AD-A256 112

(2)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)

2. REPORT DATE
June 1990

3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
The Voronio Diagram for The Euclidean Traveling Salesman Problem is Piecemeal Quartic and Hyperbolic

5. FUNDING NUMBERS

6. AUTHOR(S)
Cronin, T.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army CECOM Signals Warfare Directorate
Vint Hill Farms Station
Warrenton, VA 22186-5100

8. PERFORMING ORGANIZATION REPORT NUMBER
92-TRF-0043

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Statement A: Approved for public release, distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The Euclidean traveling salesman problem (ETSP) is a special case of the general traveling salesman problem (TSP). Given a set of cities and the associated costs between pairs of cities, the goal of the TSP is to find the optimal tour which visits every city exactly once, except for the start city, which is revisited at tour's end. Unlike the TSP which utilizes a general cost function to link cities, the ETSP employs the Euclidean distrance between cities as the metric, and equates optimality with shortest tour length. The objective of this research is to attempt to rigorously characterize the underlying geometry of ETSP tour construction, and subsequently to pursue an algorithm for an exact solution the problem for city databases of modest size.

14. SUBJECT TERMS
Artificial Intelligence, Data Fusion, Topology

15. NUMBER OF PAGES
29

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to **stay within the lines to meet optical scanning requirements.**

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	**PR** - Project
G - Grant	**TA** - Task
PE - Program Element	**WU**- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - DOD - Leave blank.
DOE - DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA - NASA - Leave blank.
NTIS - NTIS - Leave blank.

Block 13. Abstract. Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

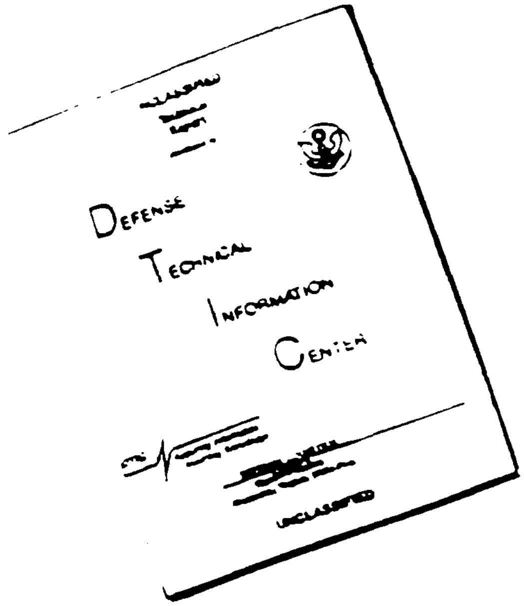
Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS only).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

The Voronoi Diagram for The Euclidean Traveling Salesman Problem is Piecemeal Quartic and Hyperbolic

T.M. Cronin
CECOM Center for Signals Warfare
Warrenton VA 22186-5100

Abstract. It is shown that the Voronoi diagram for the Euclidean traveling salesman problem is piecemeal quartic and hyperbolic. Previous attempts to leverage the traditional (linear) Voronoi diagram upon the problem have failed; in particular, counterexamples have demonstrated that the optimal tour need not traverse the Voronoi dual. In this paper, the shortest tour is treated as the union of a set of perturbations of the convex hull, with interior cities added incrementally, one at a time. A perturbation is defined to be the union of a new city with a subpath which connects two adjacent hull vertices to a set (possibly null) of previously entered interior cities. The length of a perturbation is therefore equal to the sum of two variable distances, minus the sum of a set of fixed distances. This length is called the elliptic length of the perturbation. Beginning with the convex hull, a single city is randomly added to the interior, and the hull is perturbed to capture the new city in optimal fashion. For a perturbation of a specific elliptic length, this quantity determines an ellipse symmetric about a hull segment, with foci at the segment endpoints. Any other hull perturbation of the same length defines another ellipse symmetric about some other hull segment. As the perturbation length is allowed to vary continuously from zero to infinity, a set of confocal ellipses is produced about each hull segment, and the intersection across all other such sets produces a set of quartics. For the special case in which hull segments share an endpoint (focus), the locus is a hyperbola. Each hull segment is bounded by those quartics for which the segment is the source of minimal perturbation length. The region of the plane thus bounded is called the quartic Voronoi cell for that segment. *There is a quartic cell corresponding to* each hull segment, and the union of all such cells forms the Voronoi diagram of the hull. Now, if a random city is injected into the hull, and the city is observed to lie in a specific Voronoi cell, we know that to produce the minimal tour, the city must be connected to the endpoints of the hull segment corresponding to the cell, and in turn the endpoints of the hull segment must be disconnected. If one maintains the proper canonical forms to alter the topology of the perturbation space when a new city is added, the technique may be extended to accommodate multiple interior cities. The quartic Voronoi diagram is shown to differ from the traditional Voronoi diagram in three distinct ways: it depicts shortest tour connectivity rather than point-to-point proximity; its cell boundaries are quartic and hyperbolic rather than linear; and the diagram is bounded by the convex hull rather than being unbounded (although this last constraint may be relaxed to add new cities outside the hull). A naively derived ceiling function demonstrates that an unsupervised perturbation approach is of exponential complexity, with a scaling factor as a function of the size of the hull. By resorting to an algorithm which exploits the canonical forms, it is shown that this bound may be diminished to $O[n^3]$. The algorithm is demonstrated for a database consisting of the forty-eight capitals of the contiguous United States. Open research issues include whether the technique may be extended to accommodate a hull which encloses an arbitrary number of cities, and whether the ceiling function may be further reduced.

Statement of the Problem. The Euclidean traveling salesman problem (ETSP) is a special case of the general traveling salesman problem (TSP). Given a set of cities and the associated costs between pairs of cities, the goal of the TSP is to find the optimal tour which visits every city exactly once, except for the start city, which is revisited at tour's end. Unlike the TSP which utilizes a general cost function to link cities, the ETSP employs the Euclidean distance between cities as the metric, and equates optimality with shortest tour length. The objective of this research is to attempt to rigorously characterize the underlying geometry of ETSP tour construction, and subsequently to pursue an algorithm for an exact solution to the problem for city databases of modest size.

Background. The traveling salesman problem has been an outstanding research issue for over a century, and has been approached computationally since the end of the second world war [L1]. It is important to differentiate between the TSP and TSP decision; the former requests a list of cities ordered as they appear in the optimal tour, whereas the latter seeks a yes or no answer to the question "is there a tour of cost k or less"? In 1972, it was proven that TSP decision is NP-complete [K2]; and in 1976 it was shown that ETSP decision with discretized distance is also NP-complete [P4, G1]. The ETSP with non-discretized distance is NP-hard in the strong sense [G1]. The failure of the ETSP to yield to known problem-solving strategies has caused the vast majority of researchers to abandon the search for an exact algorithm, and instead strive for fast approximation techniques. Many heuristic algorithms have been developed to date; they include: ***k-opt edge exchange*** [L3, J4], ***branch-and-bound*** [L4], ***simulated annealing*** [M1, J4], ***neural networks*** [F1, J4], ***genetic algorithms*** [B4]; ***and elastic bands*** [D4] A preeminent researcher in the field is of the opinion that for consistently high quality solutions on databases of very large scale, the Lin-Kernighan edge exchange algorithm has few competitors [J5]. Chapters 5-7 of reference [L1] provide valuable suggestions for evaluating the performance of some of the heuristic methods

An Historical Perspective of the Euclidean Traveling Salesman Problem and the Voronoi Diagram.

Operations research has been the historical forum for ETSP. There have been very few efforts possessing a computational geometry flavor. In the seventies, the issue was raised about whether a ETSP optimal tour must necessarily traverse adjacent cells of the Voronoi diagram [S2]. This conjecture has subsequently been answered in the negative. A counterexample for a degenerate case was discovered in 1983 [K1], and one for the general case was skillfully crafted five years later [D3]; the latter counterexample is portrayed at Figure 1. It will be shown in this paper that the fundamental reason for the difficulty in applying the traditional Voronoi diagram to ETSP is that the search space imposed by a perturbation of Euclidean distances is non-linear (in particular, it is quartic), whereas the traditional Voronoi cell possesses linear boundaries. It will also be shown that with modifications, the traditional Voronoi diagram may be extended to portray the optimal tour for an n-city problem, given the optimal tour for n-1 cities.

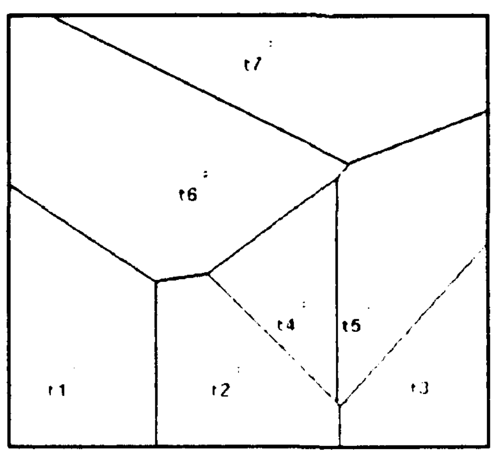


Figure 1. The traditional Voronoi diagram, computed for the Dillencourt dataset. In the optimal tour, t4 and t3 are connected, but their respective Voronoi cells are not, which counterindicates Shamos' conjecture that an optimal tour must traverse the Voronoi dual.

The Euclidean Shortest Tour as a Perturbation of the Convex Hull.

In 1957, Barachet proved that there exists an optimal tour which preserves the relative order of the points on the convex hull [B1]. This result implies that the shortest tour may be expressed as a hull deformation produced by an excursion into the interior, to capture points which do not lie on the hull (Figure 2). In 1977, a heuristic was developed to utilize the hull as an initial starting tour, and to attach interior cities based on a two step procedure [S4]. First, the sum of the distances from an interior city to the endpoints of an existing segment are computed, and the length of the segment subtracted; across all existing segments, the minimal such expression associates the city with a particular segment. The next step involves selecting a city to be inserted based on the maximal angle formed with its associated segment. The procedure iterates until a Hamiltonian cycle is formed utilizing all interior cities. Finally, an arbitrating function decides if the resultant cycle is sufficiently accurate. Analysis of the method has indicated that it is superior to some methods which do not utilize the hull [G3]. Nevertheless due to the fact that the approach is only approximate, the tour produced is generally suboptimal, and it is not well understood why the heuristic performs as it does.

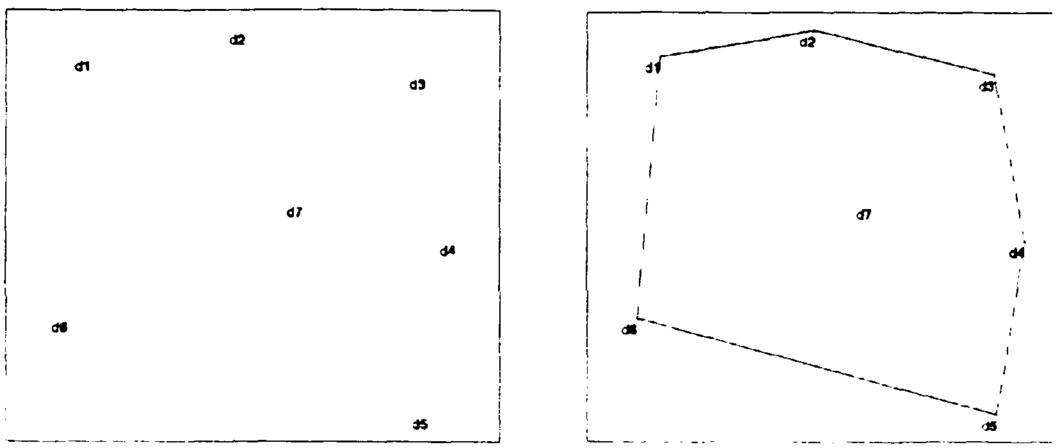


Figure 2. What is the shortest tour connecting cities d1-d7? We know that the tour must preserve the order of the cities lying on the the convex hull, so a natural way to proceed is to perturb the hull. The problem thus reduces to finding the optimal (shortest) way to attach city d7 to a pair of adjacent hull vertices.

Terminology and Notation.

In the following discussion, we shall call any excursion into the interior from two adjacent hull vertices a *perturbation* of the hull. It is important to note that a perturbation of the hull entails the corresponding loss of the segment which connects the two adjacent hull vertices. If a hull segment is unperturbed it is called a *null perturbation*. A *tour* is defined to be the union of a set of perturbations in counterclockwise order, as they appear about the hull. A convention will be adopted to represent certain perturbation concepts: a perturbation is denoted by the letter "π"; a tour is denoted by the letter "τ"; the length of perturbation π_i or tour τ_j is denoted respectively "len π_i" or "len τ_j"; the Euclidean distance between points p, q is denoted "d(p, q)"; the set of cities lying on the convex hull is denoted "H"; the convex hull itself is denoted "τ_H"; the number of cities in set S, also called the *order* of S, is denoted $|S|$. In concluding this introductory section, we formalize the definition of a perturbation, and prove three minor counting theorems.

Definition. *Hull Perturbation.* Given convex hull τ_H ordered with counterclockwise orientation, and the set I of interior cities. A *perturbation of the hull* π_k is an ordered subpath $\pi_k = h_k \cup I_j \cup h_{k+1}$; $I_j \subseteq I$. [Note that I_j may be the null set, in which case π_k is a null perturbation].

Theorem. When computing the Euclidean shortest tour, the number of perturbations of the convex hull cannot exceed the rank of the hull. Proof: By definition, a hull perturbation is an excursion into the interior of the hull which connects two adjacent hull vertices to a subset of the interior. Without regard to order, there are n ways to connect n adjacent hull vertices (the first to the second; the second to the third; ..., the nth to the first), producing a set of n hull segments. Each of these segments may be the source of a perturbation.

Theorem. If the size of the set of interior cities [I] exceeds the size of the hull [H], then the shortest Euclidean tour must contain a hull segment perturbation of order at least [I] - [H] + 1. Proof: from the pigeonhole principle, since there are more interior cities than hull segments, some hull segment must be assigned at least [I] - [H] + 1 interior cities.

Theorem. If the size of the hull [H] exceeds the size of the set of inner cities [I], then there must exist at least [H - I] hull segments which remain unperturbed when constructing the shortest Euclidean tour. Proof: again, from the pigeonhole principle, since there are fewer interior cities than hull segments, at least [H - I] hull segments must be null perturbations.

An Arbitrary Hull Enclosing a Single Interior City.

Since we know intuitively that the shortest Euclidean tour may be represented as a perturbation of the hull, let us proceed with the simplest case by introducing a single interior city into an arbitrary hull. It is natural to derive under what conditions a perturbation initiated from a given hull segment to the new city results in the shortest tour, versus one initiated from another segment. If len τ_H represents the sum of the lengths of the segments which comprise the convex hull, and p is an arbitrary point introduced into the hull interior, then to produce the shortest tour, one is interested in minimizing the expression len τ_H + $[d(p, h_i) + d(p, h_{i+1}) - d(h_i, h_{i+1})]$ $\forall h_i \in H$. The boundaries of equal hull perturbation are those for which len π_i = len π_j, for distinct elements h_i, $h_j \in H$. To formally characterize the boundaries of equal perturbation requires that the expressions for perturbations initiated from two distinct hull segments be set equal to each other, and the resulting equation solved.

The Elliptic Distance of a Point to Two Other Points.

During traveling salesman problem solving, the operation which decrements the length of a segment from the sum of the the distances from the segment endpoints to an arbitrary point is sufficiently fundamental to be given a special name, which we will call the *elliptic distance*.

Definition. The *elliptic distance* of a point p to two points q, r, denoted $d_e(p,q,r)$, is defined to be:

$$d_e(p,q,r) = d(p, q) + d(p, r) - d(q, r) \qquad [1]$$

Derivation of the Quartic Locus for the Single Interior City Problem.

Theorem. In general, the Voronoi diagram of the convex hull of the set of cities for the Euclidean traveling salesman problem has quartic edges.

Proof: Let p be an arbitrary point on the interior of convex hull τ_H, and let π_i, π_j be perturbations of two distinct hull segments such that len π_i = len π_j.

$$\text{len } \pi_i = [\, d(p, h_i) + d(p, h_{i+1}) - d(h_i, h_{i+1})] \text{ for some } h_i \in H, \text{ and}$$
$$\text{len } \pi_j = [\, d(p, h_j) + d(p, h_{j+1}) - d(h_j, h_{j+1})] \text{ for some } h_j \in H.$$

Let len π_i = len π_j = k_{ij}, which represents some specific elliptic length. Thus,

$$d_e[\,p, h_i, h_{i+1}] = d_e[\,p, h_j, h_{j+1}\,] = k_{ij} \qquad [2]$$

Equation [2] describes two ellipses, the first with major axis aligned with the hull segment having endpoints h_i, h_{i+1}, and the second aligned with the hull segment having endpoints h_j, h_{j+1}. The endpoints of the hull segments are the respective foci of the ellipses. The distances involving point p are variable, while the distances on the hull are constant. Let us represent the two ellipses as follows:

$$x^2/a^2 + y^2/b^2 = 1 \qquad [3]$$
$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \qquad [4]$$

Equation [3] denotes one of the two ellipses of interest after it has been rotated and translated to be in standard form about the origin. Equation [4] represents the second ellipse with the coefficients A, B, C, D, and E determined *in the coordinate system* of [3]. To characterize the locus of equal perturbation, we are required to simultaneously solve [3] and [4].

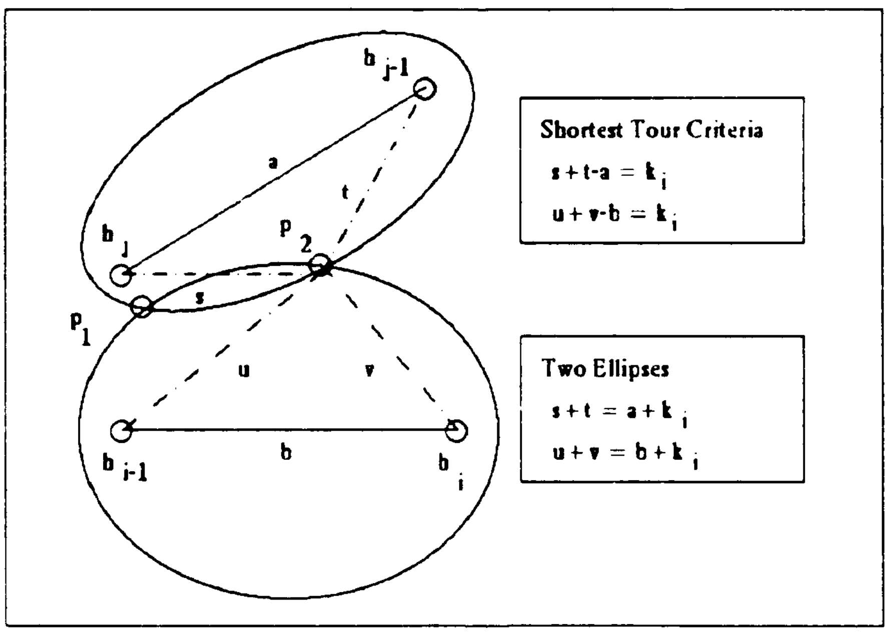


Figure 3. The locus of equal perturbation between hull segments is obtained by intersecting two ellipses.

From [3], we obtain: $y = +/- \, (b/a)\sqrt{a^2 - x^2}$ [5]

Substituting the positive root for y in [4] yields

$$Ax^2 + Bx(b/a)\sqrt{a^2 - x^2} + C(b/a)^2(a^2 - x^2) + Dx + E\sqrt{a^2 - x^2} + F = 0 \quad [6]$$

Factoring, and moving the radical to the right side of the equation produces

$$[A - (b^2/a^2)C]x^2 + Dx + b^2C + F = -(b/a)\sqrt{a^2 - x^2}\,(Bx + E) \quad [7]$$

Squaring both sides of [7] to clear the radical, and gathering coefficients with respective powers of x results in the equation:

$$\begin{aligned}
&[A^2 - (2b^2AC/a^2) + b^4C^2/a^4 + (b^2/a^2)B^2] && x^4 \\
&+ [2AD - (2b^2CD/a^2) + (2b^2BE/a^2)] && x^3 \\
&+ [2Ab^2C + 2AF - (2b^4C^2/a^2) - (2b^2CF/a^2) + b^2E^2/a^2 - b^2B^2 + D^2] && x^2 \\
&+ [2b^2CD + 2DF - 2b^2BE] && x \\
&+ b^4C^2 + 2b^2CF - b^2E^2 + F^2 = 0. && [8]
\end{aligned}$$

QED. Thus the locus of equal perturbation for inserting a random city into the hull is defined by a quartic equation, with coefficients expressed in terms of the parameters for two ellipses, where the ellipses are symmetric about segments formed by linking two cities.

A Graphic Depiction of a Simple Quartic Space.

Figure 4 illustrates an example of a quartic space imposed on a four city database. The segment containing z1 and z2 is fixed in the plane. The segment containing z3 and z4 is allowed to pivot about z4, with z3 being rotated counterclockwise ninety degrees through an angle θ, in increments of ten degrees. We are trying to find the locus such that a perturbation from the segment z1-z2 is equal to a perturbation from segment z4-z3. Initially, when z3 is on the x-axis, the locus is a horizontal line lying halfway between the two segments. For the sake of argument, when the angle $\theta = 0$, let the locus be the line $y = k$. At ten degrees, the locus lifts slightly from the horizontal and develops curvature. At about forty-five degrees, the locus develops a prominent maximum, and also mainfests two inflection points; it visually resembles the planar curve known as the Witch of Agnesi. Beyond forty-five degrees, the locus gradually loses its smooth maximum and develops a pronounced cusp; the fact that four distinct roots exist is now apparent. Finally, at ninety degrees of rotation, the locus becomes a diagonal line connecting z1 with the initial position of z3.

Note that a peculiar phenomenon has occurred. Although the segment containing z3 has been allowed to rotate ninety degrees (from the line $y = 0$ to the line $x = 0$), the corresponding locus has rotated only forty-five degrees (from the line $y = k$ to the line $y = -x + 2k$). Therefore, the angular range of the output is only half that of the input. It should also be noted that realistically, the behavior which produces the cusp does not occur, for when city z3 is rotated beyond a certain critical angle, it is absorbed by segment z1-z2, and the shortest tour becomes z1-z3-z2-z4-z1, rather than z1-z2-z3-z4-z1

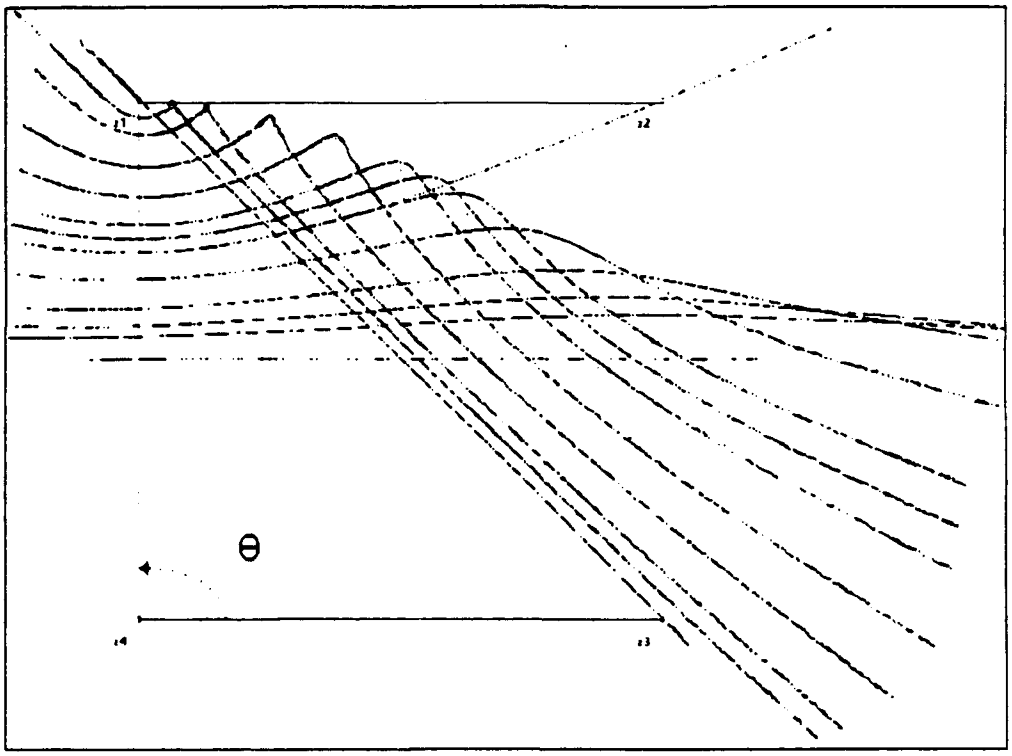


Figure 4. A plot of a quartic locus, using z4 as a pivot, while rotating z3 counterclockwise.

Two Hull Segments Sharing an Endpoint Produce a Hyperbolic Locus.

In this section, a corollary to the theorem is proven to show that in a special case which frequently occurs during traveling salesman problem solving, the locus of equal perturbation is hyperbolic rather than quartic. The simplest hypothesis to maintain when inserting another city into the hull is that the current set of perturbations will merely be extended by the new city, without radically altering the topology. As will become apparent below, the simple extension of a perturbation is arbitrated by an intrapath hyperbolic discriminator; the more complex operation of reasoning between perturbations requires a quartic discriminator.

Corollary. When two hull segments share a city, the Voronoi edge is a semi-hyperbola

Proof: Refer to Figure 5 below. Let the endpoints of the segments be respectively h_{i-1}, h_i and h_i, h_{i+1}. Let a be the length of the hull segment connecting h_i and h_{i+1}, and b the length of the segment connecting h_i and h_{i-1}. Let p be an arbitrary point on the locus. Let x be the distance from h_{i+1} to p, y the distance from h_i to p, and z the distance from h_{i-1} to p. We proceed to derive the locus:

$$
\begin{aligned}
d_e(h_i, h_{i-1}) &= d_e(h_i, h_{i+1}) \\
x + y - a &= y + z - b \\
x - a &= z - b \\
x - z &= a - b \qquad [9]
\end{aligned}
$$

QED. Equation [9] represents a semi-hyperbola passing through hull vertex h_i, with foci at h_{i-1} and h_{i+1}. It is bowed toward the longer of the two hull segments. In the case when the two hull segments are of equal length, the semi-hyperbola degenerates to a line.

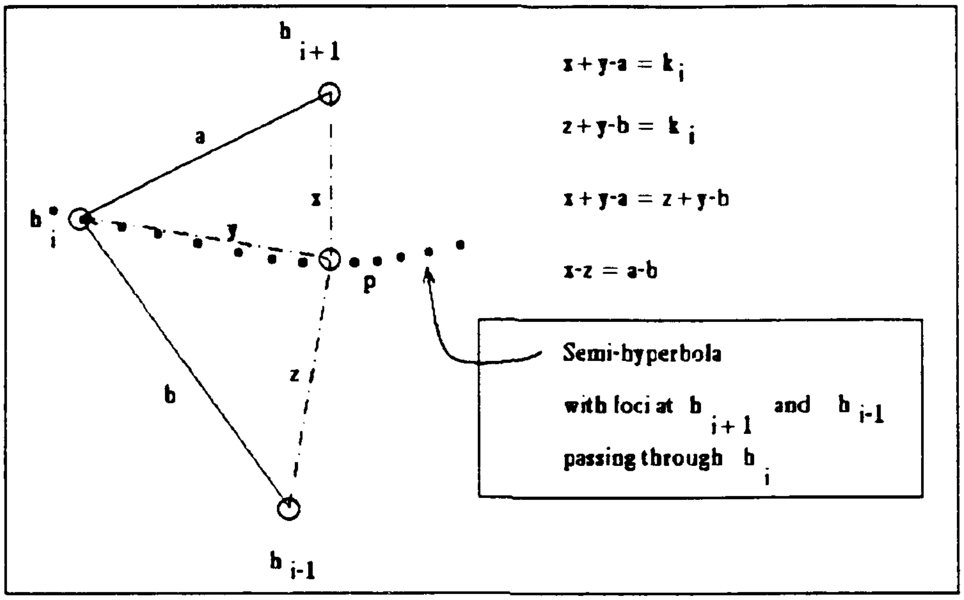


Figure 5. A hyperbolic locus results when two hull segments share an endpoint.

The Quartic Voronoi Diagram to Determine an Exact Solution to the Single Interior City Problem.

It is clear that to develop the delimiters of equal hull perturbation, we are required to develop a subset of the union of quartics and hyperbolas induced by the elliptic distance between pairs of hull segments. This subset is called the quartic Voronoi diagram of the hull, and is defined as follows:

Definition. Given convex hull τ_H, the quartic Voronoi diagram of the hull, VorQ (τ_H), is defined to be:

$$\mathrm{VorQ}(\tau_H) = \{\, x \mid d_e(x, h_i, h_{i+1}) = d_e(x, h_j, h_{j+1}) \text{ for some } h_j \in H,$$
$$\text{and } d_e(x, h_j, h_{j+1}) < d_e(x, h_k, h_{k+1}) \ \ \forall k \neq j \,\}.$$

For the seven city example introduced above, VorQ (τ_H) is displayed at Figure 6. City d7 is properly contained within the quartic Voronoi cell corresponding to the segment connecting d5 and d6, implying the segment must be perturbed to capture d7. Note that prior to introducing city d7, the existing optimal tour is the convex hull. The Voronoi diagram consists of some edges passing through the existing tour's vertices (the cities on the hull), and some which do not pass through any of the cities in the search space. In the former case, the edges are composed of hyperbolas, whereas in the latter case the edges are pieces of quartic curves. It is important to keep this concept in mind, because it will be revisited when the diagram is extended in general to accommodate a new city deposited into an arbitrary optimal tour space. The generalization will be seen to function in the following manner: if the existing optimal tour is extended simply by inserting a new city between two cities in the tour, the locus of equal perturbation which arbitrates the decision is a semi-hyperbola; otherwise, a more complex decision process must be invoked to reason across the perturbation space, and the locus is a quartic polynomial.

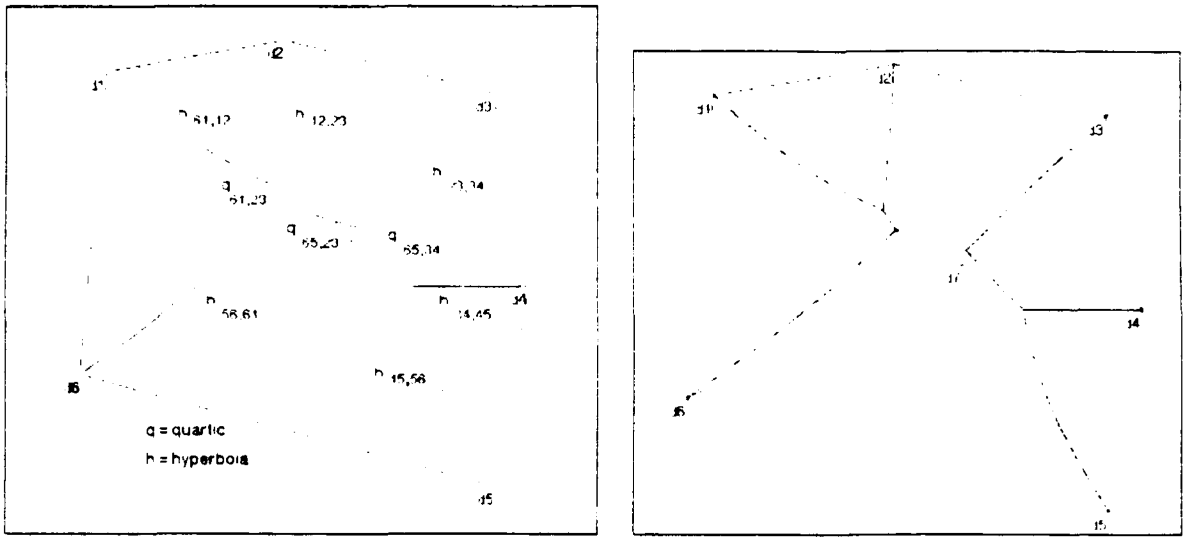


Figure 6 (computed quartics and hyperbolas). The quartic Voronoi diagram of the hull as a connectivity map. The curves passing through the hull vertices are hyperbolas; whereas the others are quartics. City d7 resides within the quartic Voronoi cell corresponding to the segment connecting d5 and d6, and therefore the optimal connection is as shown at the right.

A Two-City-in-a-Hull Example.

We will continue in this vein by commiting to computer memory the optimal perturbation for city d7, and introducing yet another city into the perturbed space. Figure 7 depicts an instance of a two-city-in-a-hull quartic Voronoi diagram. City d7 has been fixed, after having discovered its optimal perturbation (d6-d7-d5) in a previous step. A new city (p) is about to be introduced. At the left, the quartic edges demarcating optimal connectivity of p with d2, d3, d5, d6, and d7 are displayed, with the globally relevant pieces highlighted. Similar plots may be obtained for other pairs of vertices; for the complete interaction between segment d5-d6 and each of the other five hull segments, please refer to the Appendix (where d7 appears under the alias of d8). When computed across all relevant pairs, the quartic Voronoi diagram emerges (right). As an example, if p happens to fall in the quartic cell labeled with the descriptor "6p75", the optimal tour must insert p between cities d6 and d7, in the already established perturbation d6-d7-d5. However, if p happens to fall in the cell labeled "1p6;675", two distinct perturbations are required: the one involving p is issued from segment d1-d6; whereas the one involving d7 is issued from segment d6-d5.

Note that in the vicinity of city d7, a single perturbation from some hull segment is sufficient to ensure optimality. However, if city p happens to reside in one of the quartic cells beyond this region, it is necessary to perturb two hull segments to achieve optimality. Perhaps the most intriguing aspect of the Voronoi diagram as a connectivity map is that it partitions the plane into cells which indicate precisely how to maintain optimality when inserting an arbitrary city into the current tour. What this really means is that one can predict how to attach a new city to an optimal tour, without specifying the coordinates of the city ahead of time. The implications are profound, for if an efficient algorithm can be designed to construct (or perhaps merely reflect) the quartic diagram for arbitrarily large sets of cities, it follows that a dynamic programing approach is sufficient to solve the problem exactly. The fourth-order complexity inherent to the loci of tour constraints in large part explains why those approaches which subscribe to Dantzig's linear simplex have to date failed to solve the problem.

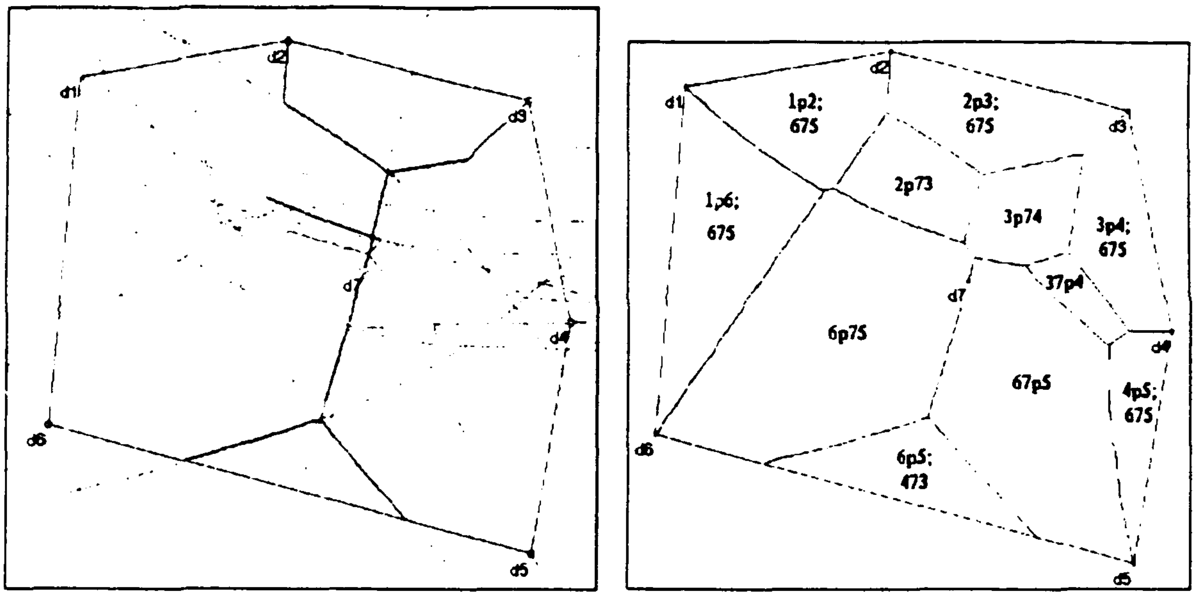


Figure 7 (computed quartics and hyperbolas). The quartic Voronoi diagram for two interior cities.

Nested Hull Traversal, Outside-in vs. Inside-out.

Since the theory is based on a perturbation of the convex hull, there is an obvious requirement to secure an algorithm which efficiently computes the hull. It is shown in reference [K3] that in the plane, the hull may be optimally computed in O [n * log h] time. With the traveling salesman problem, we interpret n as the total number of cities, and h as the number of cities on the convex hull. For our implementation, we will compute the entire nested hull decomposition, sometimes called the "onion" [E2]. The purpose of computing the onion is to gain control of the search space by attempting to insert the cities uniformly into the hull, to limit generation of "greedy" perturbations. A greedy perturbation occurs when by mere virtue of having probed sufficiently far into the hull, a perturbed hull segment continues to absorb cities which rightfully belong to another perturbation. Reference [C2] demonstrates that a planar nested hull structure may be constructed in O [n * log n] time. However, in the implementation described below, we will utilize an algorithm due to Eddy [E1], with time complexity O [n^2], but with average run time O [n].

Recall that by convention we order a hull with counterclockwise orientation, smallest ordinate first. If we label the outer hull with index 0, and label each inner hull with an ordinal number formed by incrementing the index by 1, it is seen that during processing, a city will be inserted based on a primary key equal to the ordinal number of its hull, and a secondary key equal to its relative counterclockwise position within its hull. The exception to this rule is to reject the insertion if it causes some unprocessed city to be bypassed.

A rather startling twist to the outside-in approach is based on the fact that the quartic loci of equal perturbation extend both inside and outside a hull. If we start with the innermost hull (the core of the onion), in theory we should be able to probe outward one hull at a time, maintaining optimality as we proceed. This technique, which we will call the *inside-out* approach, is in fact as valid as the other. An experiment detailed below demonstrates that both approaches are indeed capable of producing the optimal tour.

The Topology of Quartic Voronoi Space, in the Context of the Shortest Tour.

In this section we develop the canonical forms required to maintain quartic incremental optimality. Three primitive operations will be informally introduced, and then developed more rigorously.

The quartic Voronoi diagram partitions the plane into cells, the boundaries of which demarcate the locus of the shortest tour among various combinations of subtours. It is intuitively obvious that if a newly in oduced city lies within an arbitrarily small neighborhood of an existing optimal tour, the new tour can be formed simply by extending the old space to the new city. This topological structure, which we shall call ***hyperbolic extension space***, is computationally the simplest hypothesis to be entertained when introducing a new city. Hyperbolic extension space preserves an existing tour by extending an existing perturbation to encompass the new city.

A markedl, different topology is manifested when an extended perturbation interplays with another perturbation, which is located two hull segments backward (forward) to compel the issuing of a new perturbation at the preceding (subsequent) hull segment, which is called a ***shunt to the left*** (***shunt to the right***). This topology is called ***quartic shunt space***. It addresses the issue of maintaining optimality in a radial fashion; i.e., in a manner roughly orthogonal to the convex hull which defines the baseline tour. An intuitive way to describe this canonical form is that it acts as a monitor of flanking behavior on both sides of a perturbation, and cedes the flank to a neighboring hull segment when necessary to maintain optimality.

Because of the existence of the two distinct topologies, it is necessary to maintain separate computational hypotheses in parallel (Figure 8). An extension occurs if a new city lies in one of the extension cells in the lower portion of the diagram at the top. However, if the city lies within a Voronoi shunt cell as indicated at the top, a transition to quartic shunt space occurs when two existing perturbations are bridged (diagram at bottom). This not uncommon spatial phenomenon may radically alter the global shape of the tour, and must be hypothesized every time a new city is processed, to guarantee tour optimality.

The final topology deals with the issue of perturbation encroachment. There are instances when a perturbation probes suffciently far into the hull that for the sake of optimality it is necessary for it to claim cities from another perturbation. This spatial phenomenon produces the third topology which we call *quartic interchange space*. Quartic interchange space consists of those Voronoi cells which indicate that cities from one or more perturbations are to be exchanged into an extended or shunted perturbation. Quartic interchange is invoked after hyperbolic extension and quartic shunting, which are performed in parallel. It can be an operation of *quadratic complexity*, because an existing perturbation may be broken into sections and nullified by the operation, with separate sections being absorbed by separate perturbations. Quartic interchange is particularly relevant when using the outside-in nested hull approach, because at certain moments in time perturbations from across the hull begin to collide with those on the near side, and the interaction must be arbitrated to preserve optimality.

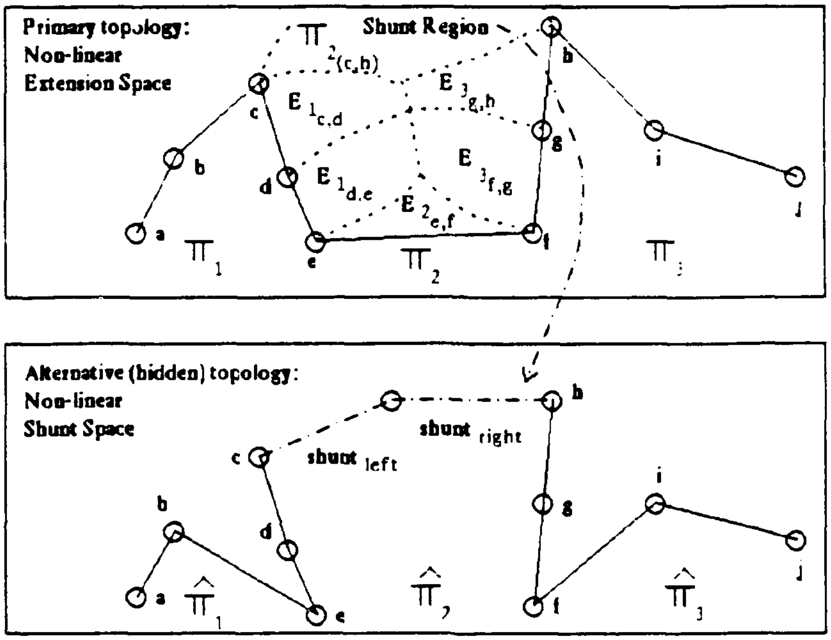


Figure 8. Maintaining two quartic topologies in parallel.

Assume that after k cities have been optimally connected to the convex hull, we would like to know under what conditions it is possible to simply extend the tour to a new city, vs. radically altering the tour by permitting the new city to link others which are currently non-adjacent. It is obvious that if the new city is within a small *spatial neighborhood of the existing tour*, optimality is preserved by simply inserting the city into the tour between two cities. The question of which two cities is governed by a set of hyperbolas which pass through the endpoints of the segments that connect the ordered list of cities defining the current optimal tour. Whether or not the new city is within a suitable neighborhood of the current tour is arbitrated by a set of quartic curves which discriminate if a shunting operation is in order.

Hyperbolic Extension Space.

It is a simple matter to connect a new city to an existing perturbation, if that is what is desired (it will be seen below in the section on quartic shunting that optimality is not always preserved by simply extending a perturbation). The city is connected to those two cities in the perturbation for which the elliptic distance is minimal. In other words, an existing perturbation should be extended to a new point if and only if the length of the perturbation plus the elliptic length of the optimal extension to the point is less than the corresponding sum for all other perturbations. For example, referring to Figure 9, if a new city is found to reside in quartic Voronoi cell "jk", it must be connected to cities j and k, while at the same time the segment joining j to k must be deleted. In this way, perturbation π_3 is extended to capture the new city.

There are cases when a specific perturbation requires reordering to maintain optimality: namely, when some city is nearer to the city to be inserted than either of the endpoints of the best segment found when minimizing the elliptic distance across all segments in the perturbation. However, the algorithm required to implement this operation is of linear time complexity, and does not detract from the performance of the general extension philosophy.

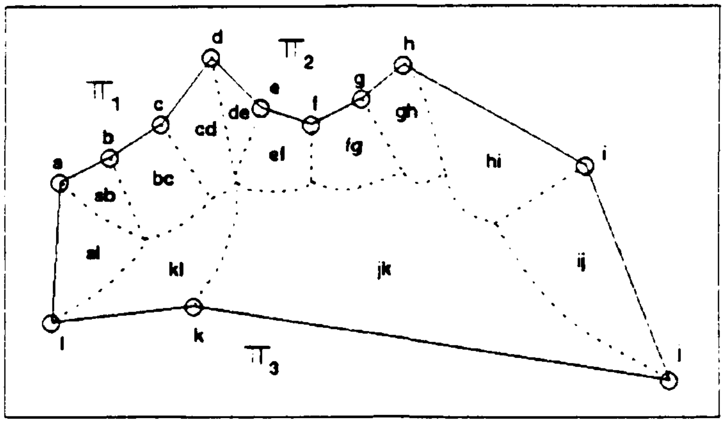


Figure 9 (estimated quartics and hyperbolas). Extending an existing perturbation is straightforward. When a new city enters the system, it is connected to the perturbation for which the elliptic distance to a segment is minimal, across all perturbations.

Quartic Shunt Space.

A shunt to the left is a bridging operation which connects a new city to the hull segment to the left of its extended perturbation, whereas a shunt to the right connects the city to the hull segment succeeding it. The left shunt is formed by connecting the city to its nearest neighbor two perturbations to the left, and then following the respective perturbation subpaths down to the hull vertices of the perturbation at the left. Any cities which become detached by this process must be reconnected to the perturbation space. The quartic shunt operator is a powerful tool, useful for merging two perturbations of the same parity into one with opposite parity, lying between the other two.

An example of a quartic shunt to the left is shown at Figure 10 (the data is a handcrafted approximation of a graphic depicted on p. 224, reference [PS]. At the left side of the figure, city c12 has just been introduced. Hyperbolic extension space calls for a perturbation of hull segment c5-c4, indicated by the dotted lines. However, quartic shunt space calls for a shunt to be formed between c12 and c16, which is the nearest neighbor two hull segments to the left of c12's extended perturbation. The endpoints of the shunt are followed down from c12 and c16 respectively to c5 and c6. City c13, which is left dangling by the shunt operation, is optimally reattached to the perturbation space by connecting it to segment c6-c7.

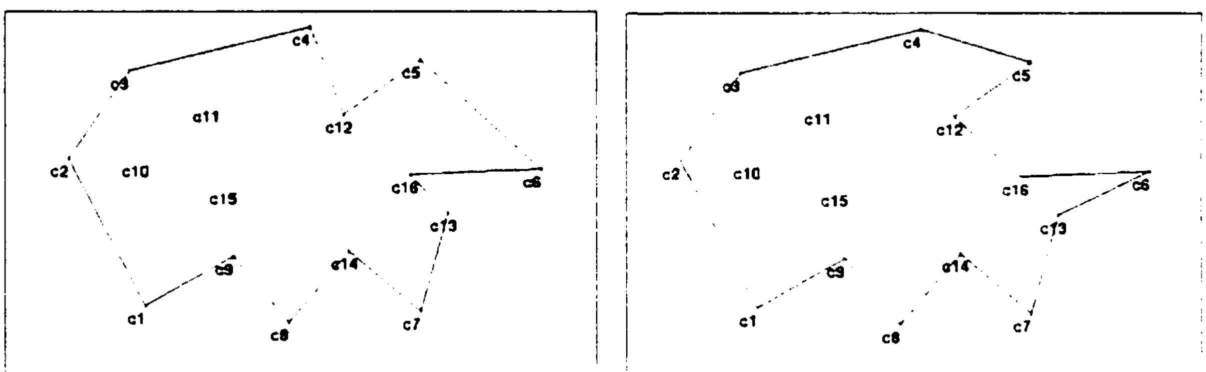


Figure 10. A quartic snunt to the left, using the Preparata and Shamos dataset.

Quartic Interchange Space.

Quartic interchange space dictates when a new city's perturbation, whether it be an extension or shunt, has encroached sufficiently far into the hull to encompass cities which earlier were optimally installed in some other perturbation. Every time a new city is processed, it must be hypothesized that the extending perturbation may now have encroached deep enough into the hull to begin influencing perturbations on the other side. While at some time in the past it may have been legitimate to have constructed a cross-hull perturbation to maintain optimality, it may now be time to partially or completely "undo" the perturbation by swapping some of its cities to the near side of the hull.

Quartic interchange is iterative. The currently extended perturbation is compared to all other perturbations in the space. If an exchange of cities is warranted, it is permitted to occur, and the revised perturbation space is subjected to interchange once again. This action is repeated until no improvement is obtained.

The General Voronoi Diagram for the Euclidean Traveling Salesman Problem.

Earlier, we proved that the Voronoi diagram of the convex hull has quartic edges, but possesses hyperbolic edges between adjoining hull segments. For the general case, this concept may be extended. The generalized Voronoi diagram partitions the plane into three types of cells: hyperbolic extension cells; quartic shunt cells; and quartic interchange cells. Before the k^{th} city is introduced, one computes the Voronoi diagram for the set of previously introduced k-1 cities. As in the convex hull case, computation must once again resort to an elliptic distance comparison, except now three different types of tour topologies must be hypothesized, rather than the single hypothesis entertained by introducing a single city into the hull. The space is once again quartic, because to obtain the boundaries of equal perturbation to the k^{th} city, two variable distances are added, and the sum of a set of fixed distances (the length of a specific hypothesized subpath) is subtracted. Rather than reasoning with perturbed segments on the hull, one must reason with tangible segments which are part of an existing tour, hypothesized segments which form shunts between perturbations, and hypothesized segments which form interchange links with other perturbations. To formalize, the quartic Voronoi diagram indicated by the optimal tour for k-1 cities, denoted $\mathrm{VorQ}(\tau_{k-1})$, is a function ψ of five arguments:

$$\mathrm{VorQ}(\tau_{k-1}) = \psi\,[\,\tau_{k-1}, \mathrm{Ext}_H(\tau_{k-1}), \mathrm{Shunt}_L(\tau_{k-1}), \mathrm{Shunt}_R(\tau_{k-1}), \mathrm{Inter}(\tau_{k-1})\,]$$

where:

τ_{k-1}	=	the optimal tour for k-1 cities;
$\mathrm{Ext}_H(\tau_{k-1})$	=	the hyperbolic extension space induced by τ_{k-1}
$\mathrm{Shunt}_L(\tau_{k-1})$	=	the left quartic shunt space induced by τ_{k-1}
$\mathrm{Shunt}_R(\tau_{k-1})$	=	the right quartic shunt space induced by τ_{k-1}
$\mathrm{Inter}(\tau_{k-1})$	=	the quartic interchange space induced by τ_{k-1}

The Principle of Quartic Incremental Optimality.

It is clear we are proceeding with a strategy akin to dynamic programming. Namely, when adding a new city to the interior, we attach the city to an existing optimal tour in a way indicated by the quartic Voronoi diagram computed just prior to the city's introduction. Formally, the shortest tour τ_k for k cities is a function Δ of two arguments:

$$\tau_k \quad = \quad \Delta\left((c_{k_x}, c_{k_y}), \mathrm{VorQ}(\tau_{k-1}) \right); \qquad k = [H] + 1, \ldots, n;$$

where:

τ_k	=	the optimal tour containing k cities;
(c_{k_x}, c_{k_y})	=	the coordinates of the k^{th} city to be introduced;
τ_{k-1}	=	the optimal tour containing k-1 cities;
$\mathrm{VorQ}(\tau_{k-1})$	=	the quartic Voronoi diagram prescribed by k-1 cities;
$[H]$	=	the order of the convex hull H;
n	=	the total number of cities to be processed.

Unique vs. Multiple Numbers of Distinct Optimal Tours as a Function of the Quartic Space.

Proper containment within a quartic Voronoi cell guarantees a unique tour. In nondegenerate cases, there can be no more than three unique tours because quartic Voronoi edges converge in groups of three just as linear Voronoi edges do If a newlv introduced city is situated at a Voronoi junction (a point where three quartics come together), three optimal tours exist; whereas if the city lies on a quartic but not on a junction, two optimal tours exist. In the degenerate case when cities are equispaced in the plane, there may be more tours than in the nondegenerate case. For example, a hull consisting of a regular polygon containing k sides produces k optimal tours if the introduced city lies at the center of the polygon, because the center is at the intersection of k Voronoi edges (degenerate hyperbolas). As a point of interest, it can be seen that if one allows the number of vertices of a regular polygon to approach infinity, so does the number of optimal tours connecting the hull to the center. However, in this case, the limiting form of the polygon is a circle, and a paradox arises, because the Euclidean distance between adjacent hull vertices approaches zero as the number of distinct optimal tours rises to infinity.

Future Work on A Proof of the Admissibility of Δ.

Mathematical induction will be used in an attempt to show that Δ is admissable. For the case of inserting a single city into the hull (i.e., k = 1), the shortest tour is trivially depicted by VorQ (τ_H), so the initial step of the inductive proof is satisfied. What remains to be shown is that the sequencing of the operations of hyperbolic extension, quartic shunting, and quartic interchange preserves optimality.

Summary of the Generalization of the Voronoi Diagram to ETSP.

During the first decade of the century, Voronoi's intention was to develop a mathematical structure which could be used to rapidly associate a query point with the nearest point contained within a known two-dimensional constellation of points [V2]. In decades of subsequent work, the dimensionality constraint has been relaxed, as well as the specification that both the query object and known objects be points [P5, E2]. This paper has focused on an exact solution to the Euclidean traveling salesman problem, and consequently has introduced a new distance metric, known as the elliptic distance, used to compute the distance of a floating query point from two fixed points. The limiting form of this metric, when intersected with that from another perturbed segment, induces a quartic structure on the Voronoi diagram. This non-linear search space permits a feasible testbed arena for the problem which the traditional Voronoi diagram cannot provide

Dillencourt's nondegenerate counterexample [D3] to Shamos' conjecture that the shortest tour must traverse the Voronoi dual is shown at the left of Figure 11. At the right is the quartic Voronoi diagram, which depicts connectivity of the three interior cities to the convex hull. All the curves indicated are hyperbolas, except the short one plotted between segment t1-t7 and segment t2-t3, which is a quartic locus. The shortest tour is t1-t2-t4-t3-t5-t7-t6-t1, but the linear Voronoi dual does not permit t3 and t4 to be connected. Note that in this instance, if one merely attaches each of the three interior cities to the hull segment indicated by the quartic Voronoi cell in which it resides, using the hyperbolic extension operator, the shortest tour is obtained. Of course, the actual computer run invokes the processes of quartic shunting and quartic interchange, but in this case the fourth-order operators fail to improve the tour produced by hyperbolic extension space.

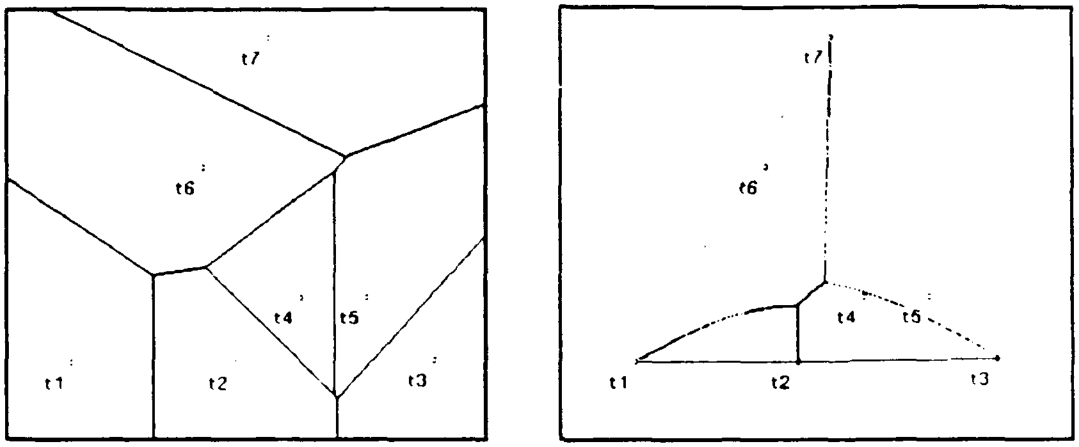


Figure 11 (computed edges). The Voronoi diagram for the Dillencourt data (left), and its one-city-in-a-hull *quartic Voronoi diagram* (right). This data is the first known nondegenerate counterexample to Shamos' conjecture that the shortest Euclidean tour must traverse adjacent Voronoi cells. In the optimal tour, t4 is connected to t3. It is apparent that t4 and t3 can be connected in the qurtic diagram, but not in the linear one.

The traditional Voronoi diagram is a proximity map, where at a glance it can be seen which object in a search space is nearest to a query point. The quartic Voronoi diagram is a connectivity map, which displays shortest tour connectivity information for the k^{th} city, as a function of a constellation of k-1 fixed cities. It has been shown that the process of intersecting an infinite set of confocal ellipses symmetric about an existing ETSP link with those about another link produces a quartic curve. The quartic curve, which in practice frequently reduces to a hyperbola because of the tendency of extension space to dominate during nested hull traversal, serves the same role as the perpendicular bisector does in the traditional Voronoi diagram. Thus, instead of being piecemeal linear, the extended Voronoi structure for ETSP is piecemeal quartic. The final discrepancy between the traditional Voronoi diagram and the quartic diagram deals with the issue of boundedness. Traditional Voronoi diagrams are unbounded; i.e., cells on the perimeter of the diagram are permitted to extend to infinity. However, for the Euclidean traveling salesman problem, the quartic Voronoi diagram is bounded by the convex hull of cities, so that no cell is unbounded. Nevertheless, this is not to say that the boundedness constraint cannot be loosened to incrementally add new cities exterior to the hull, which is the philosophy behind the inside-out nested hull approach (an example of this technique will be elaborated upon in an example appearing below, in which the innermost hull is used as a baseline optimal tour from which to add cities incrementally to the exterior). Table 1 summarizes the three distinctions between the traditional, linear Voronoi diagram, and its quartic counterpart designed for exact solution of the Euclidean traveling salesman problem.

Traditional Usage	⟶ Extension to ETSP
Proximity Map	Shortest-tour Connectivity Map
Cell Boundaries are Line Segments	Cell Boundaries are Piecemeal Hyperbolic and Quartic
Perimeter of Diagram is Unbounded	Perimeter of Diagram is Bounded by Convex Hull

Table 1. Extension of the Voronoi Diagram to the Euclidean Traveling Salesman Problem.

The Computational Complexity of the Hull Perturbation Approach.

a. The Time Complexity of Blind Search.

The convex hull of a set of cities serves as a control structure from which to initiate perturbations. In this section we naively derive an expression for the time complexity of the approach, if the tack is taken to blindly generate perturbations from hull segments in an arbitrary fashion. The derivation hinges upon making a substitution at an opportune moment when the binomial coefficients are manifested. It is hoped that future research will lend insight into techniques to improve the naive bound. The number of perturbations in the optimal tour cannot exceed the size of the hull, because a perturbation is defined to be an excursion into the interior from a hull segment, the number of which is equal to the number of hull vertices. Let H be the set of cities on the convex hull, and I be the set of cities lying on the interior of the hull. Let the rank of H be h, and the rank of I be i. If n is the total number of cities, then $n = h + i$. From each hull segment, the set of interior cities may be visited zero at a time, one at a time, two at a time, ... or i at a time. Thus the total number of computations required to find the shortest tour is:

$$h * {}_iC_1 + h * {}_iC_2 + \dots + h * {}_iC_i =$$

$$h * ({}_iC_1 + {}_iC_2 + \dots + {}_iC_i) =$$

$$h * (2^i) =$$

$$h * (2^{n-h}) =$$

$$\frac{h * 2^n}{2^h}$$

Some observations may be made about this naive bound. Note that a large hull is desirable, because the effect of the denominator is to diminish the 2^n term. Although the complexity is exponential, it is an order-of-magnitude improvement over a brute force approach, which is of factorial complexity.

b. The Time Complexity of Quartically-Controlled Search, using Nested Hull Traversal.

As a city is processed during nested hull traversal, there are three general phases of computing which must be performed in sequence. The first is a linear time operation to extend the current topology by minimizing the elliptic distance from all perturbations to the new city, which includes reordering a perturbation if necessary. The second phase involves two linear time operations, to construct the left and right quartic shunt topologies, after which they are compared with the extension topology to render the one with shortest tour length. Finally, the quartic interchange space is computed, which is a quadratic operation, because the left and right tour edges produced by the insertion of the new city act as windows to possibly absorb whole groups of cities from perturbations on the far side of the hull. Therefore, to process each new city, the worst case time complexity is quadratic. The sum of a set of quadratic expressions in k, where k ranges from 0 to n, is a closed form expression equal to $n * (n + 1) * (2n + 1) / 6$.

In summary, a computer implementation of the principle of quartic incremental optimality requires O[n * log n] preprocessing time to compute the nested hull decomposition, O[n] storage for intercity distances and optimal partial tours, and O[n^3] time complexity to maintain incremental optimality.

An Example: The Forty-eight Capital Problem.

The shortest tour connecting the forty-eight capitals of the contiguous United States remained an intriguing open problem until Shen Lin obtained an optimal solution in 1985 [A1, A2]. Each coordinate of the database represents the location of a Bell telephone office in the capital of a state. The principles of nested hull traversal and quartic incremental optimality were leveraged against this database. The nested hull structure for this data is exhibited at Figure 12. The implementation to derive the convex hull is based on an iterative enhancement made by the author to an algorithm developed in the seventies by W.F. Eddy [E1].

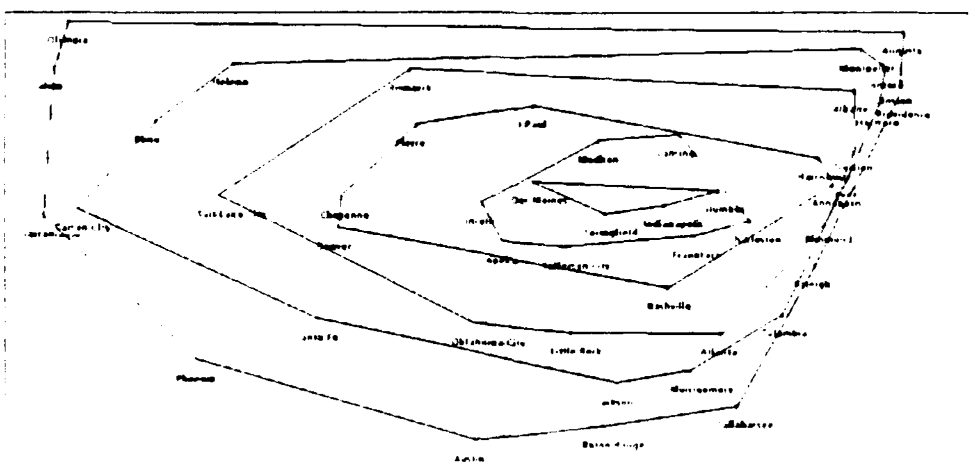


Figure 12. The Nested Hull Structure of the Forty-eight Capitals.

Working from the Outer Hull to the Inner.

First, the problem was attacked by starting with a baseline tour consisting of the outer convex hull, and probing inwards. Each nested hull is traversed in counterclockwise order to insert new cities, before the next inner hull is processed. A temporal history of incremental optimality is shown in Figure 13. The inserted city and the quartic function triggered are listed below each graphic. The interesting cases are those which are not mere hyperbolic extensions, but those which also involve quartic shunts and exchanges. The most dramatic quartic shunt occurs in row five, column seven, when the introduction of Springfield, Illinois produces a shunt to the right. Springfield is originally processed by the hyperbolic extension operator, which compels attachment to the perturbation which contains Lincoln, Nebraska. However, quartic shunt space produces a shorter tour by conjoining Springfield with a perturbation to the right containing Frankfort, Kentucky. Another interesting iteration occurs in row four, column six, when the introduction of Cheyenne, Wyoming into hyperbolic extension space causes the transposition of Bismarck, North Dakota with Pierre, South Dakota. Subsequently, quartic interchange causes Salt Lake City , Utah to be drawn out of its perturbation with Carson City, Nevada into the perturbation containing Cheyenne. The algorithm correctly terminates with Lin's optimal tour, shown in row six, column two.

Working from the Inner Hull to the Outer.

Next, the same data was processed by starting with the innermost nested hull and probing outward. Because the quartic Voronoi edges extend through the hull vertices both on the inside and the outside, the nested hull technique is theoretically valid in either direction. Figure 14 illustrates the optimal subtours produced by the algorithm when starting with the innermost hull and probing successively outward through the outer hulls. In this case, the innermost hull contains only four cities, so the original number of perturbations is four. There is an interesting tradeoff on time complexity when working with fewer perturbations. Again, note that the optimal tour is produced.

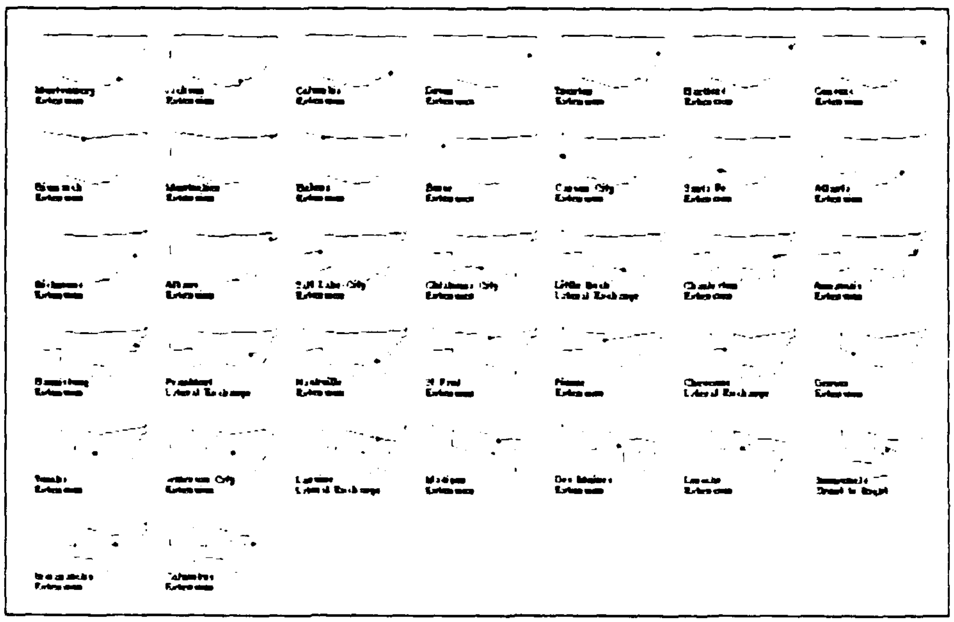


Figure 13. Working inwards from the outer hull, employing quartic nested hull traversal.

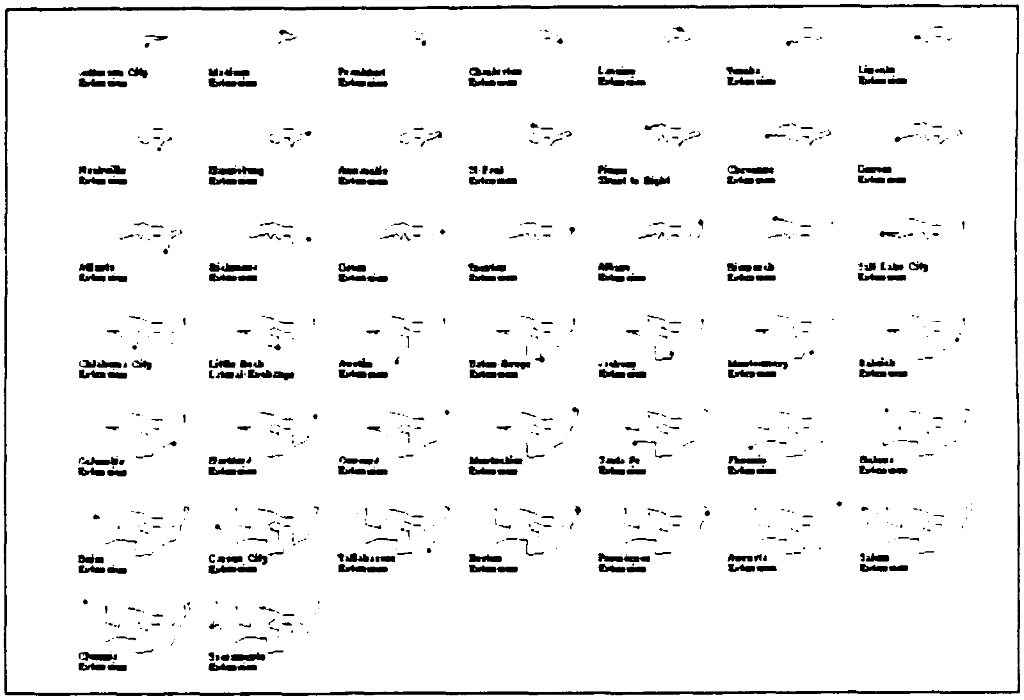


Figure 14. Working outwards from the inner hull, employing quartic nested hull traversal.

Summary.

The chief result of the research to date is a proof that the underlying search space (the Voronoi diagram) for the Euclidean traveling salesman problem is non-linear; specifically, the space is quartic when reasoning across subtours, and hyperbolic when reasoning within subtours. These facts become apparent when one realizes that reasoning about shortest tours is a process which inherently involves the intersection of a pair of ellipses, the foci of which are defined by pairs of cities. Ellipse intersection is an operation which in the worst case produces a fourth-order equation (quartic). In the special case in which two ellipses share a focus, the locus is a semi-hyperbola. The discovery of the non-linear search space has prompted the author to devise an algorithm which utilizes three operators to constrain search: hyperbolic extension; quartic shunting; and quartic interchange. To limit the generation of greedy perturbations, cities are gradually inserted in an incremental fashion, according to their position within the nested hull structure of the city database. The new knowledge about the non-linear search space has resulted in an O[n^3] solution to optimality of a forty-eight city problem. The solution is obtained both by beginning with the outer convex hull and probing inward, or by starting with the innermost hull and probing outwards.

Future Directions of the Research.

It is desirable to pursue a rigorous proof of the theory of quartic incremental optimality; the proof will proceed by induction. Also, to facilitate further empirical analysis, the theory as it currently stands will continue to be developed and leveraged against several large databases of cities for which the optimal tour is known, in an attempt to find examples which counterindicate the algorithm. Short-term plans include runs against a 127-city database [R1], and a 532-city benchmark for which the optimal solution has been developed [P1]. Subsequently, the runtime for the experimental data will be plotted as a function of the number of cities, to determine if the algorithmic ceiling function is of cubic order as predicted by the analysis.

Acknowledgments

The research has greatly benefited from discussions with Gerald Andersen, Richard Antony, Jacob Barhen, Chris Bogart, Paul Broome, Jagdish Chandra, Douglas Chubb, Michael Dillencourt, Martin Groetschel, Andrew Harrell, Robert Hein, William Jackson, Shen Lin, Sanjoy Mitter, Al Pollard, John Pfaltz, Carl Russell, Paul Tseng, Franz-Erich Volter, and David Willshaw. I am once again indebted to Richard Antony for his customarily excellent job of technical editing and commentary. A special note of gratitude to David S. Johnson of AT&T Bell Laboratories for incisive, constructive criticism of the principle of incremental optimality, and for providing the coordinate data for the forty-eight capital problem. I am grateful to Geoffrey Fox and Robert Bixby for an invitation to a very productive workshop entitled "Computational Aspects of the Traveling Salesman Problem", held at Rice University, 22-24 April 1990. Herbert Edelsbrunner was a source of valuable insights into issues arising when computing the convex hull. Thanks to Gerhard Reinelt for discussions on the Voronoi diagram, and for establishing a library of traveling salesman problem city databases, which will allow researchers worldwide to share a common testbed of benchmark problems for which the optimal solutions are either known or bounded.

APPENDIX

The appendix consists of a series of five computer plots which graphically portray the quartic (fourth-order polynomial) loci which exist naturally when hypothesizing a solution to the two-city-in-a-hull Euclidean traveling salesman problem described in the main body of text. These loci were discovered empirically by the author during the summer of 1989; it was only later after several months of research that a proof was obtained to demonstrate algebraically that the loci are actually comprised of quartic and hyperbolic curves. It may be of interest to some readers to know how the plots were obtained. An algorithm was designed to capture the knowledge about the possible ways (permutations) to connect city d8 and one other arbitrary city to the convex hull (d1-d2-d3-d4-d5-d6). In the eight city example, there are six possible topologies to compare between any two hull segments: the two ways to attach d8 and the arbitrary city to each of the two hull segments (which yields a subtotal of four), and the two ways to attach one city to one segment and the second to the other segment. In general, this means that there are fifteen quartic loci (the combination of six topologies taken two at a time) among which to arbitrate when hypothesizing a shortest tour. The algorithm was encoded in Lisp and run on an artificial intelligence computer workstation. A set of experiments were conducted as follows: the computer mouse was moved about its pad on the desk, which caused the cursor to move about the monitor screen displaying the constellation of cities. If the length of a specific arrangement of cities was within one unit of that of another arrangement, a black dot was plotted to the screen at the position of the cursor. This action provided positive feedback to the author, who dynamically readjusted the position of the mouse to obtain "more black dots" in a continuous fashion, until an entire quartic curve manifested itself. When a point in time was reached in which it became obvious that no more loci were forthcoming, the session was terminated, and another pair of segments was selected for experimentation. In the set of graphics selected for exhibit here, one of the pair is always segment d6-d5. Also note that what in the text was referred to as city "d7" is here called city "d8". It should also be pointed out that in exhibits A-3, A-4, and A-5, the quartic plots are superimposed over the solution to the one-city-in-a-hull problem discussed in the text.

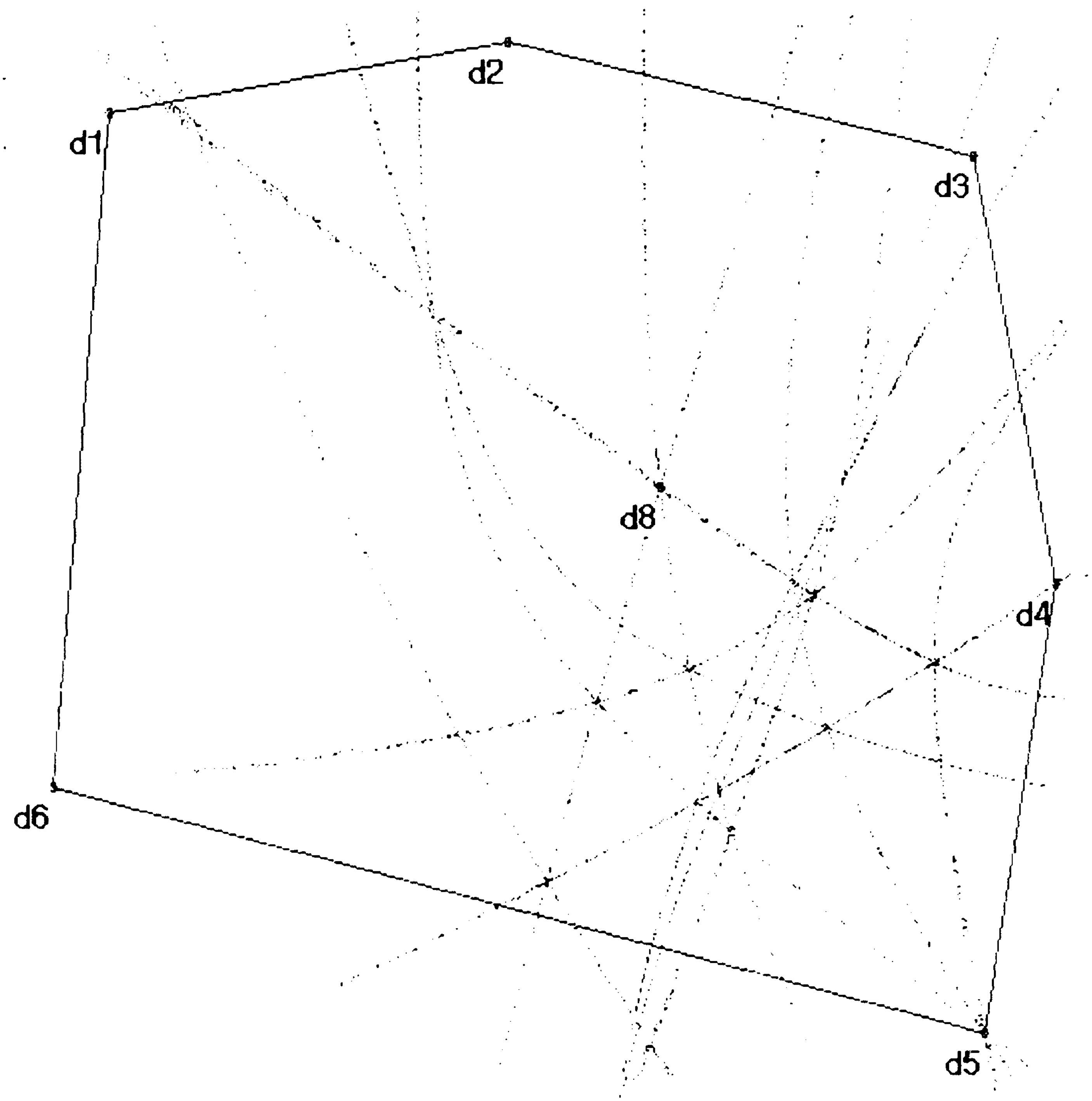


Exhibit A-1. The quartic interplay between segment d6-d5 and segment d5-d4, when attaching d8 and an arbitrary city to the convex hull to produce the shortest Euclidean tour.

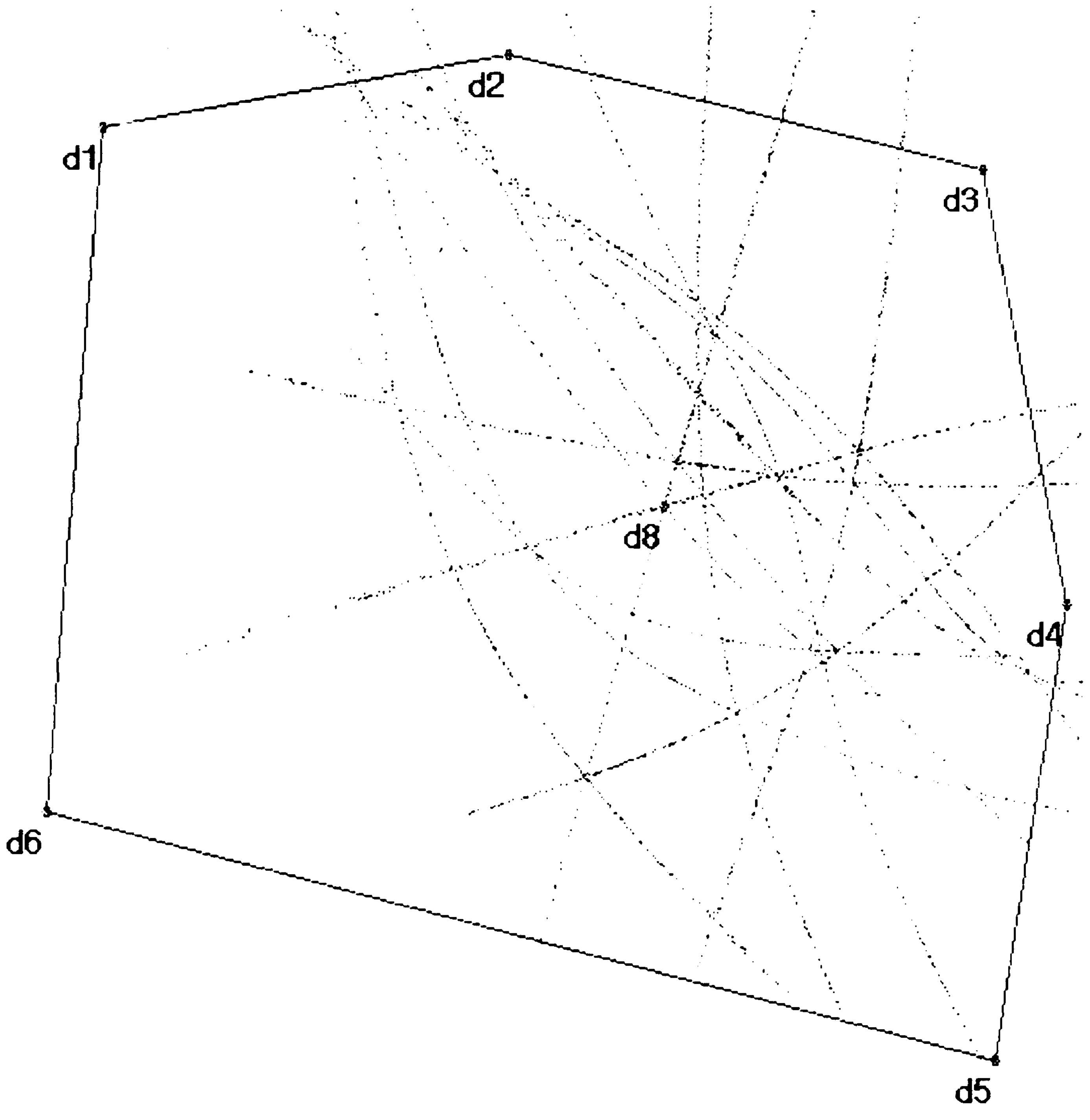


Exhibit A-2. The quartic interplay between segment d6-d5 and segment d4-d3, when attaching d8 and an arbitrary city to the convex hull to produce the shortest Euclidean tour.

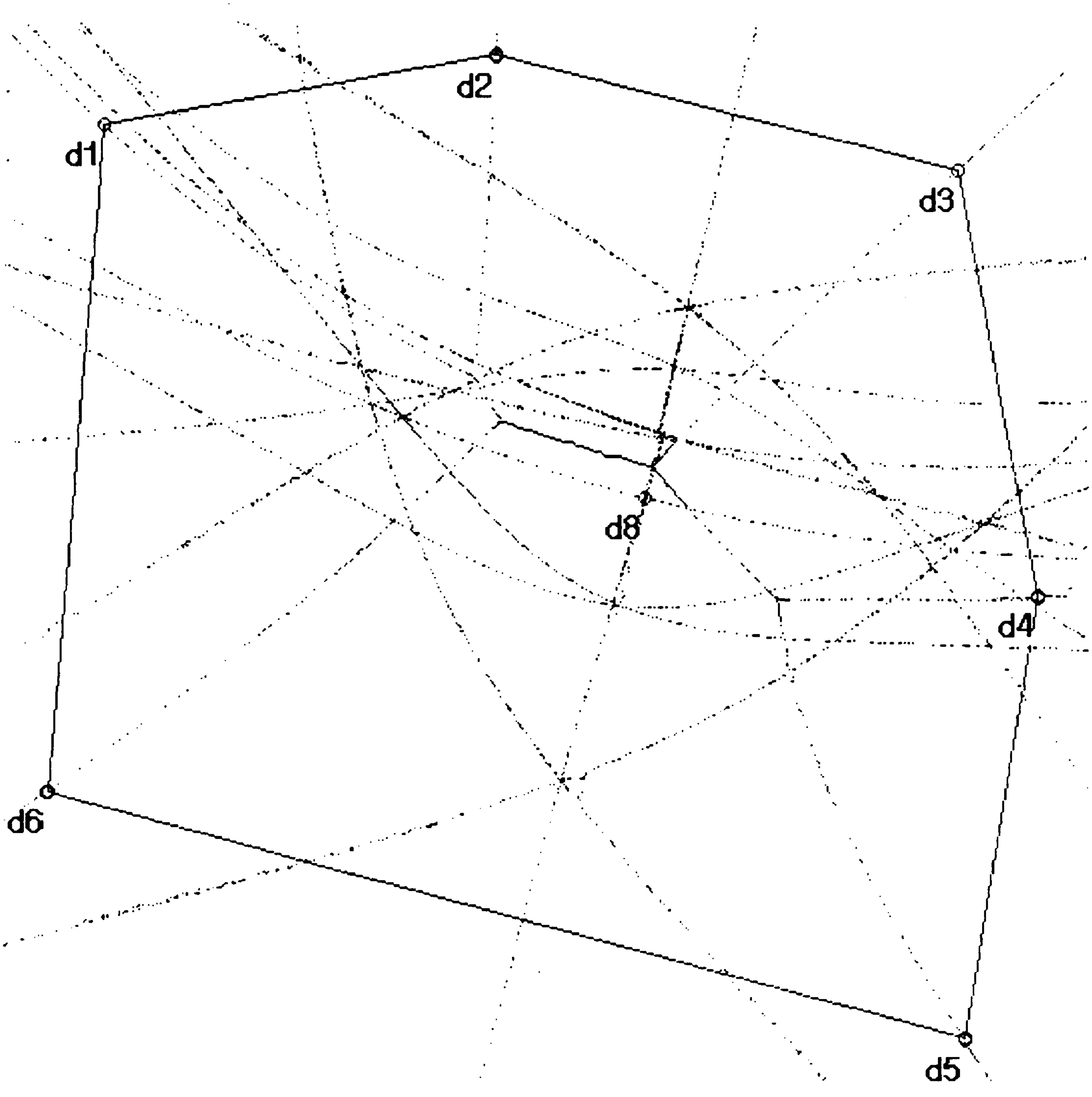


Exhibit A-3. The quartic interplay between segment d6-d5 and segment d3-d2, when attaching d8 and an arbitrary city to the convex hull to produce the shortest Euclidean tour.

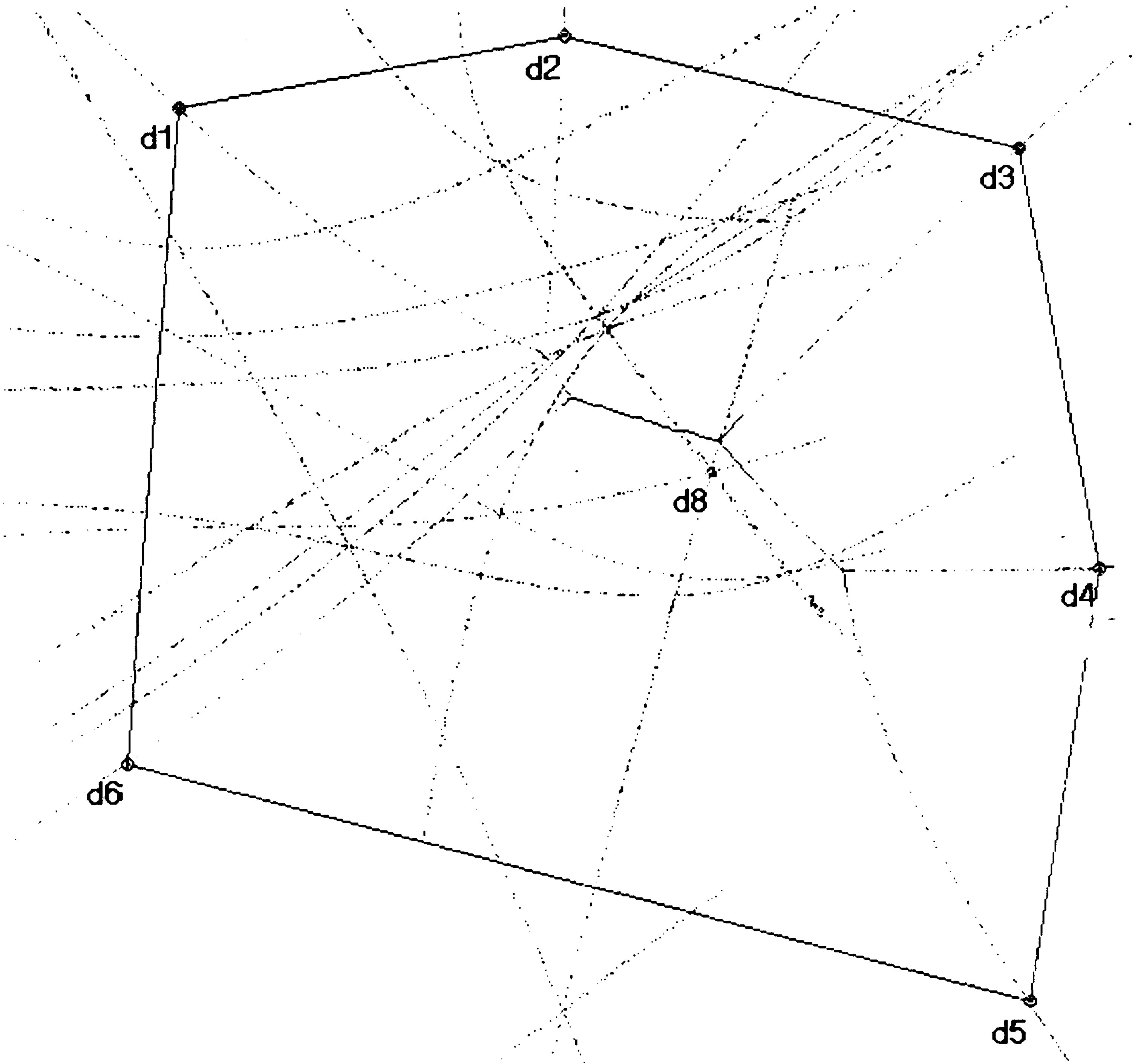


Exhibit A-4. The quartic interplay between segment d6-d5 and segment d2-d1, when attaching d8 and an arbitrary city to the convex hull to produce the shortest Euclidean tour.

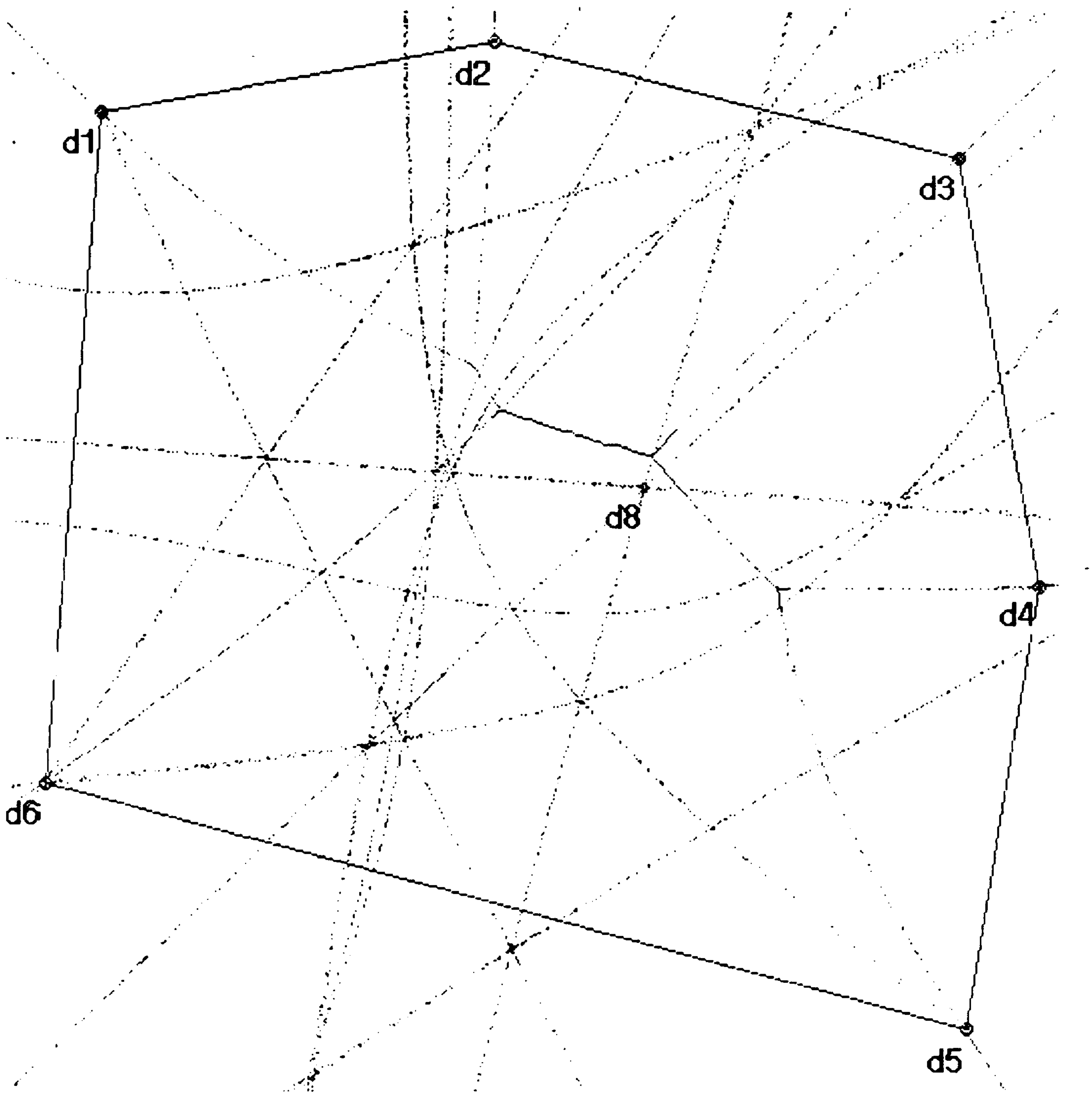


Exhibit A-5. The quartic interplay between segment d6-d5 and segment d1-d6, when attaching d8 and an arbitrary city to the convex hull to produce the shortest Euclidean tour.

Bibliography

[A1] Anonymous, "Just Keep Going", Los Angeles Times, Section II, p. 4, col. 1, 17 June 1985

[A2] Anonymous, "Here's one for the Road, and then Some", Discover, July 1985.

[B1] Barachet, L.L. "Graphic Solution of the Traveling Salesman Problem", Operations Research 5, 1957, pp. 841-845

[B2] Bellmore, M., and G.L. Nemhauser, "The Traveling Salesman Problem: A Survey", Operations Research 16, 1968, pp. 538-558.

[B3] Bern, M.W., and R.L. Graham, "The Shortest Network Problem", Scientific American, January 1989, pp. 84-89

[B4] Brady, R.M., "Optimization strategies gleaned from biological evolution", Nature 317, 1985

[C1] Christofides, N., "Worst-case Analysis of a New Heuristic for the Travelling Salesman Problem", Report 388, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh PA, 1976.

[C2] Chazelle, B., "On the Convex Layers of a Convex Set", IEEE *Trans. Inform. Theory* IT-31, 1985.

[C3] Crowder, H., and M.W. Padberg, "Solving Large-Scale Travelling Salesman Problems to Optimality", Management Science 26, 1980, pp. 495-509.

[D1] Dantzig, G.B., D.R. Fulkerson, and S.M. Johnson, "Solution of a Large-scale Traveling-Salesman Problem", Operations Research 2, 1954, pp. 393-410.

[D2] Dantzig, G.B., D.R. Fulkerson, and S.M. Johnson, "On a Linear Programming, Combinatorial Approach to the Traveling-Salesman Problem", Operations Research 7, 1959, pp. 58-66.

[D3] Dillencourt, M.B., "Graph-Theoretical Properties of Algorithms Involving Delaunay Triangulations", University of Maryland, Center for Automation Research, Technical Report TR-369, September 1988

[D4] Durbin, R., and D. Willshaw, "An analogue approach to the travelling salesman problem using an elastic net method", Nature 326, 16 April 1987

[E1] Eddy, W.F., "A new convex hull algorithm for planar sets", ACM Transactions on Mathematical Software 3, 1977.

[E2] Edelsbrunner, H., Algorithms in Combinatorial Geometry, Springer-Verlag, Berlin Germany, 1987.

[E3] Edelsbrunner, H., G. Rote, and E. Welzl, "Testing the Necklace Condition for Shortest Tours and Optimal Factors in the Plane", Proceedings of the 14th International Colloquium on Automata, Languages, and Programming, T. Ottmann, ed., Karlsruhe, Federal Republic of Germany, July 1987

[E4] Eisenhart, L.P., Coordinate Geometry, Dover Publications Inc., New York NY, 1939

[F1] Fox, G., Private communication, Semi-annual Concurrent Processing technology review of the Joint Tactical Fusion Office, Jet Propulsion Laboratory, Pasadena CA, March 1990.

[G1] Garey, M.R., R.L. Graham, and D.S. Johnson, "Some NP-complete Geometric Problems", Eighth Annual Symp. on Theory of Comput., May 1976, pp. 10-22.

[G2] Garey, M.R., and D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W.H. Freeman and Company, New York NY, 1979

[G3] Golden, B.L., and W.R. Stewart, "Empirical Analysis of Heuristics", in The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, eds. E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, John Wiley & Sons, Chichester UK, 1985

[H1] Harel, D., The Science of Computing: Exploring the Nature and Power of Algorithms, Addison-Wesley, Reading MA, 1987, 1989

[H2] Held, M., and R.M. Karp, "The Traveling Salesman Problem and Minimum Spanning Trees: Part II", Mathematical Programming I, 1971, pp. 6-25.

[J1] Johnson, D.S., and C.H. Papadimitriou, "Computational Complexity", in The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, eds. E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, John Wiley & Sons, Chichester UK, 1985

[J2] Johnson, D.S., and C.H. Papadimitriou, "Performance Guarantees for Heuristics", in The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, eds. E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, John Wiley & Sons, Chichester UK, 1985

[J3] Johnson, D.S., "More approaches to the travelling salesman guide", Nature 330, December 1987, p. 525

[J4] Johnson, D. S., Set of viewgraphs entitled "The Traveling Salesman Problem", AT&T Bell Laboratories, Murray Hill NJ, September 1989.

[J5] Johnson, D. S., Private communication, and set of viewgraphs entitled "How to Beat Lin-Kernighan", AT&T Bell Laboratories, Murray Hill NJ, April 1990.

[K1] Kantabutra, V., "Traveling salesman cycles are not always subgraphs of Voronoi duals", Information Processing Letters 16(1):11-12, January 1983.

[K2] Karp, R.M., "Reducibility among Combinatorial Problems", in Complexity of Computer Computations, ed. R.E. Miller and J.W. Thatcher, Plenum Press, New York NY, 1972.

[K3] Kirkpatrick, D.G., and R. Seidel, "The Ultimate Planar Convex Hull Algorithm?", SIAM Journal Computing 15, 1986.

[K4] Krolak, P., and W. Felts, "A Man-Machine Approach Toward Solving the Traveling Salesman Problem", Communications of the ACM 14, May 1971, pp. 327-335.

[L1] Lawler, E.L., J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, eds., The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, John Wiley & Sons, Chichester UK, 1985.

[L2] Lin, S., "Computer Solutions of the Traveling Salesman Problem", Bell Systems Technical Journal 44, 1965, pp. 2245-2269.

[L3] Lin, S., and B.W. Kernighan, "An Effective Heuristic Algorithm for the Traveling Salesman Problem", Operations Research 21, 1973, pp. 498-516.

[L4] Little, J.D.C., K.G. Murty, D.W. Sweeney, and C. Karel, "An algorithm for the traveling salesman problem", Operations Research 11, 1963, pp. 972-989.

[M1] Mitter, S., Private communication, Semi-annual U.S. Army review of the Center for Intelligent Control, Massachusetts Institute of Technology, Cambridge MA, September 1989.

[P1] Padberg, M., and G. Rinaldi, "Optimization of a 532-city Symmetric Traveling Salesman Problem By Branch and Cut", Operations Research Letters, Vol. 6, No. 1, March 1987.

[P2] Padberg, M., and G. Rinaldi, "A Branch and Cut Algorithm for the Resolution of Large-scale Symmetric Traveling Salesman Problems", New York University technical report, April 1989.

[P3] Papadimitriou, C.H., "The Euclidean Traveling Salesman Problem is NP-complete", Theoretical Computer Science 4, 1977, pp. 237-244.

[P4] Papadimitriou, C.H., and K. Steiglitz, "Some Complexity Results for the Traveling Salesman Problem", Eighth Annual Symp. on Theory of Comput., May 1976, pp. 1-9.

[P5] Preparata, F.P., and M.I. Shamos, Computational Geometry: An Introduction, Springer-Verlag, New York NY, 1985.

[R1] Reinelt, G., "TSPLIB - A Traveling Salesman Problem Library", Institute of Mathematics, University of Augsburg, Augsburg Germany, 20 March 1989.

[S1] Sedgewick, R., Algorithms, Addison-Wesley, Reading MA, 1983.

[S2] Shamos, M.I., "Computational Geometry", Ph.D. Thesis, Department of Computer Science, Yale University, New Haven CT, 1978.

[S3] Steiglitz, K., An Introduction to Discrete Systems, John Wiley & Sons, New York NY, 1974.

[S4] Stewart, W.R., "A Computationally Efficient Heuristic for the Traveling Salesman Problem", Proceedings 13th Annual Meeting of S.E. TIMS, 1977.

[V1] Varaiya, P.P., Notes on Optimization, Von Nostrand Reinhold Company, New York NY, 1972.

[V2] Voronoi, G., "Nouvelles applications des parametres continus a la theorie des formes quadratiques", J. Reine Angew. Math 134, 1908, pp. 97-178.

[Y1] Young, J.W., T. Fort, and F.M. Morgan, Analytic Geometry, Houghton Mifflin, Cambridge MA, 1936.