_t} = e.$$ 

LEMMA 3.2. Let $\{(R_t, S_t)\}$ be a family of density $e$. Then for each $F > 0$, there is a circle $C$ of area $F$ and a number $t$ such that $|C \cap S_t| \geq |eF|^5$.

Proof. For simplicity, let $e = 1$. Assume for contradiction that for all $C$ of area $F$, and for all $t$, $|C \cap S_t| < F$. Let $F = I + h$, where $I$ is an integer and $0 < h \leq 1$, and let $\delta = h/F > 0$. Then for each $t$, each circle of area $F$ contains at most $I = F(1 - \delta)$ points of $S_t$. Let $\epsilon > 0$ be such that $(1 - \epsilon)^2 > (1 - \delta)(1 + \epsilon)$, and let $r = (F/\pi)^{1/2}$ (i.e., $\pi r^2 = F$). By Propositions 3.1 and 3.2 there exist $t$ and $m$ ($m$ depends on $t$) such that,

(i) $|N(C, m)/m^2 - F| < \epsilon F$ for every circle $C$ of area $F$;

(ii) $|S_{t,r}/A_t - 1| < \epsilon$;

(iii) $|N(R_t, m)/m^2 - A_t| < \epsilon A_t$.

Let $G_{m,t} = G_m \cap R_t$ and let $C_{m,t}$ be the set of all circles of radius $r$ (and area $F$) whose centers belong to $G_{m,t}$. For each $x \in S_t$ and $C \in C_{m,t}$ let

$$n_x = |\{C \mid C \in C_{m,t}, x \in C\}|,$$

$$n_c = |\{x \mid x \in S_t \cap C\}|.$$

$[X]$ denotes the smallest integer not smaller than $X$. 
Note that if \( x \) is in \( R_{t,r} \), then

\[ n_x = |\{u \mid u \in G_m, \delta(u, x) \leq r\}|, \]

and that, under the assumption that the lemma is false, \( n_c \leq F(1 - \delta) \) for all \( C \in \mathbb{C}_{m,t} \). Finally, let

\[ P = \{(x, C) \mid C \in \mathbb{C}_{m,t}, x \in C \cap S_t\} \quad \text{and let} \quad p = |P|. \]

We derive a contradiction by computing \( p \) by two different methods:

**Method 1:**

\[ p = \sum_{C \in \mathbb{C}_{m,t}} n_c \leq F(1 - \delta) |\mathbb{C}_{m,t}| < F(1 - \delta) m^2 (1 + \varepsilon) A_t. \]

The last inequality follows from (iii), since \( |\mathbb{C}_{m,t}| = N(R_t, m) \).

**Method 2:**

\[ p = \sum_{x \in S_t} n_x \geq \sum_{x \in S_t} n_x \geq |S_{t,r}| m^2 F(1 - \varepsilon) > A_t (1 - \varepsilon) m^2 F(1 - \varepsilon). \]

The second inequality follows from (i), and the last inequality from (ii).

By combining the above result and cancelling equal terms, we get 
\( (1 - \varepsilon)^2 < (1 - \delta)(1 + \varepsilon) \), which contradicts the assumption on \( \varepsilon \).

**Proof of Theorem 3.2.** By Lemma 3.1(b), it is enough to show that for every circle \( C \) there is a sparse set \( S \) whose density in \( C \geq 4c_2/\pi \). For each positive real number \( t \), let \( S_t \) be a sparse set of width \( t \) and of maximum possible cardinality, and let \( R_t \) be a set of constant width containing \( S_t \). Then, by Lemma 3.1(a) and the definitions, the family \( \{(R_t, S_t)\} \) has density \( \geq 4c_2/\pi \) (see Example 2). Let \( C \) be a given circle of area \( F \). Then by Lemma 3.2, there is a replica \( C' \) of \( C \) and a number \( t \) such that \( |C' \cap S_t| \geq |F \cdot 4c_2^2/\pi| \geq F \cdot 4c_2^2/\pi \). This implies that the density of \( C' \cap S_t \) in \( C' \) is at least \( 4c_2^2/\pi \). Since \( S_t \) is a sparse set, so is \( C' \cap S_t \). The theorem follows.

**Corollary.** For each \( n \),

\[ \delta(n) \geq (\pi^2/12)^{1/4} m^{-1/2} \quad \text{and} \quad l(n) \geq (\pi^2/12)^{1/4} n^{1/2}. \]

**Proof.** By Theorems 3.1 and 3.2, using the result (due to Thue and others, see [2]) that \( \lim_{n \to \infty} \delta(n) \sqrt{n} = (\pi^2/12)^{1/4} \).
4. An Upper Bound on $l^k(n)$

In this section first we derive an upper bound on $l^k(n)$ expressed in terms of $\delta^k(n)$, and then derive from it an upper bound expressed in terms of $k$ and $n$ only. We show that our result improves a result of Few on longest sets in unit $k$-dimensional cubes [5] for almost all $k$'s, and then use the technique of Few to improve our result for $k = 2$.

**Lemma 4.1.** For each $k$ and $n$ ($k, n \geq 2$),

$$l^k(n) \leq l^k(n - 1) + (3 - \sqrt{3}) \delta^k(n).$$

**Proof.** As before, we shall prove the lemma for $k = 2$, since the generalization to arbitrary $k$ will be obvious. Let $V_n \in \mathbb{V}_n$. We shall prove that

$$l(V_n) \leq l(n - 1) + (3 - \sqrt{3}) \delta(n).$$

Let $x, y \in V_n$ be such that $\delta(x, y)$ is minimized. Then $\delta(x, y) \leq \delta(n)$. Let $z$ be the median of the interval $(x, y)$ and let $V_{n-1} = [V_n - \{x, y\}] \cup \{z\}$. Then $l(V_{n-1}) \leq l(n - 1)$. Hence, it suffices to prove that $l(V_n) \leq l(V_{n-1}) + (3 - \sqrt{3}) \delta(x, y)$. Let $H$ be an optimal TSP circuit in $V_{n-1}$, and let $u, v \in V_{n-1}$ be such that the path $(u - z - v)$ is included in $H$ (see Fig. 4.1). A TSP circuit $H'$ for $V_n$ is obtained by replacing $(u - z - v)$ in $H$ by either $(u - x - y - v)$ or $(u - y - x - v)$, whichever is shorter.

Without loss of generality assume that $(u - x - y - v)$ is the shorter one, that is,

$$l(u - x - y - v) \leq l(u - y - x - v).$$

Then

$$l(V_n) \leq l(H') = l(H) \mid l(u \ x \ y \ v) \ l(u \ z \ v).$$

Hence, it suffices to show that

$$l(u - x - y - v) - l(y - z - v) \leq (3 - \sqrt{3}) \delta(x, y).$$

**Figure 4.1**
Note also that $x$ and $y$ were chosen so that
\[
\delta(x, y) = \min\{\delta(s, t) \mid s \neq t, \{s, t\} \subseteq \{u, v, x, y\}\}. \tag{2}
\]
For a given pair of points $(u, v)$, let $f(u, v) = l(u - x - y - v) - l(u - z - v)$, and let $M = \max\{f(u, v) \mid u, v \text{ satisfy (1) and (2)}\}$. (Note that even if $u, v, x, y$ are not restricted to be coplanar, $f(u, v)$ is maximized when $u, v, x, y$ are coplanar.) To prove the lemma, it suffices to show that $M \leq (3 - \sqrt{3}) \delta(x, y)$.

Let $\vec{a}, \vec{b}$ be such that $\vec{a}, x, \vec{b}, y$ form a parallelogram, in which $\delta(\vec{a}, x) = \delta(\vec{a}, y) = d(\vec{b}, x) = d(\vec{b}, y) = \delta(x, y)$ (see Fig. 4.2). Then
\[
f(\vec{a}, \vec{b}) = 3\delta(x, y) - \sqrt{3} \delta(x, y) - (3 - \sqrt{3}) \delta(x, y).
\]
Hence, the lemma will follow if we can show that $M = f(\vec{a}, \vec{b})$. To prove this, we prove the following claim:

**Claim.** $f(\vec{a}, \vec{b}) \geq f(u, v)$ for all $u, v$ which satisfy (1) and (2).

**Proof of the Claim.** Note that $f(u, v)$ can be written as
\[
f(u, v) = f_1(u) + f_2(v) + \delta(x, y),
\]
where
\[
f_1(t) = \delta(t, x) - \delta(t, z); \quad f_2(t) = \delta(t, y) - \delta(t, z).
\]

The claim now follows by the following observations:

**Observation (a).** Let $t$ be any point, and let $t'$ be in the arc $(t \quad z)$, (see Fig. 4.3). Then, by the triangle inequality, $f_1(t') \geq f_1(t)$ and $f_2(t') \geq f_2(t)$.
Observation (b). Let \( r = \delta(x, y) \), and let \( C_x, C_y \) be circles of radius \( r \) and centers \( x, y \), respectively. Let \( D_x, D_y \) denote the discs bounded by \( C_x, C_y \). Then, by (2), \( u \) and \( v \) cannot lie in the interior of \( D_x, D_y \) (see Fig. 4.4).

Let \( D = D_x \cup D_y \), and let \( C \) be the boundary of \( D \). Observations (a) and (b) imply that \( f_1(v) \) and \( f_2(u) \) (and hence also \( f(u, v) \)), are maximized for \( u \) and \( v \) which satisfy (2) when both \( u \) and \( v \) are in \( C \).

Observation (c). Let \( L \) be the line containing \( (x - y) \), and let \( a \) be the unique point in \( C_x \cap C \cap L \), and \( b \) be the unique point in \( C_y \cap C \cap L \) (see Fig. 4.4). Then, when \( u \) moves from \( a \) to \( b \) along \( C \), \( f_1(u) \) increases monotonically and \( f_2(u) \) decreases monotonically.

Observation (d): If both \( u \) and \( v \) are in \( C \), then \( l(u - x - y - v) \leq l(u - y - x - v) \) is equivalent to \( \delta(u, x) \leq \delta(v, x) \). This, together with Observations (b) and (c), implies that under constraints (1) and (2), \( f(u, v) \) is maximized when \( \delta(u, x) = \delta(v, x) \), that is, when \( u \) is the reflection of \( v \) in \( L \). Note that in this case \( f_1(u) = f_1(v) \) and \( f_2(u) = f_2(v) \). Moreover, by symmetry, \( f(u, v) \) attains its maximum for some \( u, v \) in \( C \cap C_x \). Thus, the problem of computing \( M \) reduces to the following maximization problem:

\[
\begin{align*}
\text{maximize} & \quad f_1(u) + f_2(u) = f_1(u) + f_2(v), \\
\text{subject to} & \quad u \in C \cap C_x.
\end{align*}
\]

Using polar coordinates by substituting \( u = (r, \alpha) \) (see Fig. 4.5), we get

\[
\begin{align*}
\text{maximize} & \quad r(1 + 2 \cos(\alpha/2) - (5 + 4 \cos \alpha)^{1/2}) = f(\alpha), \\
\text{subject to} & \quad 0 \leq \alpha \leq 2\pi/3, \quad (r = \delta(x, y)).
\end{align*}
\]
OPTIMAL TSP CIRCUITS

(In polar coordinates, \( \delta(u, x) = r \), \( \delta(u, y) = 2r \cos(\alpha/2) \), and \( \delta(u, z) = r(5 + 4 \cos \alpha)^{1/2}/2 \)). Taking the derivative with respect to \( \alpha \), we get

\[
f'(\alpha) = r(-\sin(\alpha/2) + 2 \sin \alpha/(5 + 4 \cos \alpha)^{1/2}).
\]

Since \( f'(\alpha) > 0 \) for \( \alpha \in (0, 2\pi/3] \), \( f \) attains its maximum when \( \alpha = 2\pi/3 \), i.e., when \( u = \bar{u}, v = \bar{v} \). This completes the proof of the claim, and hence of the lemma.

**Theorem 4.1.** For all \( k \) and \( n \)

\[
l_k(n)/(n\delta^k(n)) \leq (3 - \sqrt{3})k/(k - 1) + o(1).
\]

**Proof.** Let \( c_k = \lim_{n \to \infty} \delta^k(n) n^{1/k} \) (the existence of \( c_k \) follows from Theorem 3.2'). Then

\[
\delta^k(n) - c_k n^{-1/k} + o(n^{-1/k}).
\]

The theorem will follow if we can show that

\[
l_k(n) \leq [(3 - \sqrt{3})k/(k - 1)] c_k n^{1/k} + o(n^{1/k}).
\]

By Lemma 4.1, \( l_k(n) \leq l_k(n - 1) + (3 - \sqrt{3}) \delta^k(n) \). Hence,

\[
l_k(n) \leq (3 - \sqrt{3})[\delta^k(2) + \cdots + \delta^k(n)] = (*).
\]

Let \( \delta^k(x) \) be a continuous, nonincreasing real extension of \( \delta^k(n) \). Then \( \delta^k(x) = c_k x^{-1/k} + o(x^{-1/k}) \). We get that

\[
l_k(n) \leq (*) < (3 - \sqrt{3}) \int_2^n \delta^k(x) \, dx
\]

\[
= (3 - \sqrt{3}) \int_2^n [c_k x^{-1/k} + o(x^{-1/k})] \, dx
\]

\[
= [(3 - \sqrt{3})k/(k - 1)] c_k n^{1/k} + o(n^{1/k}).
\]

**Theorem 4.2.** For all \( k \), \( l_k(n) < [(3 - \sqrt{3})k/(k - 1)] n^{1/k} + o(n^{1/k}). \)

**Proof.** In view of Theorem 4.1, it is enough to show that for large enough \( n \), \( \delta^k(n) < n^{-1/k} \).

Let \( \delta^k(n) = d \). Then it is possible to pack \( n \) disjoint \( k \)-dimensional spheres of diameter \( d \) in a \( k \)-dimensional sphere of diameter \( 1 + d/2 \). It is known that the ratio between the sum of the volumes of the small spheres and the volume of the large sphere cannot exceed some constant \( E_k < 1 \). The ratio
between the volume of one small sphere and the volume of the large sphere is
\[ \frac{d}{(1 + d/2)} \] . Hence
\[ n \left( \frac{d}{(1 + d/2)} \right)^k \leq E_k, \]
or, equivalently,
\[ d^k \leq (1 + d/2)^k E_k/n. \]
For large enough \( n \), \( (1 + d/2)^k E_k < 1 \) (recall that \( d = \delta^k(n) \)) which implies that, for large enough \( n \), \( \delta^k(n) < n^{-1/k} \).

A problem similar to the one discussed in this section was discussed in [5, 8], where a bound on the length of the shortest road through \( n \) points in the \( k \)-dimensional unit cube was investigated. In [5] it was shown that this bound cannot exceed
\[ [k(2(k - 1))^{(1-k)/2k} + o(1)] n^{1-1/k}. \]  
(4.1)
Using an argument similar to the one in Theorem 4.2, but replacing the \( k \)-dimensional sphere of diameter 1 + \( d/2 \) by a \( k \)-dimensional cube of side length 1 + \( d/2 \), and using the fact that the volume of the \( k \)-dimensional sphere of radius 1 is \( \pi^{k/2}/\Gamma(k/2 + 1) \), one can show that this bound cannot exceed
\[ [(2 - \sqrt{3})k/(k - 1)](\Gamma(k/2 + 1))^{1/k}(\pi^{-1/2}) + o(1) \]  
(4.2)
For large \( k \) we have
\[ (4.1) \approx [(k/2)^{1/2} + o(1)] n^{1-1/k} \approx 0.7071 \sqrt{k} n^{1-1/k} \]
and
\[ (4.2) \approx [(3 - \sqrt{3})(2k/\pi e)^{1/2} + o(1)] n^{1-1/k} \approx 0.6136 \sqrt{k} n^{1-1/k} \]
In fact, (4.2) gives a better bound than (4.1) already for \( k = 7 \) (a constant of 2.370(-) vs. a constant of 2.413(+)). However, for \( k < 7 \) the technique used in [5] provides a better bound on \( l^k(n) \). In particular, one can use that technique to prove

**Theorem 4.3.** \( l(n) = (\pi n/2)^{1/2} + O(1) \).

**Proof (sketch).** Let \( V_n \subset V_n \) be given. Then \( V_n \) can be embedded in a figure \( C \) of diameter 1 which, by Barbier's theorem, is of area \( \leq \pi/4 \).

Let \( t = (\pi/2n)^{1/2} \), and let \( L_0, L_1 \) be the sets of lines defined by
\[ L_0 = \{(x, y) \mid y = nt \text{ for some integer } n\}; \]
\[ L_1 = \{(x, y) \mid y = (n + \frac{1}{2})t \text{ for some integer } n\}. \]
For a point $v$ in $R^2$, let $\delta(v, L_i)$ be the shortest distance from $v$ to a line in $L_i$. Then for each $v$, $\delta(v, L_0) + \delta(v, L_1) = t/2$. Hence,

$$\sum_{v \in V_n} \delta(v, L_0) + \sum_{v \in V_n} \delta(v, L_1) = nt/2.$$ 

Hence, for $i = 0$ or $i = 1$, it holds that

$$\sum_{v \in V_n} \delta(v, L_i) \leq nt/4. \quad (4.3.1)$$

Without loss of generality, assume that (4.3.1) holds for $i = 0$. Consider the TSP circuit composed of:

(a) The line segments in $L_0 \cap C$.

(b) Portions of the boundary of $C$ connecting these line segments to a path.

(c) For each point $v$ in $V_n$, the shortest line segment connecting $v$ to the path described above, each such segment counted twice.

(d) A segment connecting the first and last points of $V_n$ traversed along the described path.

The sum of the lengths of the line segments described in (a) is equal approximately to the area of $C$ divided by $t$, and hence it is at most $\pi/4t + O(1)$.

The sum of the lengths of the segments in (b) is $O(1)$. The sum of the lengths of the segments in (c) (each taken twice) is $nt/2$, and the segment (d) is of length $\leq 1$. Altogether, the total length of the described circuit is $\pi/4t + nt/2 + O(1)$. The theorem follows. 

5. Two Related Results

Two problems which are related to the problem discussed in this paper are:

(1) Minimal tree: Given a set $V$ of $n$ points in $R^k$, find a tree (i.e., a connected graph without circuits) on $V$ such that the length of the tree, defined as the sum of the lengths of its arcs, is minimal. Denote this length by $l^*_T(V)$.

(2) Steiner tree: Given a set $V$ as above, find a set $V' \supseteq V$ such that $l^*_S(V')$ is minimal. Formally, for a given $V$, the length of the Steiner tree of $V$ is defined by

$$l^*_S(V) = \min_{V' \supseteq V} \{l^*_T(V')\}.$$
Note. The existence of a set $V' \supseteq V$ such that $l^k_T(V')$ is minimal follows from the observation that for every set $V'$ which contains $V$ there exists a set $V''$ which contains $V$ such that

$$l^k_T(V'') \leq l^k_T(V') \quad \text{and} \quad |V''| \leq 2|V| - 2.$$ 

(Thus, in the definition of a Steiner tree we can add the restriction: $|V'| \leq 2|V| - 2$, which implies that the minimum is attained.) $V''$ is constructed from $V'$ in the following manner: Let $T$ be a tree of minimal length on $V'$. Delete from $V'$ all the points which are not in $V$ and have degree at most 2 in $T$. In the resulting tree every point not in $V$ has a degree at least 3. The observations follow.

The corresponding problems for graphs of bounded diameter are: For each $n$ find:

1. $l^k_T(n) = \max_{V \in \mathcal{V}_n^k} \{l^k_T(V)\}$.

2. $l^k_S(n) = \max_{V \in \mathcal{V}_n^k} \{l^k_S(V)\} = \max_{V \in \mathcal{V}_n^k} \{\min_{V' \supseteq V} l^k_T(V')\}$

**Theorem 5.1.** For each $k$ and $n$,

$$1 - 1/n \leq \frac{l^k_T(n)}{n\delta^k(n)} \leq (3 - \sqrt{3})k/(k - 1) + o(1).$$

**Proof.** The upper bound follows immediately from the upper bound on $l^k(n)$ (Theorem 4.1). The lower bound follows from the fact that a tree on $n$ points has $n - 1$ edges and from the definition of $\delta^k(n)$. 

**Theorem 5.2.** For each $k$ and $n$,

$$\frac{1}{2} \leq \frac{l^k_S(n)}{n\delta^k(n)} \leq k/(k - 1) + o(1)$$

**Proof.** The upper bound follows from the observation that $l^k_S(n) \leq l^k_S(n - 1) + \delta^k(n)$. (In the proof of Lemma 4.1, simply add $(z - x)$ and $(z - y)$ to the Steiner tree for $V_{n-1}$, see Fig. 5.1.)
The lower bound follows from the lower bound on $l^k(n)$, by the fact that the existence of a Steiner tree of length $l$ implies the existence of a Hamiltonian circuit of length $\leq 2l$.

**APPENDIX**

*Proof of Lemma 2.7.* Let $V = V^*_2 = \{x, y, z, u, v\}$ be a longest set in $V_{5,3}$. We shall prove that $V$ satisfies conditions (1)–(3) of Lemma 2.7. As in Lemma 2.6, we may assume that $\{x, y, z\}$ are the vertices of an equilateral triangle, $T$. There are 12 Hamiltonian circuits in $V$, each uses either one or two sides of $T$. We denote as $(x - y)$-circuit a circuit that intersects $T$ with the edge $(x - y)$, as $(x - y - z)$-circuit a circuit that intersects $T$ with the edges $(x - y)$ and $(y - z)$, etc. The 12 circuits are listed below:

- $(x - y - z)$-circuits: $C_1 = (x - y - z - u - v - x)$
  $C_2 = (x - y - z - v - u - x)$
- $(y - z - x)$-circuits: $C_3 = (x - u - v - y - z - x)$
  $C_4 = (x - v - u - y - z - x)$
- $(z - x - y)$-circuits: $C_5 = (x - y - u - v - z - x)$
  $C_6 = (x - y - v - u - z - x)$
- $(x - y)$-circuits: $C_7 = (x - y - u - z - v - x)$
  $C_8 = (x - y - v - z - u - x)$
- $(y - z)$-circuits: $C_9 = (x - u - y - z - v - x)$
  $C_{10} = (x - v - y - z - u - x)$
- $(z - x)$-circuits: $C_{11} = (x - u - y - v - z - x)$
  $C_{12} = (x - v - y - u - z - x)$

Conditions (1)–(3) of Lemma 27 are equivalent to the following

![Figure A.1](attachment:figure_A.1.png)
statement: In the longest set $V$, $u$ and $v$ lie on one of $T$'s heights, such that the lengths of all the (Hamiltonian) circuits that do not intersect themselves are equal to $l(V)$. (If that height is $h_x$ and $u$ lies above $v$, then these circuits are $C_2$, $C_3$, $C_5$, $C_6$, $C_8$, and $C_{11}$.)

The proof of the lemma proceeds by the following two claims:

Claim 1. Both $u$ and $v$ lie on one of $T$'s heights.

Proof. We shall assume that Claim 1 is false, and derive a contradiction. Denote by $N_x$ the set of points $t$ in $T$ s.t. $\delta(t, x) \leq \min\{\delta(t, y), \delta(t, z)\}$. $N_y$ and $N_z$ are defined similarly (see Fig. A.1). We consider three cases:

Case (i). For some $t \in \{x, y, z\}$, exactly one out of $\{u, v\}$ is an interior point of $N_t$ (see Fig. A.2). Without loss of generality we may assume that $u$ is an interior point of $N_u$, and that $u$ is in $N_v$ (not necessarily as an interior point). Then we have the following inequalities:

$$\delta(u, z) < \min\{\delta(u, y), \delta(u, x)\}$$
$$\delta(v, y) < \min\{\delta(v, x), \delta(v, z)\}.$$

The above inequalities imply that $C_2$, $C_3$, $C_4$, and $C_5$ are longer than $C_6$, and hence are not optimal. They also imply that $C_8$ and $C_{11}$ are longer than $C_8$, hence $C_9$ and $C_{11}$ are not optimal either. But all the remaining circuits use the edge $(u - z)$, which must therefore be an essential edge, in contradiction with Lemma 2.3(a)\(^6\).

Case (ii). Both $u$ and $v$ are interior points of some $N_t$ for some $t$. Without loss of generality $t = z$. Then

$$\delta(u, z) < \min\{\delta(u, y), \delta(u, x)\}$$
$$\delta(v, z) < \min\{\delta(v, y), \delta(v, x)\}.$$

\(^6\) Actually, with the straightforwards generalization of Lemma 2.3(a) to longest sets in $V_{n,k}$. 
We may also assume that

\[ \delta(u, y) + \delta(v, x) \leq \delta(v, y) + \delta(u, x). \]

Then \( C_i \) is shorter than \( C_9, C_{10}, C_{11}, C_{12} \) and is not longer than \( C_8 \). \( C_3 \) and \( C_4 \) are longer than \( C_2 \) and \( C_1 \), respectively. Thus, the only candidates for optimal circuits are \( C_1, C_2, C_5, C_6, C_7, \) and \( C_8 \). To show that this yields a contradiction we use an extension of the idea in Lemma 2.3(a): Let \( S = \{(z-u), (z-v)\} \) and \( L = \{(x-u), (x-v), (y-u), (y-v)\} \). Each of the candidate circuits above uses one or two edges of \( S \), and the same number of edges of \( L \). By removing \( u \) and \( v \) a small distance \( h \) away from \( z \) along \( (z-u) \) and \( (z-v) \), respectively, the lengths of the edges in \( S \) increase by \( h \), the lengths of the edges in \( L \) decrease by less than \( h \), and the lengths of the rest of the edges do not decrease. Hence, the lengths of all the candidate circuits above increase. By taking \( h \) small enough, a set \( V^1 \) which is longer than \( V \) is obtained, a contradiction.

**Case (iii).** \( u \) and \( v \) are not interior points of any \( N_i \); that is, \( u \) and \( v \) lie on different heights of \( T \), as shown in Fig. A.3, with \( u \) on \( h_y \) and \( v \) on \( h_x \). In this case we have:

\[ l(C_2) = l(C_3) = l(C_6) < \min \{l(C_1), l(C_4), l(C_5)\} \]

and

\[ l(C_8) < l(C_i) \quad \text{for} \quad i = 7, 9, 10, 11, 12. \]

**Subcase (iii.1)** \( l(C_8) < l(C_2) \). In this case \( C_8 \) is a unique optimal circuit, in contradiction to Lemma 2.3(a), (b).

**Subcase (iii.2)** \( l(C_2) < l(C_8) \). In this case the optimal circuits are \( C_2, C_3, C_6 \). This means that \( (u-v) \) is essential, in contradiction the Lemma 2.3(a).

**Subcase (iii.3)** \( l(C_2) = l(C_8) \). The proof of this case is a little more
involved, and is sketched below. The equality above is easily seen to be equivalent to

\[ \delta(z, u) + \delta(z, v) = 1 + \delta(u, v) \]  

(\star)

It also implies that \( C_8 \) is optimal and hence

\[ l(V) = 1 + 2(\delta(u, z) + \delta(v, z)). \]  

(\star\star)

Hence, if \( V \) is longest then (\star\star) is maximized under the constraint (\star). We shall show that this cannot happen if both \( u \) and \( v \) are different from the center \( O \).

Let \( a = \varepsilon uzO, \beta = \varepsilon vzO, \alpha_1 = \varepsilon vuO, \beta_1 = \varepsilon uvO \) (see Fig. A.3). Assume first that \( a = \beta \) (and hence \( a_1 = \beta_1 \)). Then

\[ \delta(z, u) = \delta(z, v) = \frac{1}{2} \cos(30 - \alpha) \]

\[ \delta(u, v) = 2\delta(z, u) \sin \alpha = \frac{\sin \alpha}{\cos(30 - \alpha)} \]

Thus, by (\star)

\[ \frac{1}{\cos(30 - \alpha)} = 1 + \frac{\sin \alpha}{\cos(30 - \alpha)} \]

and hence \( \alpha \approx 5.2644 \), which implies that \( l(V) = l(C_8) \approx 3.2020 \), which is smaller than \( l(V^*) \approx 3.2175 \), where \( V^* \) is the set defined in Lemma 2.7.

Second, assume that \( a > \beta \), and hence \( a_1 < \beta_1 \). For a small \( h \), let \( u(h) \) and \( v(h) \) be the points on \( h_1 \) and \( h_2 \) for which

\[ \delta(z, u(h)) = \delta(z, u) + h \]

\[ \delta(z, v(h)) = \delta(z, v) - h \]

For infinitesimal \( h \) we have

\[ \delta(u(h), u) \approx \frac{h}{\sin(30 - \alpha)}, \]

\[ \delta(v(h), v) \approx \frac{h}{\sin(30 - \beta)} \]

and

\[ \delta(u(h), v(h)) - \delta(u, v) \approx \delta(v(h), v) \cos \beta_1 - \delta(u(h), u) \cos \alpha_1. \]

\footnote{All angles are measured in degrees.}
Thus, if we define

\[ \frac{d(u, v)}{d(h)} = \frac{\delta(u(h), v(h)) - \delta(u, v)}{h} \]

then we get

\[ \lim_{h \to 0} \left( \frac{d(u, v)}{d(h)} \right) = \frac{\cos \beta_1}{\sin(30 - \beta)} - \frac{\cos \alpha_1}{\sin(30 - \alpha)} < 0. \tag{since \( \alpha > \beta \) and \( \alpha_1 < \beta_1 \).} \]

Let \( V(h) = \{x, y, z, u(h), v(h)\} \), and let \( C_i(h) \) be \( C_i \) in \( V(h) \). For small negative \( h \), \( C_8(h) \) and \( C_2(h) \) are shorter than all other circuits in \( V(h) \). By the definition of \( u(h) \) and \( v(h) \), \( l(C_8(h)) = l(C_8) \) and by the inequality above,

\[ l(C_2(h)) = l(C_2) + \delta(u(h), v(h)) - \delta(u, v) > l(C_2). \]

Since \( l(C_2) = l(C_8) \), this implies that \( l(C_2(h)) > l(C_8(h)) \). Thus

\[ l(V(h)) = l(C_8(h)) = l(C_8) = l(V) \]

and hence, by the assumption on \( V \), \( V(h) \) is longest. But \( V(h) \) has a unique optimal circuit—\( C_8(h) \)—in contradiction with Lemma 2.3. This completes the proof of Claim 1.

**Note.** The proof above implies not only that \( u \) and \( v \) lie on one of the heights, but also that they are not interior points of \( N_t \) for \( t \in \{x, y, z\} \) (see Fig. A.4).

Assume that \( u \) and \( v \) lie on \( h_x \), and that \( u \) lies above \( v \) (as in Fig. A.4). It is easily observed that the \((x - y - z)\)-circuits are reflections through \( h_x \) of the \((y - z - x)\) circuits, that the \((x - y)\)-circuits are (similar) reflections of the \((z - x)\)-circuits, and that \( C_5 \) is a reflection of \( C_6 \). Using this, one can verify that the only candidates for optimal circuits are \( C_2, C_5, C_8 \), and their
reflections. If all of them are optimal, then (2) and (3) of Lemma 2.7 must be satisfied, and hence the set \( V \) is the one described in Lemma 2.7 (and shown in Fig. 2.8), as claimed. Thus, it remains to show that the circuits \( C_2, C_5, \) and \( C_8 \) are optimal.

**Claim 2.** \( C_2, C_5, \) and \( C_8 \) are optimal in \( V \).

**Proof:** Assume that the claim is false. Then the longest set in \( V_{3,1}, V \), differs from the set \( V^* \) in Fig. 2.8. Let \( V^* = \{x, y, z, u^*, v^*\} \). By the note above \( u \) cannot lie above \( u^* \). We consider three cases, according to the locations of \( u \) and \( v \).

**Case (i):** \( u = u^* \) and \( v \) lies above \( v^* \). Then, by the triangle inequality

\[
\delta(u^*, v^*) + \delta(v^*, z) > \delta(u, v) + \delta(v, z),
\]

hence

\[
l(V) < l(C_2) < l(C_2^*) = l(V^*)
\]

which contradicts the assumption on \( V \).

**Case (ii).** \( u = u^* \) and \( v \) lies below \( v^* \). Then

\[
\delta(y, v) < \delta(y, v^*),
\]

hence

\[
l(V) < l(C_8) < l(C_8^*) = l(V^*),
\]

which again contradicts the assumption on \( V \).

**Case (iii).** \( u \) lies below \( u^* \). In this case \( \delta(u, x) > \delta(u, y) \), which implies that \( C_2 \) is longer than \( C_5 \), and hence is not optimal. Hence, the only candidates for optimality are \( C_5, C_8 \) and their reflections through \( h_x \). If any of them is not optimal then Lemma 2.3 is violated. It follows that both \( C_5 \) and \( C_8 \) must be optimal. Thus

\[
l(C_5) = 2 + \delta(y, v) + \delta(u, v) + \delta(u, z) = l(C_8)
\]

\[
= 1 + 2\delta(y, v) + \delta(u, z) + \delta(u, z) + \delta(u, x).
\]

In order to use the Lagrange multipliers theorem we restate the above equality, after rearranging and cancelling equal terms, in terms of the scalar variables \( u' \) and \( v' \) whose range of definition is \( h_x \) as follows: The constraint

\[
\delta(v', z) + \delta(u', x) = 1 + \delta(u', v')
\]

is satisfied at \( u' = u \) and \( v' = v \). The constraint (\&) is also satisfied at
Let \( u' = u_1 \) and \( v' = v_1 \), where \( v_1 \) is at the bottom \( B \) of \( h_x \) (see Fig. A.5), and \( u_1 \) satisfies \( \tan(\xi u_1 yz) = (\sqrt{3} - 1)/2 \) (we leave the verification of this fact to the reader). Let \( V_1 = \{x, y, z, u_1, v_1\} \). Then since \( l(V_1) < l(V^*) \), \( V_1 \) is not optimal; hence \( V_1 \neq V \), which implies that \( v_1 \neq v \). It follows that both \( u \) and \( v \) are internal points of the segments \((O, B)\). Hence, by the assumption that \( V \) is longest, the function

\[
l(V) = l(C_x) = 2 + \delta(y, v') + \delta(u', v') + \delta(u', z) = f(u', v')
\]

attains a local maximum, under the constraint (\&), at \( u' = u \) and \( v' = v \). Let \( \alpha' = \xi u'yz, \beta' = \xi v'yz, \alpha = \xi uyz \) and \( \beta = \xi vyz \). Then multiplying (\&) and (\&\&) above by 2 and expressing them in terms of \( \alpha' \) and \( \beta' \), we get (after rearranging terms) that under the constraint

\[
\frac{1}{\cos \beta'} - 2 \tan \alpha' + \tan \beta' = \text{Const}_1
\]

the function

\[
\frac{1}{\cos \beta'} + \frac{1}{\cos \alpha'} + \tan \alpha' - \tan \beta' + \text{Const}_2
\]

does not attain a local maximum at \( \alpha' = \alpha \) and \( \beta' = \beta \). Thus, by the Lagrange multipliers theorem there exists a constant \( \lambda \) for which the partial derivatives of the function \( G(\alpha', \beta') \) defined below with respect to \( \alpha' \) and \( \beta' \) vanish at \( \alpha' = \alpha \) and \( \beta' = \beta \).

\[
G(\alpha', \beta') = \frac{1}{\cos \beta'} + \frac{1}{\cos \alpha'} + \tan \alpha' - \tan \beta' + \lambda \left( \frac{1}{\cos \beta'} - 2 \tan \alpha' + \tan \beta' \right).
\]
Taking derivatives, we get that at $\alpha' = \alpha$ and $\beta' = \beta$:

$$
\frac{dG}{d\alpha'} = \frac{\sin \alpha + 1 - 2\lambda}{\cos^2 \alpha} = 0
$$

$$
\frac{dG}{d\beta'} = \frac{(1 + \lambda) \sin \beta + \lambda - 1}{\cos^2 \beta} = 0.
$$

Thus, we get

$$
\sin \alpha = 2\lambda - 1,
$$

$$
\sin \beta = \frac{1 - \lambda}{1 + \lambda},
$$

hence

$$
\cos \alpha = 2\sqrt{\lambda - \lambda^2}
$$

$$
tg \alpha = \frac{2\lambda - 1}{2\sqrt{\lambda - \lambda^2}}
$$

$$
\cos \beta = \frac{2\sqrt{\lambda}}{1 + \lambda}
$$

$$
tg \beta = \frac{1 - \lambda}{2\sqrt{\lambda}}.
$$

Given $\alpha$ and $\beta$, the length of $C_8$ as a function of $\lambda$ is

$$
1 + \frac{1}{\cos \beta} + \frac{1}{2 \cos \alpha} + \frac{\sqrt{3} - \text{tg} \alpha}{2}
$$

$$
= 1 + \frac{1 + \lambda}{2\sqrt{\lambda}} + \frac{1}{4\sqrt{\lambda - \lambda^2}} + \frac{\sqrt{3}}{2} - \frac{2\lambda - 1}{4\sqrt{\lambda - \lambda^2}}
$$

$$
= \frac{2 + \sqrt{3}}{2} + \frac{1 + \lambda}{2\sqrt{\lambda}} + \frac{1 - \lambda}{2\sqrt{\lambda - \lambda^2}}.
$$

In a similar way, the length of $C_5$ is shown to be

$$
2 + \frac{\lambda}{2\sqrt{\lambda - \lambda^2}} + \frac{\sqrt{\lambda}}{2}.
$$

Since $l(C_4) = l(C_8)$, we obtain

$$
\frac{1}{\sqrt{\lambda}} + \frac{1 - 2\lambda}{\sqrt{\lambda - \lambda^2}} + \sqrt{3} - 2 = 0.
$$
Solving the above for \( \lambda \), we get
\[
\lambda \approx 0.7089055
\]
which implies that \( \alpha \approx 24.70 \), \( \beta \approx 9.81 \), and that
\[
l(V) = l(C_5) = l(C_8) \approx 3.201257 < l(V^*) \approx 3.2175
\]
and hence \( V \) is not longest. This contradiction completes the proof of Claim 2, and hence of Lemma 2.7.  

**ACKNOWLEDGMENTS**

I would like to thank Yaron Gold for stimulating discussions which led to this research. I would also like to thank an anonymous referee for correcting some errors in the original manuscript and for numerous suggestions which improved the presentation of the paper.

**REFERENCES**