Simplicity and Hardness of the Maximum Traveling Salesman Problem under Geometric Distances*

Sándor P. Fekete†

Abstract

Recently, Barvinok, Johnson, Woeginger, and Woodroofe have shown that the *Maximum TSP*, i. e., the problem of finding a traveling salesman tour of maximum length, can be solved in polynomial time, provided that distances are computed according to a polyhedral norm in $\mathbb{R}^d$, for some fixed d. The most natural case of this class of problems arises for rectilinear distances in the plane $\mathbb{R}^2$, where the unit ball is a square. With the help of some additional improvements by Tamir, the method by Barvinok et al. yields an $O(n^2 \log n)$ algorithm for this case by making elegant use of geometry, graph theory, and optimization, including some rather powerful tools.

In this paper, we present a simple algorithm with $O(n)$ running time for computing the length of a longest tour for a set of points in the plane with rectilinear distances. The algorithm does not use any indirect addressing, so its running time remains valid even in comparison based models in which sorting requires $\Omega(n \log n)$ time, which implies the same lower bound on verifying a Hamiltonian cycle. In addition, our approach gives a simple characterization of *all* optimal solutions. These results give a good idea what makes the (polyhedral) max TSP so much easier than its minimization counterpart.

Resolving the complexity status of the max TSP for Euclidean distances in spaces of fixed dimension has been stated by Barvinok et al. as a main open problem. In this paper, the results on simplicity are complemented by a proof that the Maximum TSP under Euclidean distances in $\mathbb{R}^d$ for any fixed $d \geq 3$ is NP-hard, shedding new light on the well-studied difficulties of Euclidean distances. In addition, our result implies NP-hardness of the Maximum TSP under polyhedral norms if the number k of facets of the unit ball is not fixed. As a corollary, we get NP-hardness of the Maximum Scatter TSP for geometric instances, where the objective is to find a tour that maximizes the shortest edge. This resolves a conjecture by Arkin, Chiang, Mitchell, Skiena, and Yang in the affirmative.

1 Introduction

The Traveling Salesman Problem (TSP) is one of the classical problems of combinatorial optimization: Given a set $\{v_1, v_2, \ldots, v_n\}$ of vertices together with the distance $d(v_i, v_j)$ between every pair of distinct vertices v_i, v_j, the goal is to find a permutation π of the vertices (a "tour") that minimizes (Minimum TSP) or maximizes (Maximum TSP) the total tour length

$$d\left(v_{\pi(n)}, v_{\pi(1)}\right) + \sum_{i=1}^{n-1} d\left(v_{\pi(i)}, v_{\pi(i+1)}\right).$$

Geometric instances of the TSP have always been of particular interest: vertices v_i correspond to points $p_i = (x_1, \ldots, x_d)$ in space $\mathbb{R}^d$, and distances $d(v_i, v_j)$ are given by some geometric norm $\|p_i - p_j\|$. The most common norms considered include the Euclidean norm L_2 and the Manhattan norm L_1, which are both special cases of the L_p norms. The L_1 norm is also an example of a *polyhedral norm*, where the set of points at distance 1 from the origin is given by a centrally symmetric polyhedron with k facets.

Two key questions regarding the complexity of the Minimum TSP on geometric instances have been answered. Itai, Papadimitriou, and Swarcfiter [11] showed that the Minimum TSP is NP-hard for any fixed dimension $d \geq 2$ and any L_p or polyhedral norm. On the other hand, Arora [2, 3] and Mitchell [13] showed that all these geometric instances allow a polynomial-time approximation scheme (PTAS), i. e., a sequence of algorithms A_s that compute a solution within a factor of $1 + \frac{1}{s}$ of the optimum, in time that is polynomial for any fixed s. This highlights the special role of geometry, since it is well known that for general Minimum TSP instances, no PTAS can exist. (See Trevisan [17], and Papadimitriou and Yannakakis [14] for such results in more restricted situations.)

*Work on this paper was partially supported by the Hermann-Minkowski-Minerva Center for Geometry at Tel Aviv University, while the author was visiting the Center in March 1998, and by DFG travel grant FE 407/3-1 for a visit to the USA in June 1998.

†Center for Applied Computer Science, Universität zu Köln, 50923 Köln, GERMANY, sandor@zpr.uni-koeln.de

A third question that is still unresolved is the issue whether the TSP under Euclidean distances is a member of the class NP, allowing polynomial time verification of a good solution. The difficulty of this question arises from the fact that there are no polynomial bounds known on the accuracy that is necessary for comparing a sum of square roots to a given integer. See the book [12] for a discussion.

The complexity of the Maximum TSP for geometric distances has been less clear. Barvinok [5] showed that there is a PTAS for the Maximum TSP under all metrics in $\mathbb{R}^d$, for any fixed d. Very surprisingly, Barvinok, Johnson, Woeginger, and Woodroofe [6] showed that under polyhedral norms with a fixed number k of facets on the unit ball, the Maximum TSP is indeed polynomial. Making elegant use of geometry, graph theory, and optimization, including some rather powerful tools, their method yields an $O(n^{f-2} \log n)$ algorithm, where f is the number of facets of the unit ball defining the polyhedral norm. For the natural case of points in the plane under rectilinear distances, this yields a running time of $O(n^2 \log n)$.

In this paper, we present a simple algorithm with $O(n)$ running time for computing the length of a longest tour for a set of points in the plane with rectilinear distances. The basic tool of the algorithm is a rectilinear "minimum star", i. e., a connection of all points to one center point, such that the total rectilinear distance to the center is minimized. A minimum star in $\mathbb{R}^2$ can be found in $O(n)$ time by two median computations, one for each coordinate, using the linear time method by Blum, Floyd, Pratt, Rivest, and Tarjan [8]. With the help of the star center, it is possible to determine the optimal value of an optimal 2-factor in linear time, i. e., a set of subtours of minimum total length covering the points. More precisely, we can show that all points can be covered by at most two subtours, with a total length equal to twice the length of a minimum star. The property that makes these subtours optimal is the fact that any connection between two points has the same length as a connection via the center of the star – in other words, the triangle inequality is kept tight.

Using a subtour constraint of the integer linear programming formulation of the TSP, we can give an upper bound on the total length of a maximum tour, and a simple proof that there are many tours meeting this bound. It turns out that any optimal tour arises from an optimal 2-factor by one cheapest possible 2-exchange, i. e., a merging of the two subtours by swapping two edges. Using geometry, an optimal 2-exchange can be found in linear time. Conversely, it follows that any optimal tour can be found in this way. This yields a compact characterization of the set of optimal tours, and a proof that the number of optimal tours is $\Omega\left(\left(\frac{n}{4}!\right)^4\right)$, i. e., exponentially large. The linear time algorithm can be generalized to all cases where the unit ball is a symmetric quadrangle.

The method by Barvinok et al. yields a polynomial approximation scheme for the Maximum TSP under arbitrary norms, since any unit ball can be approximated sufficiently well by a polyhedron with a bounded number of faces. *"The complexity of the Maximum TSP with Euclidean distances and fixed d remains unsettled, however..."* (Barvinok et al.) In the second part of this paper, this question is settled with a proof that the Maximum TSP with Euclidean distances is NP-hard for any fixed $d \geq 3$. One of the consequences is NP-hardness of the Maximum TSP for polyhedral norms with an unbounded number of facets on the corresponding unit ball. Another consequence concerns the so-called *Maximum Scatter TSP*, where the objective is to find a tour that maximizes the shortest edge. The Maximum Scatter TSP was first considered by Arkin, Chiang, Mitchell, Skiena, and Yang [1], and the complexity for geometric instances was stated as an open problem. Our result implies NP-hardness for Euclidean instances in 3-dimensional space.

The rest of this paper is organized as follows. In Section 2, the linear time algorithm for the Maximum TSP under rectilinear distances in the plane is described. Section 3 contains the NP-hardness proof for the Maximum TSP under Euclidean distances in $\mathbb{R}^3$. Section 4 concludes with extensions and open problems.

2 An $O(n)$ Algorithm

In this section, we describe a linear time algorithm for determining the length of an optimal tour for the Maximum TSP under rectilinear distances in the plane.

2.1 Stars and Matchings

Our construction uses properties of so-called *stars*; a star for a given set of vertices V is a minimum Steiner tree with precisely one Steiner point (the *center*) that contains all vertices in V as leafs. The total length of the vertices in a star is an upper bound on any matching in V, since any edge (v_i, v_j) in the matching can be mapped to a pair of edges (v_i, c) and (c, v_j) in the star, and by triangle inequality, $d(v_i, v_j) \leq d((v_i, c) + d(c, v_j)$. The worst case ratio between the total length of a minimum star $\min S(P)$ and a maximum matching $\max M(P)$ has been posed by Suri as an open problem [15] for the case of Euclidean distances. This ratio plays a crucial role in different types of optimization problems. See the paper by Fingerhut, Suri, and Turner [10] for applications in the context of broadband communication networks.

Also, Tamir and Mitchell [16] have used the duality between minimum stars and maximum matchings for showing that the core of certain matching games have a nonempty core. A solution to Suri's problem [15] and several extensions can be found in a forthcoming paper by Fekete and Meijer [9].

For rectilinear distances, determining the length $\min S(P)$ of a minimum length star (also known as the *Weber problem*) can be determined in linear time. Since we will need some basic properties, we give a short proof of the following lemma:

LEMMA 2.1. *For a given set of n points $P = \{p_1, \ldots, p_n\}$ with $p_i = (x_i, y_i)$, an optimal star center c can be determined in linear time.*

Proof. For any star center $c = (x_c, y_c)$, let $P_x^- := \{p_i \in P \mid x_i < x_c\}$, $P_x^+ := \{p_i \in P \mid x_i > x_c\}$, $P_y^- := \{p_i \in P \mid y_i < y_c\}$, $P_y^+ := \{p_i \in P \mid y_i > y_c\}$. It is easy to see that for an optimal star center, we get the local optimality conditions

$$n_x^- := |P_x^-| \le \frac{n}{2} \tag{2.1}$$

$$n_x^+ := |P_x^+| \ge \frac{n}{2} \tag{2.2}$$

$$n_y^- := |P_y^-| \le \frac{n}{2} \tag{2.3}$$

$$n_y^+ := |P_y^+| \ge \frac{n}{2}. \tag{2.4}$$

This implies that we can compute an optimal star center by choosing x_c as a median of the x_i, and y_c as a median of the y_i. It was shown by Blum, Floyd, Pratt, Rivest, and Tarjan [8] how to compute a median of n numbers in $O(n)$ time. The claim follows. ∎

It should be noted that for Euclidean distances, the problem of determining $\min S(P)$ is considerably harder: It was shown by Bajaj [4] that an optimal star center for five points in the plane is in general not solvable by radicals over the field of rationals. This implies that an algorithm for computing $\min S(P)$ must use stronger tools than provided by constructions by straight edge and compass.

The following argument from the paper [9] leads to Theorem 2.1 and is used with friendly permission by Henk Meijer. Since some of these basic properties are closely connected to the results of this paper, we give a brief sketch and introduce some basic notation that will also be used later. It should be noted that independent from our work, Theorem 2.1 appears in the paper by Tamir and Mitchell [16] as Theorem 8. The basic idea is that the coordinates of an optimal star center subdivide the plane into four quadrants. If ties are broken in the right way, the number of points in opposite quadrants is roughly the same. See Figure 1.

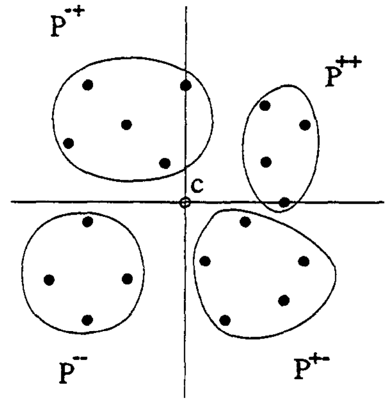


Figure 1: The four quadrants and their point sets

We assume without loss of generality that x_c and y_c are smallest possible, i. e., taken from the set of coordinates x_i and y_i. Let $n_x^0 := |P_x^0| := |\{x_i \mid x_i = x_c\}|$, and $n_y^0 := |P_y^0| := |\{y_i \mid y_i = y_c\}|$. By picking any subset of P_x^0 of size $\lceil \frac{n}{2} \rceil - n_x^-$ and joining it with P_x^-, we get a set $P_x^{-/0}$ of size $\lceil \frac{n}{2} \rceil$; the remaining $\lfloor \frac{n}{2} \rfloor$ points form the set $P_x^{0/+}$. Similarly, we get the partition into $P_y^{-/0}$ of size $\lceil \frac{n}{2} \rceil$, and $P_y^{0/+}$ of size $\lfloor \frac{n}{2} \rfloor$. Define the following *quadrant sets*: $P^{--} := P_x^{-/0} \cap P_y^{-/0}$, $P^{-+} := P_x^{-/0} \cap P_y^{0/+}$, $P^{+-} := P_x^{0/+} \cap P_y^{-/0}$, and $P^{++} := P_x^{0/+} \cap P_y^{0/+}$. The sets P^{--} and P^{++} are *opposite quadrant sets*, as are P^{-+} and P^{+-}. Two quadrant sets that are not opposite are called *adjacent*.

Now let $n^{--} := |P^{--}|$, etc. We get the following conditions:

LEMMA 2.2. *If n is even, then opposite quadrant sets contain the same number of points, i. e.,*

$$n^{--} = n^{++} \tag{2.5}$$

$$n^{-+} = n^{+-}. \tag{2.6}$$

If n is odd, the numbers of points must satisfy

$$n^{--} = n^{++} + 1 \tag{2.7}$$

$$n^{-+} = n^{+-}. \tag{2.8}$$

Proof. From the definition of the quadrant sets, it follows that

$$n^{--} + n^{-+} = n^{+-} + n^{++} \tag{2.9}$$

$$n^{--} + n^{+-} = n^{-+} + n^{++}. \tag{2.10}$$

if n is even. From this the claim follows easily.

For the odd case, the definition of the quadrant sets yields the conditions

(2.11) $n^{--} + n^{-+} = n^{+-} + n^{++} + 1$

(2.12) $n^{--} + n^{+-} = n^{-+} + n^{++} + 1.$

(2.13)

This implies the claimed conditions. ∎

When matching points in P^{--} with points in P^{++}, and points in P^{-+} with points in P^{+-}, we get the property

(2.14) $L_1(v_i, v_j) = L_1(v_i, c) + L_1(c, v_j)$

for any edge $(v - i, v_j)$ in the matching. Thus, it follows that

THEOREM 2.1. *If n is even, then for rectilinear distances in the plane, we have* $\max M(P) = \min S(P)$.

This can be used for obtaining an $O(n)$ algorithm for computing $\max M(P)$ for rectilinear distances in the plane: any matching connecting only vertices that lie in diagonally opposite quadrants is optimal. (This is also noted in the paper [9].)

2.2 Subtours

A 2-factor for a set of vertices is a set of edges that covers each vertex exactly twice. Since any tour is a 2-factor, a maximal length 2-factor is an upper bound for the length of a tour. Using triangle inequality, it is straightforward to see that twice the length of a star is an upper bound for the length of any 2-factor (i. e., a set of edges that covers each vertex exactly twice). Achieving tightness for this bound is the main stepping stone for our algorithm. Following the arguments of the preceding section, we prove the following three lemmas. We start with the easiest case:

LEMMA 2.3. *Suppose n is even and two of the quadrant sets are empty. Then there is a feasible tour of length* $2 \min S(P)$, *which is optimal.*

Proof. Any edge (v_i, v_j) between opposite quadrant sets satisfies property (2.14). If the number of points in two opposite quadrant sets is the same, we can get a tour by jumping back and forth while there are unvisited points in these quadrant sets. ∎

For a less trivial case, we can get two subtours meeting our upper bound:

LEMMA 2.4. *Suppose n is odd and $|P_x^0 \cup P_y^0| = 1$, or n is even. If $n^{++}, n^{-+} > 1$, then there is a tour $T_{--/++}$ of the points in $P^{--} \cup P^{++}$, and a tour $T_{-+/+-}$ of the points in $P^{-+} \cup P^{+-}$, such that $\ell(T_{--/++}) + \ell(T_{-+/+-}) = 2 \min S(P)$.*

Proof. If n is even, we can argue like in the proof of Lemma 2.3: There must be two subtours, one covering each pair of opposite quadrant sets.

If n is odd and there is only one point p_* in $P_x^0 \cup P_y^0$, the case reduces to n even, since $p_* = (c_x, c_y) \in P^{--}$, and p_* can be inserted into any tour of $P^{--} \setminus \{p_*\}$ and P^{++} while still guaranteeing (2.14) for any tour edge. ∎

In the following subsection, we will use this lemma to get the optimal tour value. If n is odd and there is more than one point sitting on the median axes, then we already have the optimal tour value:

LEMMA 2.5. *Suppose n is odd and $|P_x^0 \cup P_y^0| > 1$. Then there is a feasible tour of length* $2 \min S(P)$, *which is optimal.*

Proof. By conditions 2.1 and 2.3 and because x_c and y_c are smallest possible, we know that $P_x^0 \cap P_x^{-/0}$ and $P_y^0 \cap P_y^{-/0}$ each must contain a point. Therefore, consider two points $p_a \in P_x^0 \cap P_x^{-/0}$, $p_b \in P_x^0 \cap P_y^{-/0}$, with $a \neq b$. We distinguish the following cases – see Figure 2:

(a) $p_a, p_b \in P^{--}$:
By connecting p_a with a point in P^{-+}, and p_b with a point in P^{+-}, and otherwise jumping back and forth between opposite quadrant sets, we get a tour that satisfies (2.14) for any edge.

(b1) $p_a \notin P^{--}, p_b \in P^{--}$:
In this case, $p_a \in P^{-+}$. By changing the membership of p_b from P^{--} to P^{-+}, we get $|P^{--}| = |P^{++}|$ and $|P^{-+}| = |P^{+-}| + 1$. Then a tour for the modified P^{-+} and P^{+-} can be obtained like in case (a).

(b2) $p_a \in P^{--}, p_b \notin P^{--}$:
This is treated in the same way as case (b1).

(c) $p_a, p_b \notin P^{--}$:
In this case, $p_a \in P^{-+}$ and $p_b \in P^{+-}$. By changing the membership of p_a from P^{-+} to P^{++} and the membership of p_b from P^{+-} to P^{++}, we get $|P^{--}| + 1 = |P^{++}|$ and $|P^{-+}| = |P^{+-}|$, so we can get tours like in case (a). ∎

2.3 Computing optimal tours

Because of Lemma 2.4, we have a valid upper bound on the length of 2-factor, and a pair of subtours that

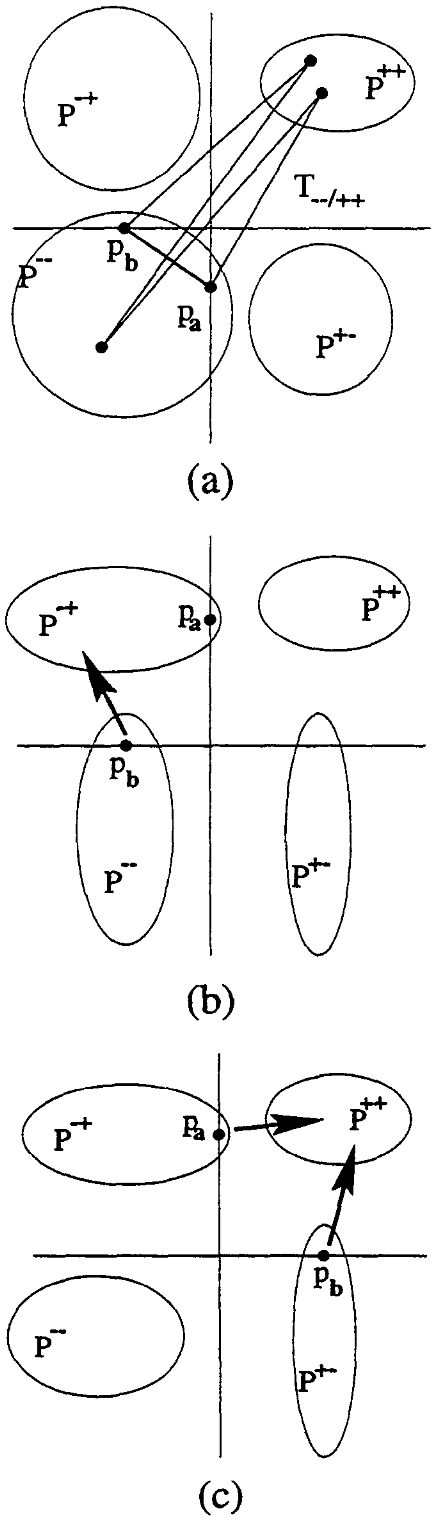


Figure 2: Getting optimal 2-factors

meet this bound, as long as all quadrant sets have a nontrivial number of points. Now we are left to argue how the upper bound has to be adjusted for connected tours, and how the adjusted bound can be met. Our discussion covers the case where there are two or more quadrant sets containing only one point.

We will argue that in any optimal tour, there must be a pair of edges $e_1 = (v_1, v_2)$ and $e_2 = (v_3, v_4)$, with $v_1 \in P^{--}$, $v_2 \in P^{-+}$, $v_3 \in P^{+-}$, $v_4 \in P^{++}$, or with $v_1 \in P^{--}$, $v_2 \in P^{+-}$, $v_3 \in P^{-+}$, $v_4 \in P^{++}$. See Figure 3. Such a pair of edges will be called a *quadrant matching.*

LEMMA 2.6. *Any tour of P contains a quadrant matching, or an edge connecting adjacent quadrant sets and an edge that stays within a third quadrant set.*

Proof. Any tour of P must contain at least two different edges e_1 and e_2 that connect adjacent quadrant sets, i. e., that connect $S_1 = (P^{--} \cup P^{++})$ to $S_2 = (P^{-+} \cup P^{+-}) = P \setminus S_1$. If $e_1 = (v_1, v_2)$ and $e_2 = (v_3, v_4)$ share a vertex $v \in S_i$, there must be other edges connecting $S_i \setminus \{v\}$ to $P \setminus (S_i \cup \{v\})$. Therefore, consider without loss of generality edges that do not share a vertex. This yields the following cases - see Figure 3:

(a) $(v_1, v_2), (v_3, v_4)$ form a quadrant matching:

In this case, there is nothing to prove.

(b) $v_1, v_3 \in P^{--}$, $v_2, v_4 \in P^{+-}$:

Since two edges adjacent to vertices in P^{+-} are already given, there can be at most $2n^{+-}-2 = 2n^{-+}-2$ edges between P^{+-} and P^{-+}, so there must be an edge connecting two vertices in P^{-+} (and we are done), or there must be two edges between P^{-+} and an adjacent quadrant set, reducing this case without loss of generality to case (c) below.

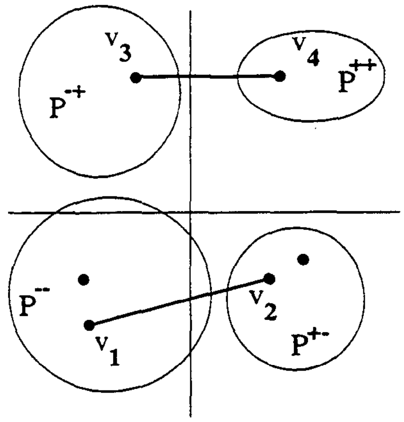
(c) $v_1, v_3 \in P^{--}$, $v_2 \in P^{+-}, v_4 \in P^{-+}$:

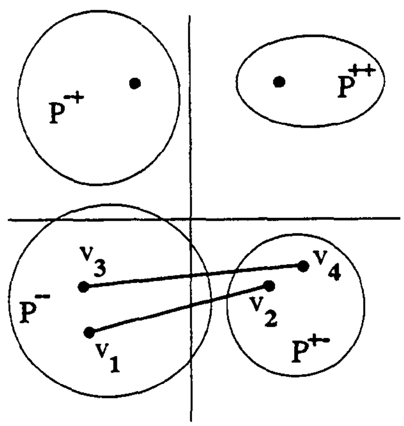
Since two edges adjacent to vertices in P^{--} are already given, there can be at most $2n^{--}-2 \geq 2n^{++}-1$ edges between P^{--} and P^{++}, so there must be an edge connecting two vertices in P^{++}, or an edge connecting P^{++} to an adjacent quadrant set. In either case, the claim follows. ∎

Now we can give an upper bound for an optimal tour with a given pair of edges.

LEMMA 2.7. *Let $e_1 = (p_1, p_2)$ be an edge connecting P^{--} and P^{+-}. Let $e_2 = (p_3, p_4)$ be an edge forming a quadrant matching with e_1, or an edge within a third quadrant P_3. Let $p_i = (x_i, y_i)$, and define $z_1 := \min\{(x_c - x_1), (x_c - x_2)\}$, $z_2 := \min\{(x_3 - x_c), (x_4 - x_c)\}$.*

Then any tour containing e_1 and e_2 has length at most $2 \min S(P) - 2z_1 - 2z_2$.





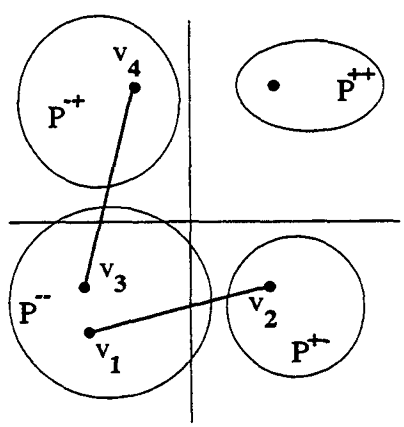


Figure 3: (a) A quadrant matching; (b+c) edges connecting adjacent quadrant sets

Proof. Since $L_1(p_1,p_2) = L_1(p_1,c)+L_1(c,p_2)-2z_1$, and $L_1(p_3,p_4) = L_1(p_3,c)+L_1(c,p_4)-2z_2$, the claim follows. ∎

By considering all pairs of edges, we get an adjusted upper bound of the tour length. For this purpose, let $Z_1 = 0$, if either $P^{--} = \emptyset$ or $P^{+-} = \emptyset$; otherwise, let $Z_1 = \min\{(x_c - x_i) \mid p_i \in P^{--} \cup P^{+-}\}$. Let $Z_2 = 0$, if either $P^{-+} = \emptyset$ or $P^{++} = \emptyset$; otherwise, let $Z_2 = \min\{(x_i - x_c) \mid p_i \in P^{-+} \cup P^{++}\}$. Similarly, let $Z_3 = 0$, if either $P^{--} = \emptyset$ or $P^{-+} = \emptyset$; otherwise, let $Z_3 = \min\{(y_c - y_i) \mid p_i \in P^{--} \cup P^{-+}\}$. Let $Z_4 = 0$, if either $P^{+-} = \emptyset$ or $P^{++} = \emptyset$; otherwise, let $Z_4 = \min\{(y_i - y_c) \mid p_i \in P^{+-} \cup P^{++}\}$. Finally, let $Z_* = \min\{Z_1 + Z_2, Z_3 + Z_4\}$.

Lemma 2.8. *Any tour of P has a length of at most $\min S(P) - 2Z_*$, and this bound can be computed in linear time.*

Proof. The claim follows immediately from Lemmas 2.6, 2.7, and the definition of Z_*. ∎

Now we can state the final claim:

Lemma 2.9. *For any point set P, there is a tour of length $\min S(P) - 2Z_*$.*

Proof. The claim is certainly true if either Lemma 2.3 or Lemma 2.5 provides the existence of a tour of length $\min S(P)$. Otherwise, consider a pair of vertices where the value Z_* is met. Without loss of generality, let this be for $p_1 \in P^{--}$ and $p_2 \in P^{++}$. Connect p_1 to any vertex in P^{+-}, and p_2 to any vertex in P^{-+}. Now it is easy to see that using only edges connecting opposite quadrant sets, we can get a tour. See Figure 4. ∎

Summarizing, we state:

Theorem 2.2. *The length of an optimal tour for the Maximum TSP under rectilinear distances in the plane can be computed in linear time.*

It is easy to conclude the following, which gives a first indication of the fundamental difference between rectilinear distances and Euclidean distances:

Corollary 2.1. *For any set of n points in the plane, there are $\Omega\left((\frac{n}{4}!)^4\right)$ many tours which are optimal for the max TSP under rectilinear distances. If the distances are Euclidean, there may only be one optimal tour.*

Proof. Any tour that can be constructed as in Lemma 2.9 is optimal, so we can choose an arbitrary permutation for each quadrant set. This yields the above lower

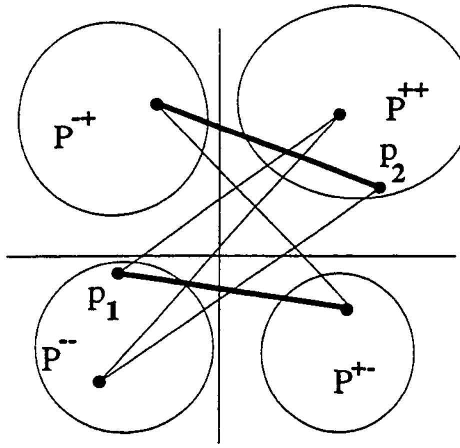


Figure 4: Getting an optimal tour

bound on the number of optimal tours. To see that there may only be one optimal tour for Euclidean distances, consider a set of $n = 2k+1$ points that are evenly distributed around a unit circle. ∎

Conversely, we see that any optimal tour can be constructed as described in Lemma 2.9.

We conclude this section by noting an observation by David Johnson:

PROPOSITION 2.1. *The above method, when augmented by a suitable linear transformation of the plane, works for all planar norms that have a symmetric quadrangle as the unit ball.*

3 An NP-Hardness Result

In this section, we establish the NP-hardness of the Maximum TSP under Euclidean distances in $\mathbb{R}^d$. The proof gives a reduction of the well-known problem *Hamilton Cycle in Grid Graphs*, which was shown to be NP-complete by Itai, Papadimitriou, and Swarcfiter [11]. A grid graph G is given by a finite set of vertices $V = \{v_1 v_2, \ldots, v_n\}$, with each vertex v_i represented by a grid point $(x_i, y_i) \in Z^2$. Without loss of generality, we may assume that G is connected, $(x_i, y_i) \in \{0, \ldots, n-1\}^2$, and that n is sufficiently large; for easier notation, we write $v_i = (x_i, y_i)$. Two vertices v_i and v_j in G are adjacent if and only if they are at distance 1, i.e., if $(x_i - x_j)^2 + (y_i - y_j)^2 = 1$. Note that any grid graph is bipartite: vertices v_i with $x_i + y_i$ even can only be adjacent to vertices v_j with $x_j + y_j$ odd, and vice versa. In the following, we will denote this partition by $V = V_e \dot{\cup} V_o$, where V_e is the set of "even" vertices, while V_o is the set of "odd" vertices.

The basic idea of the proof is to embed any grid graph G into the surface of a sphere in $\mathbb{R}^3$, such that edges in the grid graph correspond to longest distances within the point set. This can be achieved by representing the vertices in V_e by points that are relatively close to each other around a position (a, b, c) on the sphere, and the vertices in V_o are represented by points close to each other at a position on the sphere that is roughly opposite (i.e., antipodal) to (a, b, c); for simplicity of description by spherical coordinates, we will use positions that are close to the equator. Locally, the mapping of the two point sets onto the sphere is an approximation of the relative position of vertices in the grid graph. Since adjacent vertices in a grid graph have different parity, unit edges in the grid graph representation correspond to edges connecting points that are almost at opposite positions on the sphere, and vice versa.

In the following, the technical details are described. For simplicity, we use spherical coordinates and multiples of π. However, it will become clear from our discussion that we only require computations of bounded accuracy. It is straightforward to use only Cartesian coordinates that can be obtained by polynomial time approximation within the desired overall error bound of $O(n^{-4})$ for the length of an edge.

Represent each vertex v_i by a point $S(v_i)$ on the unit sphere, described by spherical coordinates (r, ϕ, θ), which translate into Cartesian coordinates by $x = r\cos\phi\cos\theta$, $y = r\sin\phi\cos\theta$, $z = r\sin\theta$. Note that, as in standard geographic coordinates, the "equator" of the sphere is given by $\theta = 0$; the angle θ describes the "latitude" of a point, while ϕ describes the "longitude". Since we will only consider points with $r = 1$, we will simply write (ϕ, θ) in spheric coordinates, but (x, y, z) in Cartesian coordinates. Let $\psi = \frac{2\pi}{n^3}$. Now any vertex $v_i \in V_e$ is represented by a point $S(v_i) = (x_i\psi, y_i\psi)$. Any vertex $v_i \in V_o$ is represented by a point $S(v_i) = (\pi + x_i\psi, -y_i\psi)$.

LEMMA 3.1. *There is a small constant $\varepsilon_n = O(n^{-4})$, which can be computed in polynomial time, such that for the three-dimensional Euclidean distance $L_2(S(v_i), S(v_j))$ between two points $S(v_i)$ and $S(v_j)$, the relation $L_2(S(v_i), S(v_j)) \geq 2 - \frac{\psi}{2} - \varepsilon_n$ holds if and only if v_i and v_j are adjacent in G.*

Proof. Since the diameter of the grid graph cannot exceed n, it is easy to see that $L_2(S(v_i), S(v_j)) \leq n\psi = O(n^{-2})$ whenever v_i and v_j have the same parity. Therefore, consider $v_i \in V_e$ and $v_j \in V_o$. Then

$$[L_2(S(v_i), S(v_j))]^2 =$$

$$
\begin{aligned}
&= [L_2(\ (\cos(x_i\psi)\cos(y_i\psi), \\
&\qquad \sin(x_i\psi)\cos(y_i\psi), \\
&\qquad\qquad \sin(y_i\psi)\ ), \\
&\quad (\cos(\pi + x_j\psi)\cos(-y_j\psi), \\
&\qquad \sin(\pi + x_j\psi)\cos(-y_j\psi), \\
&\qquad\qquad \sin(-y_j\psi)\ )\ )\ ]^2 \\
&= [\cos(x_i\psi)\cos(y_i\psi) + \cos(x_j\psi)\cos(y_j\psi)]^2 \\
&\quad + [\sin(x_i\psi)\cos(y_i\psi) + \sin(x_j\psi)\cos(y_j\psi)]^2 \\
&\quad + [\sin(y_i\psi) + \sin(y_j\psi)]^2 \\
&= \Big[\ \left(1 - \tfrac{(x_i\psi)^2}{2} + O\left((x_i\psi)^4\right)\right)\cdot \\
&\qquad \cdot\left(1 - \tfrac{(y_i\psi)^2}{2} + O\left((y_i\psi)^4\right)\right) \\
&\quad + \left(1 - \tfrac{(x_j\psi)^2}{2} + O\left((x_j\psi)^4\right)\right)\cdot \\
&\qquad \cdot\left(1 - \tfrac{(y_j\psi)^2}{2} + O\left((y_j\psi)^4\right)\right)\Big]^2 \\
&\quad + \Big[\ \left(x_i\psi - O\left((x_i\psi)^3\right)\right)\cdot \\
&\qquad \cdot\left(1 - \tfrac{(y_i\psi)^2}{2} + O\left((y_i\psi)^4\right)\right) \\
&\quad + \left(x_j\psi - O\left((x_j\psi)^3\right)\right)\cdot \\
&\qquad \cdot\left(1 - \tfrac{(y_j\psi)^2}{2} + O\left((y_j\psi)^4\right)\right)\Big]^2 \\
&\quad + \left[y_i\psi - O\left((y_i\psi)^3\right) + y_j\psi - O\left((y_j\psi)^3\right)\right]^2 \\
&= \Big[2 - \tfrac{(x_i\psi)^2}{2} - \tfrac{(y_i\psi)^2}{2} \\
&\qquad - \tfrac{(x_j\psi)^2}{2} - \tfrac{(y_j\psi)^2}{2} + O\left(n^{-8}\right)\Big]^2 \\
&\quad + \left[x_i\psi + x_j\psi + O\left(n^{-6}\right)\right]^2 \\
&\quad + \left[(y_i\psi)^2 + (y_j\psi)^2 + 2y_iy_j\psi^2 + O\left(n^{-8}\right)\right] \\
&= 4 - (x_i - x_j)^2\psi^2 - (y_i - y_j)^2\psi^2 + O\left(n^{-8}\right).
\end{aligned}
$$

Since v_i and v_j have different parity, we have $(x_i-x_j)^2\psi^2+(y_i-y_j)^2\psi^2 = \psi^2$, if v_i and v_j are adjacent in G, and $(x_i - x_j)^2\psi^2 + (y_i - y_j)^2\psi^2 \geq 5\psi^2$, if v_i and v_j are not adjacent in G, so the claim follows. ▮

From Lemma 3.1, it is straightforward to conclude that there is a tour of length at least $2n - n\frac{\psi}{2} - n\varepsilon_n$, if and only if the grid graph G is Hamiltonian.

We summarize:

THEOREM 3.1. *Maximum TSP under Euclidean distances in $\mathbb{R}^d$ is an NP-hard problem if $d \geq 3$.*

There are several implications of this result. It was pointed out by Joe Mitchell that there is a close connection to the Maximum TSP under polyhedral norms, where the number of facets k is not fixed: Since we only need to consider the $O\left(n^2\right)$ directions for connections between points, we can replace the Euclidean distances L_2 by a polyhedral norm with $O\left(n^2\right)$ facets.

COROLLARY 3.1. *The Maximum TSP under a polyhedral norm having a unit ball with k facets in $\mathbb{R}^d$ is an NP-hard problem, if $d \geq 3$ and k is part of the input.*

Another easy consequence concerns the *Maximum Scatter TSP*, which was first considered by Arkin, Chiang, Mitchell, Skiena, and Yang [1]. In this problem, the objective is to find a tour that maximizes the length of the shortest edge. Arkin et al. gave an NP-hardness proof for the general case and a 2-approximation that uses only triangle inequality. The complexity for geometric instances was left as an open problem. Using the above construction and Lemma 3.1, we get:

COROLLARY 3.2. *The Maximum Scatter TSP under Euclidean distances in $\mathbb{R}^d$ is an NP-hard problem if $d \geq 3$.*

Finally, it is straightforward with the above construction to show the following:

COROLLARY 3.3. *The Maximum TSP and the Maximum Scatter TSP on the $(d-1)$-dimensional surface of the d-dimensional unit sphere S^{d-1} under geodesic distances are NP-hard for $d \geq 3$.*

It was noted by Joe Mitchell that another corollary can be derived by using an approximation with $O\left(n^2\right)$ facets:

COROLLARY 3.4. *The Maximum TSP and the Maximum Scatter TSP on the $(d-1)$-dimensional surface of a d-dimensional convex polytope with an unbounded number of facets under geodesic distances are NP-hard for $d \geq 3$.*

4 Conclusion

We have given a linear time algorithm for the Maximum TSP under rectilinear distances in $\mathbb{R}^2$, improving the running time of $O(n^2 \log n)$ by Barvinok, Johnson, Tamir, Woeginger, and Woodroofe [6], and beating the lower bound of $\Omega(n \log n)$ under the decision tree model for verifying a Hamiltonian cycle. This result has no easy generalization to higher dimensions, since the partition into orthants by an optimal star center may not induce a "balanced" partition of the point set, such that we have subsets of equal size in opposite orthants.

Example. Consider P with $\frac{n-1}{4}$ points in each of the orthants $\{q = (x,y,z) \mid x > 0, y > 0, z > 0\}$, $\{q = (x,y,z) \mid x < 0, y < 0, z > 0\}$, $\{q = (x,y,z) \mid x < 0, y > 0, z < 0\}$, $\{q = (x,y,z) \mid x > 0, y < 0, z < 0\}$, plus the point $(0,0,0)$. It is easy to see that $(0,0,0)$ is the unique optimal star center. No connection of points in different orthants keeps the triangle inequality tight.

However, it should be possible to improve on the complexity by Barvinok et al. for L_1 distances in $\mathbb{R}^3$ by using some of our ideas. (Since the unit ball for

L_1 distances in $\mathbb{R}^3$ is an octahedron, the resulting complexity is $O(n^6 \log n)$.)

We have also shown that the Maximum TSP under Euclidean norm in $\mathbb{R}^d$ is NP-hard for any fixed $d \geq 3$. This shows that the complexity of an optimization problem is not just a consequence of its combinatorial structure or its geometry, but may be ruled by the structure of the particular distance function that is used. The result has similar implications for closely related problems.

The case $d = 2$ remains open; in the light of our results, it seems more likely that this problem is NP-hard, even though its counterpart with rectilinear distances turned out to be extremely simple. However, it is much harder to use strictly local arguments for geometric maximization problems, so a proof of NP-hardness may have to use a more involved construction.

CONJECTURE 4.1. *The Maximum TSP for Euclidean distances in the plane is an NP-hard problem.*

Acknowledgments

I would like to thank Henk Meijer for discussions of maximum matchings and minimum stars, which turned out to be of crucial importance for finding the linear time algorithm. Thank you to Peter Winkler and Bill Pulleyblank for helpful discussions, and to David Johnson for discussions and a proof of great sportsmanship. Also many thanks to Joe Mitchell for very helpful discussions and suggestions, to Estie Arkin, Arie Tamir, and Katja Wolf for useful suggestions, and to Volker Kaibel for lending me his ear. The basic construction for the hardness proof was inspired by my visit to Australia last year; many thanks to Mike Houle for the invitation!

References

[1] E. M. Arkin, Y.-J. Chiang, J. S. B. Mitchell, S. S. Skiena, and T.-C. Yang. On the Maximum Scatter TSP. *Proceedings of the 8th ACM-SIAM Symposium on Discrete Algorithms*, 1997, 211–220.

[2] S. Arora. Polynomial time approximation schemes for Euclidean TSP and other geometric problems. *Proceedings of the 37th Annual IEEE Symposium on the Foundations of Computer Science*, 1996, 2–12.

[3] S. Arora. Nearly linear time approximation schemes for Euclidean TSP and other geometric problems. *Proceedings of the 38th Annual IEEE Symposium on the Foundations of Computer Science*, 1997, 554–563.

[4] C. Bajaj. The algebraic degree of geometric optimization problems. *Discrete and Computational Geometry*, **3** (1988), 177–191.

[5] A. I. Barvinok. Two algorithmic results for the traveling salesman problem. *Mathematics of Operations Research*, **21** (1996), 65–84.

[6] A. I. Barvinok, D. S. Johnson, G. J. Woeginger, and R. Woodroofe. The maximum traveling salesman problem under polyhedral norms. *Proceedings of the 6th International Integer Programming and Combinatorial Optimization Conference (IPCO VI)*, Springer LNCS 1412, 1998, 195–201.

[7] M. Ben-Or. Lower bounds for algebraic computation trees. *Proceedings of the 15th ACM Symposium on the Theory of Computing*, 1983, 80–86.

[8] M. Blum, R. W. Floyd, V. R. Pratt, R. L. Rivest, and R. E. Tarjan. Linear time bounds for median computations. *Proceedings of the 4th ACM Symposium on the Theory of Computing*, 1972, 119–124.

[9] S. P. Fekete and H. Meijer. On maximum matchings and minimum stars. Manuscript, 1998.

[10] J. A. Fingerhut, S. Suri, and J. S. Turner. Designing least-cost nonblocking broadband networks. *Journal of Algorithms*, **24** (1997), 287–309.

[11] A. Itai, C. Papadimitriou, and J. L. Swarcfiter. Hamilton paths in grid graphs. *SIAM J. Computing*, **11** (1982), 676–686.

[12] E. L. Lawler, J. K. Lenstra, A. H. G. Rinnoy Kan, and D. B. Shmoys. *The Traveling Salesman Problem*, Wiley, Chichester, 1985.

[13] J. S. B. Mitchell. Guillotine subdivisions approximate polygonal subdivisions: Part II – A simple polynomial-time approximation scheme for geometric k-MST, TSP, and related problems. To appear, *SIAM Journal of Computing*. Available at `http://ams.sunysb.edu/~jsbm/jsbm.html`.

[14] C. H. Papadimitriou and M. Yannakakis. The traveling salesman problem with distances one and two. *Mathematics of Operations Research*, **18** (1993), 1–11.

[15] S. Suri. Problem # 5. Problem session of the *14th ACM Symposium on Computational Geometry*, 1998.

[16] A. Tamir and J. S. B. Mitchell. A maximum b-matching problem arising from median location models with applications to the roommates problem. *Mathematical Programming*, **80** No.2 (1998), 171–194.

[17] L. Trevisan. When Hamming meets Euclid: The approximability of geometric TSP and MST. *Proceedings of the 29th Annual ACM Symposium on the Theory of Computing*, 1997, 21–29.