Equidistribution in All Dimensions of Worst-Case Point Sets for the TSP

Timothy Law Snyder[1]
Department of Computer Science
Georgetown University
Washington, DC 20057

J. Michael Steele[2]
Department of Statistics
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104

ABSTRACT

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 $S^{(n)}$.

Keywords: Equidistribution, worst-case, non-linear growth, traveling salesman, rectilinear Steiner tree, minimum spanning tree, minimum-weight matching.

---

[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.

AMS Subject Classifications: 68R10, 05C45, 90C35, 68U05

1. Introduction

In this note 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

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

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

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

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

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 $n \geq 2$,

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

If we let $\rho_{\mathrm{TSP}}(n) = \max\{\, \mathrm{TSP}(S) : S \subset [0,1]^d, |S| = n\,\}$, then the usual considerations of continuity and compactness show that there are n-sets S for which $\mathrm{TSP}(S) = \rho_{\mathrm{TSP}}(n)$ (cf. Moran (1984), P. 115); 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* $\{S^{(n)} : 2 \le n < \infty\}$ *is a sequence of worst-case* TSP *point sets with* $S^{(n)} \subset [0,1]^d$*,* $d \ge 2$*, and* $|S^{(n)}| = n$*, then for any rectangular parallelepiped* $R \subset [0,1]^d$*, we have*

$$\lim_{n\to\infty} \frac{1}{n}|S^{(n)} \cap R| = \mathrm{vol}_d(R). \tag{1.4}$$

While Theorem 1 is certainly intuitive, the proof we provide requires more than first principles; it relies essentially on the result of Steele and Snyder (1989) that there exist constants $\beta_d > 0$ such that

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

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 $m \ge 2$, we partition $[0,1]^d$ into m^d subcubes Q_i, where $1 \le i \le m^d$, each of side length $1/m$. For any rectangle R and any $\epsilon > 0$, there is an m and sets A and B such that $\cup_A Q_i \subset R \subset \cup_B Q_i$ and $\mathrm{vol}_d(\cup_{i\in B-A} Q_i) \le \epsilon\, \mathrm{vol}_d(R)$; hence to prove Theorem 1 it suffices to consider equidistribution with respect to the Q_i. Specifically, it suffices to show that for each $m \ge 2$ and $1 \le i \le m^d$, we have

$$\lim_{n\to\infty} \frac{|Q_i \cap S^{(n)}|}{n} = \frac{1}{m^d}. \tag{2.1}$$

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

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 ith subcube. We first establish a limit result concerning the $s(n,i)$ that measures their aggregate size in a way that works usefully with Hölder's inequality.

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

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

Proof.

First write (1.5) as

$$\rho_{\text{TSP}}(n) = \beta_d n^{(d-1)/d} + r(n), \text{ where } r(n) = o(n^{(d-1)/d}). \tag{2.3}$$

Let W denote a closed walk on $S^{(n)} = \{x_1, x_2, \ldots, x_n\}$; i.e, W is a sequence of edges $(x_{i_1}, x_{i_2}), (x_{i_2}, x_{i_3}), \ldots, (x_{i_{k-1}}, x_{i_k}), (x_{i_k}, x_{i_1})$ that visits each point of $S^{(n)}$ at least once and begins and ends at the same point. Even if W visits some points more than once and traverses some edges more than once, W is feasible for the traveling salesman problem on $S^{(n)}$, so $\text{TSP}(S^{(n)}) \leq \sum_{e \in W} |e|$.

We now construct a particular W on $S^{(n)}$ in the tradition of Karp (1976) and Supowit, Reingold, and Plaisted (1983). In each subcube Q_i for which $S^{(n)} \cap Q_i \neq \emptyset$, construct an optimal traveling salesman tour T_i of $S^{(n)} \cap Q_i$. This creates a set of at most m^d within-subsquare tours. We then select a point $x_i^\star$ from each T_i and let $T^\star$ be an optimal traveling salesman tour of $\{x_1^\star, x_2^\star, \ldots, x_{m^d}^\star\}$. The closed walk

W is then formed by visiting subsquares in the order specified by $T^\star$, visiting all members of subsquare Q_i by traversing T_i whenever $T^\star$ reaches $x_i^\star$.

To assess the length of W, we first note that $T^\star$ is a TSP tour of m^d points, so by (1.2), $\sum_{e \in T^\star} |e| \le c_d m^{d-1}$. This gives

$$\begin{aligned}
\rho_{\text{TSP}}(n) &= \text{TSP}(S^{(n)}) \\
&\le \sum_{e \in W} |e| \\
&= \sum_{i=1}^{m^d} \text{TSP}(S^{(n)} \cap Q_i) + \sum_{e \in T^\star} |e| \\
&\le \sum_{i=1}^{m^d} \text{TSP}(S^{(n)} \cap Q_i) + c_d m^{d-1}.
\end{aligned} \tag{2.4}$$

We now use (2.3) in (2.4) along with the fact that $\text{TSP}(S^{(n)} \cap Q_i)$ is at most $\rho_{\text{TSP}}(s(n,i))$ scaled by the subcube size $1/m$ to get

$$\begin{aligned}
\rho_{\text{TSP}}(n) &= \beta_d n^{(d-1)/d} + r(n) \\
&\le \sum_{i=1}^{m^d} \frac{\rho_{\text{TSP}}(s(n,i))}{m} + c_d m^{d-1} \\
&\le \frac{1}{m} \sum_{i=1}^{m^d} \beta_d s(n,i)^{(d-1)/d} + \frac{1}{m} \sum_{i=1}^{m^d} r(s(n,i)) + c_d m^{d-1},
\end{aligned} \tag{2.5}$$

where, for all $1 \le i \le m^d$, the value $|r(s(n,i))| \le \max_{k \le n} \{r(k)\} = o(n^{(d-1)/d})$. Since m is fixed, we cancel β_d in (2.5) to find

$$\sum_{i=1}^{m^d} s(n,i)^{(d-1)/d} \ge m n^{(d-1)/d} + h(n), \tag{2.6}$$

where $h(n) = o(n^{(d-1)/d})$. Dividing by $n^{(d-1)/d}$ and letting $n \to \infty$ thus proves half of the lemma. To obtain the other half, just apply Hölder's inequality with $p = d/(d-1)$ to $\sum_{i=1}^{m^d} s(n,i)^{(d-1)/d}$ and use $\sum_{i=1}^{m^d} s(n,i) = n$ to find that $\sum_{i=1}^{m^d} s(n,i)^{(d-1)/d} \le m n^{(d-1)/d}$.

□

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 $n_1 < n_2 < \cdots < n_k < \cdots$ there is a further subsequence $n'_1 < n'_2 < \cdots < n'_k < \cdots$ such that $a_{n'_k} \to \alpha$ as $k \to \infty$, then in fact one must have $a_k \to \alpha$ as $k \to \infty$. One easy way to see the validity of this principle is to note that if a_k does not converge to α, then there is some $\alpha' \neq \alpha$, $-\infty \leq \alpha' \leq \infty$, and some subsequence of (a_k) that converges to α'.

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

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

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

$$\sum_{i=1}^{m^d} \alpha_i = 1. \tag{2.8}$$

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

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

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

$$m = \sum_{i=1}^{m^d} \alpha_i^{(d-1)/d} \leq \Big(\sum_{i=1}^{m^d} \alpha_i\Big)^{(d-1)/d} \Big(\sum_{i=1}^{m^d} 1^d\Big)^{1/d}. \tag{2.10}$$

But, by (2.8) we see that equality holds in (2.10), and thus $\alpha_1^{(d-1)/d} = \alpha_2^{(d-1)/d} = \cdots = \alpha_{m^d}^{(d-1)/d}$, 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 $s(n,i)/n \to 1/m^d$ as $n \to \infty$, and the proof is complete.

□

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 $\mathrm{L} = \mathrm{L}(S)$ denotes the optimal cost associated with any of these, then we can define $\rho_{\mathrm{L}}(n) = \sup_{\substack{S \subset [0,1]^d \\ |S|=n}} \mathrm{L}(S)$ and let $S_{\mathrm{L}}^{(n)}$ be such that $\mathrm{L}(S_{\mathrm{L}}^{(n)}) = \rho_{\mathrm{L}}(n)$. To show that $S_{\mathrm{L}}^{(n)}$ is asymptotically equidistributed boils down to checking that L satisfies two conditions:

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

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 $\mathrm{L}(S) = \mathrm{MST}(S)$ denotes the total length of a minimum spanning tree of S, we first form a minimum spanning tree $\mathrm{MST}(S_{\mathrm{MST}}^{(n)} \cap Q_i)$ on each $S_{\mathrm{MST}}^{(n)} \cap Q_i$. 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 $S_{\mathrm{MST}}^{(n)}$. Since the lengths $\mathrm{MST}(S_{\mathrm{MST}}^{(n)} \cap Q_i)$ are no greater than the worst-case (within-subcube) lengths $\rho_{\mathrm{MST}}(s_{\mathrm{MST}}(n,i))/m$, Condition 2 follows.

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

Theorem 2. *If* $\{ S_{\mathrm{L}}^{(n)} : 1 \leq n < \infty \}$ *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* $R \subset [0,1]^d$,

$$\lim_{n \to \infty} \frac{1}{n} \left| S_{\mathrm{L}}^{(n)} \cap R \right| = \mathrm{Vol}_d(R). \tag{4.1}$$

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 $n \to \infty$. 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); we explain here how these results relate to ours. In addition to improving current bounds on the constants c_2 and c_2' in (1.2) and (1.3), their analysis of the worst-case TSP in $I\!R^2$ 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 the 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 of asymptotic equidistribution of a worst-case TSP point set required 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 we use to verify Condition 2 do not work since they require a minimality condition. Hence, since the greedy matching is not a minimum-length matching, showing equidistribution for a worst-case point set for the greedy matching problem remains an open problem.

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.

Moran, S. (1984). "On the length of optimal TSP circuits in sets of bounded diameter," *J. Comb. Theory*, Series B **37**, 113–141.

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.