

Worst-Case TSP Point Sets in All Dimensions are Asymptotically Equidistributed

Timothy Law Snyder¹
 Department of Computer Science
 Georgetown University
 Washington, DC 20057

J. Michael Steele²
 Department of Statistics
 The Wharton School
 University of Pennsylvania
 Philadelphia, PA 19104

EXTENDED ABSTRACT (June 10, 1994)

Given a set S of n points in the unit square $[0, 1]^d$, an optimal traveling salesman tour of S is a tour of S that is of minimum length. A *worst-case point set* for the Traveling Salesman Problem in the unit square is a point set $S^{(n)}$ whose optimal traveling salesman tour achieves the maximum possible length among all point sets $S \subset [0, 1]^d$, where $|S| = n$. An open problem is to determine the structure of $S^{(n)}$. We show that for any rectangular parallelepiped R contained in $[0, 1]^d$, the number of points in $S^{(n)} \cap R$ is asymptotic to n times the volume of R . Analogous results are proved for the minimum spanning tree, minimum-weight matching, and rectilinear Steiner minimum tree. These equidistribution theorems are the first results concerning the structure of worst-case point sets like $S^{(n)}$.

1. Introduction

In this paper we show that for many problems of Euclidean combinatorial optimization, the maximal value of the objective function is attained by point sets that are asymptotically equidistributed. To facilitate exposition, we focus at first on the traveling salesman problem (TSP) for a finite set S of points in the d -dimensional unit cube $[0, 1]^d$. Let $\tau(S)$ denote the set of tours that span S . The optimal TSP-cost of S is the value given by

$$\text{TSP}(S) = \min_{T \in \tau(S)} \sum_{e \in T} |e|, \quad (1.1)$$

where $|e|$ denotes the Euclidean length of the edge e .

¹ Research supported in part by Georgetown University 1991 Summer Research Award; Georgetown University 1992 Junior Faculty Research Fellowship; and Georgetown College John R. Kennedy, Jr. Faculty Research Fund.

² Research supported in part by the following grants: NSF DMS92-11634 and ARO DAAL03-91-G-0110.

For each dimension $d \geq 2$, there are constants c_d such that

$$\text{TSP}(S) \leq c_d |S|^{(d-1)/d}, \quad (1.2)$$

where $|S|$ denotes the cardinality of S . Considerable effort has been devoted to determining good bounds on c_d ; the earliest bounds are due to Few (1955), and the current records are held by Karloff (1989) and Goddyn (1990). Simply by considering the rectangular lattice one can see there are also constants $c'_d > 0$ such that for all $n \geq 2$,

$$\max_{\substack{S \subset [0,1]^d \\ |S|=n}} \text{TSP}(S) \geq c'_d n^{(d-1)/d}. \quad (1.3)$$

If we let $\rho_{\text{TSP}}(n) = \max\{\text{TSP}(S) : S \subset [0,1]^d, |S| = n\}$, then the usual considerations of continuity and compactness show that there are n -sets S for which $\text{TSP}(S) = \rho_{\text{TSP}}(n)$, and these are the *worst-case point sets* referred to in our title. We suppress ρ_{TSP} 's dependence on d to keep notation simple.

The main result obtained here is that worst-case point sets are asymptotically equidistributed in the sense made explicit in the following theorem.

Theorem 1. *If $\{S^{(n)} : 2 \leq n < \infty\}$ is a sequence of worst-case TSP point sets with $S^{(n)} \subset [0,1]^d$, $d \geq 2$, and $|S^{(n)}| = n$, then for any rectangular parallelepiped $R \subset [0,1]^d$, we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} |S^{(n)} \cap R| = \text{vol}_d(R). \quad (1.4)$$

While Theorem 1 is certainly intuitive, the proof we provide is not barehanded; it relies essentially on the result of Steele and Snyder (1989) that there exist constants $\beta_d > 0$ such that

$$\lim_{n \rightarrow \infty} \frac{\rho_{\text{TSP}}(n)}{n^{(d-1)/d}} = \beta_d. \quad (1.5)$$

The exact asymptotic result (1.5) was motivated by the classical result of Beardwood, Halton, and Hammersley (1959) for the case of random point sets, and it seems to provide just the refinement of bounds like (1.2) and (1.3) that is needed to obtain equidistribution limit theorems.

We note that a proof of Theorem 1 in dimension two using techniques different from the ones we use here is given in Snyder and Steele (1993). We also note that Theorem 1 has a close connection to some results and a conjecture of Supowit, Reingold, and Plaisted (1983). This connection will be explained more fully in Section 4, after we have developed some notation.

In the next section we prove Theorem 1; Section 3 deals with problems other than the TSP.

2. Proof of Theorem 1

For any fixed integer $m \geq 2$, we partition $[0,1]^d$ into m^d subcubes Q_i , where $1 \leq i \leq m^d$, each of side length $1/m$. By using straightforward approximation arguments one can check that to prove

Theorem 1, it suffices to consider equidistribution with respect to the Q_i . Specifically, it suffices to show that for each $m \geq 2$ and $1 \leq i \leq m^d$, we have

$$\lim_{n \rightarrow \infty} \frac{\text{TSP}(Q_i \cap S^{(n)})}{n} = \frac{1}{m^d}. \quad (2.1)$$

Our proof of (2.1) depends on the equality case of Hölder's inequality (Hardy, Littlewood, and Polya (1964), pp. 21–26). For $1 < p < \infty$ and $u_i \geq 0$, direct application of Hölder's inequality tells us that $\sum_{i=1}^k u_i = \sum_{i=1}^k u_i \cdot 1 \leq (\sum_{i=1}^k u_i^p)^{1/p} (\sum_{i=1}^k 1^{p/(p-1)})^{(p-1)/p} = (\sum_{i=1}^k u_i^p)^{1/p} k^{(p-1)/p}$. The fact that is more useful for us is that one can have equality in this bound if and only if $u_1 = u_2 = \dots = u_k$.

Let $s(n, i) = |Q_i \cap S^{(n)}|$; i.e., $s(n, i)$ is the number of points of a worst-case point set $S^{(n)}$ that appear in the the i th subcube. We first establish a limit result concerning the $s(n, i)$.

Lemma 1. *For all $m \geq 2$, we have*

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^{m^d} s(n, i)^{(d-1)/d}}{n^{(d-1)/d}} = m. \quad (2.2)$$

Proof.

This proof, included in the full version of the paper, uses (1.5); a heuristic TSP tour; the minimality of an optimal TSP tour; and Hölder's inequality to establish the necessary bounds. The proof is not included in this extended abstract; equation numberings of the full paper are retained. \square

We are now in position to prove Theorem 1. First we recall the subsequence convergence principle which says that if (a_k) is any sequence of real numbers with the property that for any integers $n_1 < n_2 < \dots < n_k < \dots$ there is a further subsequence $n'_1 < n'_2 < \dots < n'_k < \dots$ such that $a_{n'_k} \rightarrow \alpha$ as $k \rightarrow \infty$, then in fact one must have $a_k \rightarrow \alpha$ as $k \rightarrow \infty$.

Now let (n_k) be a given increasing sequence of integers. Since $0 \leq s(n, i)/n \leq 1$ for all n and i , we can find a subsequence (n'_k) of the (n_k) and m^d constants $0 \leq \alpha_i \leq 1$ such that for all $1 \leq i \leq m^d$ we have

$$\lim_{k \rightarrow \infty} s(n'_k, i)/n = \alpha_i. \quad (2.7)$$

Now, since $\sum_{i=1}^{m^d} s(n, i) = n$, we have from (2.7) that

$$\sum_{i=1}^{m^d} \alpha_i = 1. \quad (2.8)$$

Similarly, by (2.2) and (2.7), we have

$$\sum_{i=1}^{m^d} \alpha_i^{(d-1)/d} = m. \quad (2.9)$$

Now, equation (2.9) and Hölder's inequality applied with $u_i = \alpha_i^{(d-1)/d}$ and $p = d/(d-1)$ give us

$$m = \sum_{i=1}^{m^d} \alpha_i^{(d-1)/d} \leq \left(\sum_{i=1}^{m^d} \alpha_i \right)^{(d-1)/d} \left(\sum_{i=1}^{m^d} 1^d \right)^{1/d}. \quad (2.10)$$

But, by (2.8) we see that equality holds in (2.10), and thus $\alpha_1^{(d-1)/d} = \alpha_2^{(d-1)/d} = \dots = \alpha_{m^d}^{(d-1)/d}$, so applying (2.8) again, we see that $\alpha_i = 1/m^d$ for all i . By the subsequence convergence principle noted after Lemma 1, we therefore have for all $1 \leq i \leq m^d$ that $s(n, i)/n \rightarrow 1/m^d$ as $n \rightarrow \infty$, and the proof is complete. \square

4. Equidistribution in Related Problems

The method just used for the TSP can be applied to the minimum spanning tree, the minimum-length matching, and the rectilinear minimum Steiner tree. If $L = L(S)$ denotes the optimal cost associated with any of these, then we can define $\rho_L(n) = \sup_{\substack{S \subset [0,1]^d \\ |S|=n}} L(S)$ and let $S_L^{(n)}$ be such that $L(S_L^{(n)}) = \rho_L(n)$. To show that $S_L^{(n)}$ is asymptotically equidistributed boils down to checking that L satisfies two conditions:

1. $\rho_L(n) = \beta_{L,d} n^{(d-1)/d} + o(n^{(d-1)/d})$, where $\beta_{L,d} > 0$ is constant; and
2. $\rho_L(n) \leq m^{-1} \sum_{i=1}^{m^d} \rho_L(s_L(n, i)) + o(n^{(d-1)/d})$, where $s_L(n, i) = |S_L^{(n)} \cap Q_i|$.

Condition 1 has been proved for the minimum spanning tree, minimum-length matching, and rectilinear Steiner tree problems (cf., Steele and Snyder (1989) and Snyder (1992)), and Condition 2 can be verified for these problems by the method used in the proof of Lemma 1.

For example, if $L(S) = \text{MST}(S)$ denotes the total length of a minimum spanning tree of S , we first form a minimum spanning tree $\text{MST}(S_{\text{MST}}^{(n)} \cap Q_i)$ on each $S_{\text{MST}}^{(n)} \cap Q_i$. These trees can then be interconnected at total cost $o(n^{(d-1)/d})$ by adding $m^d - 1$ edges, each costing no more than c/m , where c is constant. This forms a heuristic tree on $S_{\text{MST}}^{(n)}$. Since the lengths $\text{MST}(S_{\text{MST}}^{(n)} \cap Q_i)$ are no greater than the worst-case (within-subcube) lengths $\rho_{\text{MST}}(s_{\text{MST}}(n, i))/m$, Condition 2 follows.

Checking these conditions for each of the problems yields the following.

Theorem 2. If $\{S_L^{(n)} : 1 \leq n < \infty\}$ is a sequence of worst-case point sets for the function L , where L is the minimum spanning tree, the minimum-length matching, or the rectilinear minimum Steiner tree, then, for any rectangular parallelepiped $R \subset [0, 1]^d$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} |S_L^{(n)} \cap R| = \text{Vol}_d(R). \quad (4.1)$$

5. Concluding Remarks

The asymptotic equidistribution of worst-case point sets for the problems we have considered offers some support to the conjecture of Supowit, Reingold, and Plaisted (1983) that worst-case point sets are approximated by lattices as $n \rightarrow \infty$. It is still a major open problem to resolve this conjecture.

Theorem 1 has a rather subtle relationship to some results of Supowit, Reingold, and Plaisted (1983). In addition to improving current bounds on the constants c_2 and c'_2 in (1.2) and (1.3), their study of the worst-case TSP in \mathbb{R}^2 decomposed $[0, 1]^2$ into m^2 labeled subsquares of side length $1/m$, then constructed a heuristic algorithm similar to that of Karp (1976). Supowit, Reingold, and Plaisted noted that the worst-case performance of their heuristic is attained on point sets that are equidistributed, and they used this observation to prove that the leading constant of the worst-case length of their heuristic tour is identical to the worst-case TSP constant β_2 in (1.5). This observation does not produce an equidistribution result for worst-case point sets, but it is suggestive of a result like Theorem 1. Still, a rigorous proof must take a much different path.

There are other open problems that are motivated by our results. For the Euclidean Steiner problem, the limit result for Condition 1 in Section 4 has yet to be established. We believe such a result holds, and it would imply that a worst-case point set for the Euclidean Steiner problem is asymptotically equidistributed. It is also likely that the Steiner points in the Euclidean and rectilinear cases are asymptotically equidistributed.

Another problem concerns the greedy matching. Though Condition 1 in Section 4 holds for this problem, the methods used here to verify Condition 2 fail since a greedy matching is not a matching of minimum length. Hence, this problem also remains open.

Bibliography

- Beardwood, J., Halton, J.H., and Hammersley, J. (1959). "The shortest path through many points," *Proceedings of the Cambridge Philosophical Society* 55, 299–327.

- Few, L. (1955). "The shortest path and the shortest road through n points in a region," *Mathematika* **2**, 141–144.
- Goddyn, L. (1990). "Quantizers and the worst-case Euclidean traveling salesman problem," *J. Comb. Theory, Series B* **50**, 65–81.
- Hardy, G.H., Littlewood, J.E., and Pólya, G. (1964). *Inequalities*, Cambridge University Press.
- Karloff, H.J. (1989). "How long can a Euclidean traveling salesman tour be?" *SIAM J. Disc. Math.* **2**, 91–99.
- Karp, R.M. (1976). "The probabilistic analysis of some combinatorial search algorithms," in *Algorithms and Complexity: New Directions and Recent Results*, J.F. Traub, ed., Academic Press, New York, 1–19.
- Snyder, T.L. (1992). "Worst-case minimal rectilinear Steiner trees in all dimensions," *Discrete & Comp. Geom.* **8**, 73–92.
- Snyder, T.L. and Steele, J.M. (1993). "Equidistribution of point sets for the traveling salesman and related problems," *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, Association for Computing Machinery, New York, NY, Society for Industrial and Applied Mathematics, Philadelphia, PA, 462–466.
- Steele, J.M. and Snyder, T.L. (1989). "Worst-case growth rates of some classical problems of combinatorial optimization," *SIAM J. Computing* **18**, 278–287.
- Supowit, K.J., Reingold, E.M., and Plaisted, D.A. (1983). "The traveling salesman problem and minimum matchings in the unit square," *SIAM J. Comput.* **12**, 144–156.