Shortest Paths in Two Intersecting Pencils of Lines

David Hart*

Abstract

Suppose one has a line arrangement and one wants to find a shortest path from one point lying on a line of the arrangement to another such point. We look at a special case: the arrangement consists of two intersecting pencils (sets of lines where all intersect in a point), and the path endpoints are at opposite corners of the largest quadrilateral formed. The open problem was to find a shortest path in $o(n^2)$ time. We prove here that there are only two possible shortest paths, and the shortest path can thus be computed in $O(n)$ time.

1 Introduction

In this paper, we look at finding shortest paths in an arrangement of lines, a problem that combines two areas of research—arrangements of lines and shortest path problems. We have a line arrangement and two points s and t lying on lines; we would like to find a shortest path traveling on lines of the arrangement and going from s to t. The best algorithm for solving this, known for a decade, has a worst case time of $\Theta(n^2)$. It is an open problem to find the shortest path in line arrangements in $o(n^2)$ time.

In 1996, Jeff Erickson posed a problem [7] of finding the shortest path in a line arrangement with strong constraints: the lines belong to two subsets, and all lines of each subset intersect in a single point (a set of such lines is a *pencil*). The lines radiate out from the points and intersect each other to form a grid of quadrilaterals. We would like to find a shortest path from a near corner to the opposite corner of the largest quadrilateral formed by the intersecting lines. See Figure 1.

1.1 Results

We show in this paper that, for the problem posed by Erickson, the shortest path always follows the outermost lines bounding the largest quadrilateral, giving only two choices of path, and thus making it easy to find the shortest path in $O(n)$ time. Also, the mathematics that we develop has application in simplifying the problem of shortest paths in general line arrangements (though it does not give an asymptotic time improvement).

*Department of Computer Science, University of Aarhus, Denmark. Email: dhart@daimi.au.dk

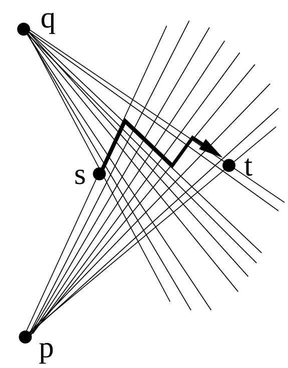

Figure 1: Illustration of the problem input, showing two intersecting pencils of lines, the path endpoints s and t, and a possible path (but not the shortest) between them.

The work in getting this result is the mathematics that proves such a simple choice of paths always gives the correct shortest path.

1.2 Background

The more general problem, of shortest paths in a line arrangement, has been an open problem for 10 years and is listed in several sources: Marc van Kreveld listed it eight years ago in a collection of open problems [11], Jeff Erickson has listed it on a web page [7], and Joe Mitchell lists it in a survey paper [12]. We (with David Eppstein) have discussed the problem and made progress on several special cases (listed below).

The known way to solve this is to apply two algorithms as subroutines. First compute the arrangement, and then apply the linear time, planar graph, shortest path algorithm of Klein, et al [9]. The computation of the arrangement produces a graph with as many as $\Theta(n^2)$ edges and vertices, so this approach takes worst case times and space $\Theta(n^2)$. The space requirement was improved in 1998 to $O(n)$ by Chen, et al [2], but the best known time bound remains $O(n^2)$.

Even for the problem posed by Erickson, no faster algorithm was known.

The first result on exact shortest paths in a line arrangement is due to C. Davis [3] in 1948. He proves that a shortest path would not travel on certain segments of the arrangement.

David Eppstein and the present author looked at special cases of arrangements. We first looked at an arrangement that consists only of vertical and horizontal segments [5] and showed that one can compute a shortest path in $O(n^{1.5} \log n)$ time and $O(n^{1.5})$ space. M. van Kreveld [10] then improved this time to $O(n \log n)$. We next looked at line arrangements where we specify that the number of different line orientations is only k (that is, many of the lines are parallel). Our algorithm finds the shortest path in $O(n \log k + k^2)$ time.

In the latter paper, we also discovered a theorem that applies to general line arrangements and makes it possible to considerably simplify the structure of the arrangement. Using this simplification of the structure of the arrangement as a starting point, the present author described an algorithm [8] for approximating the shortest path with an error bound ϵ in time $O(n \log n + (\min\{n, \frac{1}{\epsilon^2} \log \frac{1}{\epsilon}\}) \frac{1}{\epsilon} \log \frac{1}{\epsilon})$. This improved on a fixed factor of 2 approximation algorithm by Bose, et al [1].

2 Intersecting Pencils

We begin formal treatment of this problem with some definitions.

Definition 1 *Define a* pencil *to be a set of two or more lines that intersect in a single point.*

Let one pencil be the set of lines P intersecting in point p and the second be the set of lines Q intersecting in point q. Define an open half plane bounded by the line through p and q. We require (as part of the definition) that every line of P intersects every line of Q within this half plane (that is, the rays, starting at p and contained in the half plane, intersect all rays starting at q.) This gives a grid of quadrilaterals (see Figure 1). (We note that our solution applies to any intersection of two pencils, not just the quadrilateral grid we discuss here.)

2.1 Preliminaries

We will use the notation p_i and q_j for lines through points p and q; that is, the subscripts imply we are referring to lines. The line through p that makes smallest angle w.r.t. segment $\overline{pq}$ we call p_0 and the line with greatest angle we call p_n. We similarly define q_0 and $q_{n'}$. The remaining lines of each of P and Q are indexed according to the size of angle, from smallest to largest.

Definition 2 *Let the intersection of p_0 and q_0 be point s—the starting point for the path, and let the intersection of p_n and $q_{n'}$ be point t—the ending point of the path.*

Erickson's open problem is to find a shortest path, from s to t in the intersecting pencils of lines, in time $o(n^2)$.

We define *boundary lines* for any shortest path problem to be lines bounding a convex region which entirely contains the shortest path. By our problem definition, the lines are p_0, p_n, q_0, and $q_{n'}$ are boundary lines.

We summarize some facts that are intuitively obvious. We have the path endpoints lying on the corners of a large quadrilateral bounded by the boundary lines, and the path does not travel outside the quadrilateral. The simple cells in the large quadrilateral are all themselves quadrilaterals.

3 Theorem and Proof

We here state the solution to the problem.

Theorem 1 *The shortest path of Erickson's problem travels only on the boundary lines and can be computed in $O(n)$ time.*

To prove this, we next look at a trigonometric expression for a useful number we define below.

3.1 Definition of D

Suppose we have a single quadrilateral, with path endpoints s and t at opposite corners. We label the four sides as in Figure 2. Then a and b are one path and c and d are a second path from s to t. Define $D = a+b-c-d$. Clearly, if D is positive, then the upper sides are the shorter path from s to t along quadrilateral edges, and if negative then the lower sides are the shorter path.

Much of the usefulness of D comes from the following property. Suppose we divide a quadrilateral into two adjacent quadrilaterals by adding a segment across it. Let A and B be the two smaller quadrilaterals and AB be the enclosing quadrilateral, where we define path endpoints for each quadrilateral to be at the opposite corners of that particular quadrilateral, similar to the left in Figure 2. We define D_A, D_B, and D_{AB} for each of these quadrilaterals as described above. Then we have this lemma:

Lemma 1 $D_{AB} = D_A + D_B$

Proof. The sum $D_A + D_B$ adds all the segments (with correct sign) we need to obtain D_{AB}, and the segment crossing the middle of the enclosing quadrilateral AB has opposite signs in the two addends and cancels out. □

We next find a trigonometric expression for D in terms of angles between lines. See Figure 2 (right). We first give some simple equations for various distances. Solving these to eliminate some variables and simplifying gives the expression for D.

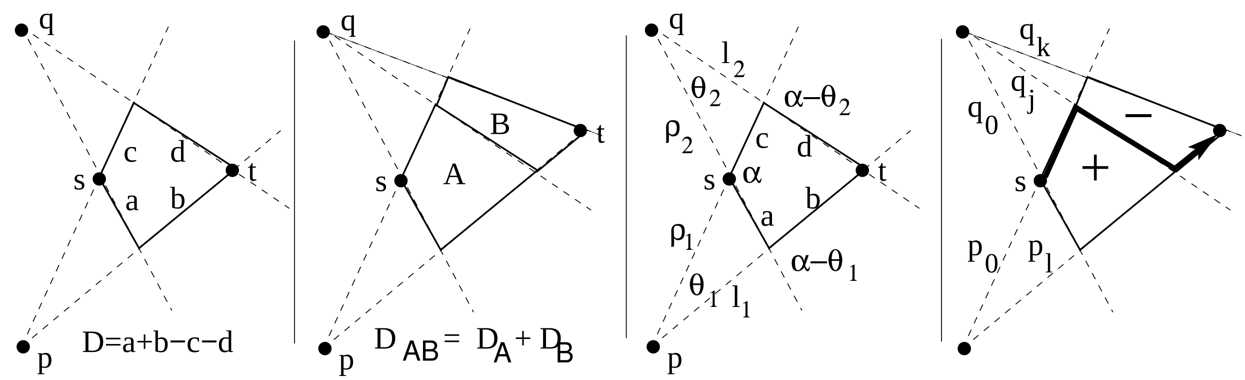

Figure 2: Definition of D (Left). Additivity of D (Middle Left). Definition of variables in trigonometric derivation (Middle Right). Assumption of two turns leads to contradiction (Right).

$$
\begin{aligned}
D &= a+b-c-d \\
l_2 \sin\theta_2 &= c \sin\alpha \\
l_1 \sin\theta_1 &= a \sin\alpha \\
l_2 \cos\theta_2 &= \rho_2 + c\cos\alpha \\
l_1 \cos\theta_1 &= \rho_1 + a\cos\alpha \\
(l_2+d)\sin\theta_2 &= b\sin(\alpha-\theta_1) \\
a &= c\cos\alpha + d\cos\theta_2 - b\cos(\alpha-\theta_1)
\end{aligned}
$$

Solving this as a linear system to eliminate a, b, c, d, l_1, and l_2 and then simplifying the resulting expression using Mathematica, we get:

$$
\begin{aligned}
D = {} & \sec\frac{\alpha-\theta_1}{2}\sec\frac{\alpha-\theta_2}{2}\times \\
& \sec\frac{\alpha-\theta_1-\theta_2}{2}\sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2}\times \\
& \{\rho_1(\cos\frac{\alpha-\theta_2+\theta_1}{2}+\cos\frac{\alpha-\theta_1-\theta_2}{2})- \\
& \rho_2(\cos\frac{\alpha-\theta_1+\theta_2}{2}+\cos\frac{\alpha-\theta_1-\theta_2}{2})\}
\end{aligned}
$$

3.2 Changes in D with Increasing θ_2

We now examine $\frac{\delta D}{\delta\theta_2}$. One could do this by taking the derivative, but the resulting function is quite ugly and difficult to analyze. Instead, we split the above expression for D into three terms and analyze these terms separately.

Define $D = f(\theta_2)(g(\theta_2) - h(\theta_2))$ in the obvious way. It is sufficient for our proof to consider only the case D positive. (Not all of the following holds if D is negative, which is another reason not to work with the function for the derivative).

Lemma 2 *Given that D is positive, $\frac{\delta D}{\delta\theta_2}$ is positive.*

Proof. We analyze the terms f, g, and h, and look at what happens as θ_2 increases. Note that $\alpha < \pi$, and all angles are divided by 2; thus, every trigonometric function involves an angle between 0 and $\frac{\pi}{2}$ and all terms are positive. Look at all trigonometric terms of f; one can see by inspection that they grow larger (or don't change) with increasing θ_2. Thus, f becomes a larger, positive multiplicative factor.

Now look at g. The derivative is $\rho_1(\sin\frac{\alpha-\theta_2+\theta_1}{2} + \sin\frac{\alpha-\theta_1-\theta_2}{2})/2$ The derivative is positive, so g is an increasing, positive number. Last, look at h. Taking the derivative with respect to θ_2, we get $\rho_2(-\sin\frac{\alpha-\theta_1+\theta_2}{2} + \sin\frac{\alpha-\theta_1-\theta_2}{2})/2$. We have $\alpha-\theta_1-\theta_2 < \alpha-\theta_1+\theta_2$ and $\sin\frac{\alpha-\theta_1+\theta_2}{2} > \sin\frac{\alpha-\theta_1-\theta_2}{2}$, so the derivative is negative and h is decreasing.

Given that $\theta_2 > \theta_1$, we have $f_2 > f_1, g_2 > g_1$, and $-h_2 > -h_1$. Then $f_2(g_2-h_2) > f_1(g_1-h_1)$, given that $g_1 - h_1 > 0$. Thus, $\frac{\delta D}{\delta\theta_2}$ is positive given that D is positive. □

3.3 Signs of Adjacent Quadrilaterals

We now prove the following lemma:

Lemma 3 *Given that a pencil Q, with 3 lines, and P, with two lines, intersect, and D_A is known to be positive, then D_B is positive (see Figure 2).*

We have adjacent quadrilaterals A and B contained in a larger quadrilateral AB, (see Figure 2), just as we had in our lemma about additivity of D. Suppose that D_A is positive. Note that quadrilateral AB has the same ρ_1, ρ_2, α, and θ_1 as A; the only difference is that AB has a larger θ_2. To obtain D_{AB}, we integrate $\frac{\delta D}{\delta\theta_2}$, which we know is positive, and we conclude that D_{AB} is a larger positive number than D_A. By additivity of D, thus further tells us that D_B must be positive. □

One can prove a symmetric statement if P is the pencil with 3 lines and we replace the word "positive" with "negative" for values of D.

4 Proof of a Single Turn

Lemma 4 *A shortest path in Erickson's problem has exactly one turn.*

Proof. Suppose for contradiction that the path makes two or more turns. Suppose wlog that the path travels initially on p_0; the other case is symmetric. The path must turn onto a line q_j before reaching q_n (otherwise the entire path has only one turn). The path must next turn onto a line p_l, since q_j does not reach t, and it must intersect a next line of Q (perhaps q_n); call this q_k.

Consider the wedge formed by p_0 and p_l. Three lines, q_0, q_j, and q_k, cross this wedge forming two quadrilaterals A and B enclosed in a larger quadrilateral AB (let A be the quadrilateral with s at one corner). The path travels above quadrilateral A and below quadrilateral B (since we said the path travels on line q_j, which is the segment crossing the middle of quadrilateral AB). The value D_A must be positive; otherwise there would be a shorter path traveling on the lower sides of A. Then by our lemma, D_B is also positive. But then a path reaching the corner of B and traveling on the upper sides of B is shorter than what we claimed was the shortest path, a contradiction. □

The immediate conclusion is the statement of Theorem 1: A shortest path in Erickson's problem can travel only on the (four) boundary lines.

5 Remarks

The solution extends to any intersection of two pencils (not just the simple quadrilateral grid here); we omit the details. Also, the mathematics we use here can be applied to simplify shortest path problems in general line arrangements, though this does not give an asymptotic time improvement.

Acknowledgments. The author thanks David Eppstein and the anonymous reviewers of CCCG for suggestions that led to a more clearly written paper.

References

[1] P. Bose, W. Evans, D. Kirkpatrick, M. McAllister, and J. Snoeyink. Approximating shortest paths in arrangements of lines. In *Proc. 8th Canad. Conf. Computational Geometry*, pp. 143-148, 1996.

[2] D. Chen, O. Daescu, X. Hu, and J. Xu. Finding an optimal path without growing the tree. *6th Annual European Sym. on Algorithms*, Springer LNCS, pp. 356-367, 1998.

[3] C. Davis. The short-cut problem. *American Mathematical Monthly*, pp. 147-150, March, 1948.

[4] H. Edelsbrunner, J. O'Rourke, and R. Seidel, Constructing arrangements of lines and hyperplanes with applications, SIAM J. Comput. 15, 1986, 341–363.

[5] D. Eppstein and D. Hart. An efficient algorithm for shortest paths in vertical and horizontal segments. In *Proc. 5th Worksh. Algorithms and Data Structures*, pp. 234–247. Lect. Notes in Comp. Sci. 1272, Springer-Verlag, 1997.

[6] D. Eppstein and D. Hart. Shortest paths in an arrangement with k-orientations. *Proceedings of the 10th ACM-SIAM Symposium on Discrete Algorithms*, pp. 310-316, 1999.

[7] J. Erickson. Shortest paths in line arrangements (entry in open problem list). `http://www.cs.uiuc.edu/~jeffe/open/algo.html\#shortpath`.

[8] D. Hart. Approximating the shortest path in line arrangements. In *Abstracts of the 13th Canadian Conference of Computational Geometry,* pp. 105-108, University of Waterloo, 2001.

[9] P. N. Klein, S. Rao, M. H. Rauch, and S. Subramanian. Faster shortest-path algorithms for planar graphs. In *Proc. 26th Symp. Theory of Computing*, pp. 27–37. ACM, 1994.

[10] M. van Kreveld. Correspondence.

[11] M. van Kreveld. Open problem in *Report on 4th Dagstuhl Seminar on Computational Geometry*, Dagstuhl Seminar Rep. 109, p. 20, 1995.

[12] J.S.B. Mitchell. Geometric shortest paths and network optimization. In J.-R. Sack and J. Urrutia, editors, Handbook of Computational Geometry. Elsevier Science, Amsterdam, 1998.