Exact Solutions for Simple Weighted Region Problems

Sada Narayanappa[2], Petr Vojtěchovský[2], Wan D. Bae[1]

[1] Department of Computer Science
[2] Department of Mathematics
University of Denver
2360 South Gaylord Street, Denver, CO 80206, U.S.A.
{snarayan, wbae}@cs.du.edu, petr@math.du.edu

Abstract. The weighted region problem (WRP) is to find an optimal path from s to t, where s and t are the vertices of a planar graph and path is allowed to go through the faces. Finding optimal paths is an important problem in robotics, computational geometry, etc. Exact solutions for WRP are known in some special cases. However, the computational complexity of WRP is unknown. We present the exact solutions of simple weighted region problem in presence of strips, triangles, and convex polygons.

Key words: shortest path, weighted region problem, Computational Geometry, Computational Complexity

1 Introduction

The *Weighted region problem (WRP)* is stated as follows: The input is a plane subdivided into polygonal regions. Each polygon is associated with a nonnegative weight α with a possibility of assigning different weights to the boundaries of the polygon. The weight α denotes the "cost per unit distance" of traveling in its associated region. The objective is to find an optimal path (least cost path) from a given point s to a given point t according to the weighted Euclidean metric.

A special case of WRP in which the weights are taken from the set $\{0, 1, \infty\}$ is addressed in [4]. It is solved exactly by constructing a graph known as a *critical graph*, which is an extension of the standard *visibility graph.* [4] also concludes with an open question about the complexity of the latter problem if the weights are taken from the set $\{1, 2\}$.

In this paper, we present exact solutions to certain specific WRPs. We also discuss the complexity of the general case. Without loss of generality, we assume the following throughout this chapter:

1. All regions are convex (easily achieved by triangulating the given subdivisions).
2. Unless stated otherwise, s and t are the vertices of the input.

3. Unless stated otherwise, the weight of the background (outside of the regions) is assumed to be one. The weight of all polygons is greater than or equal to one.
4. Unless stated otherwise, the weight of an edge is the minimum of the weights of its adjacent regions. Notice that it is imperative to postulate that the weight of an edge is less than or equal to the minimum cost of its adjacent regions. Without this assumption, an optimal path does not have to exist since it might be possible to construct a series of consecutively cheaper paths converging to an edge from the side of a cheaper polygon.

2 Preliminaries

For any region r_i, we use α_i to denote the weight of r_i. For a boundary edge e, the unit weight is defined to be $min\{\alpha_1, \alpha_2\}$, where r_1 and r_2 are the two regions incident to e. The edge e shared by regions r_1 and r_2 is denoted by $e = \cap(r_1, r_2)$. We use $int(x)$ to denote the relative interior of an edge or a region.

It has been proved in [3] that an optimal path in a WRP is piecewise linear and simple (non-intersecting), consisting of $O(n^2)$ segments, where n is the number of regions in the input. We can specify the piecewise linear paths by listing the sequence of points that represent the end points of the linear subpaths. The weighted length $||xy||$ of a line segment joining two consecutive points x and y is equal to $\alpha \cdot |xy|$, where $|xy|$ is the *Euclidean distance* from x to y, and α is either the weight of an edge e if $\overline{xy}$ lies on e or the weight of a region intersecting $\overline{xy}$. The weighted length of a path is then the sum of the weighted lengths of its subpaths. An optimal path between two points s and t is a path that has the least cost among all the paths from s to t.

While attempting to solve WRP, many authors note and take advantage of the analogy between WRP and the propagation of light. Fermat's principle in optics [1] states that light always travels from one point to another along the quickest (cheapest) route. However, it is paramount to realize that the optimality of the route asserted by Fermat's principle is valid only among all routes "that the light can actually take". But there are instances of WRP where Fermat's principle applies. Most notably the situation of Figure 1, where the optimal solution follows from Snell's law of refraction of light.

Consider an optimal path from s in a region (medium) with weight α_1, to a point t in a region with weight α_2. Let e be the edge shared by regions r_1 and r_2. This case is shown in Figure 1. If $\alpha_1 = \alpha_2$, then the shortest route is simply $\overline{st}$. Otherwise, the optimal path consists of $\overline{sx}$ and $\overline{xt}$, where x is a point on e between s and t.

The cost function, F, can be expressed as a differentiable function of x whose domain is $[0, d]$, and we want to find the absolute minimum value of F on this closed interval. Clearly,

$$F(x) = \alpha_1\sqrt{x^2 + y_s^2} + \alpha_2\sqrt{(x_t - x)^2 + y_t^2} \ .$$

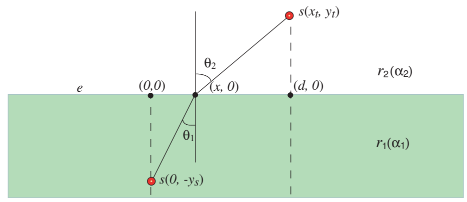


Fig. 1. Optimal path from s to t

We find the optimal value for x by setting $F'(x) = 0$. Then

$$F'(x) = \alpha_1 \frac{x}{\sqrt{x^2 + y_s^2}} - \alpha_2 \frac{x_t - x}{\sqrt{(x_t - x)^2 + y_t^2}}. \tag{1}$$

Thus $F'(x) = 0$ if and only if

$$\alpha_1 \ \sin\theta_1 = \alpha_2 \sin\theta_2 \tag{2}$$

where θ_1 and θ_2 are as shown in Figure 1.

We can see from (1) that $F'(0) < 0$ at $x = 0$ and $F'(d) > 0$. Since F' is a convex function on $[0, d]$, there exists a unique point $x \in [0, d]$ such that $F'(x) = 0$. Formula (2) is *Snell's Law* or the *Law of Refraction*.

(2) gives the optimal angle of incidence of the path from s to t. If we set $y_t = 0$ (moving t on e), and $\alpha_2 < \alpha_1$, either the optimal path is the line segment $\overline{st}$, or it must travel along e. In the latter case, $\theta_2 = \pi/2$, therefore, $\theta_1 = \sin^{-1}(\alpha_2/\alpha_1)$ is called the *critical angle* defined by e. The critical angle θ_c is defined as follows: when an optimal path goes from a region r_1 to a less expensive region r_2 then $\theta_c = \sin^{-1}(\alpha_1/\alpha_2)$.

Solving for x in (1), we obtain a 4^{th} degree polynomial. Here we scale weights by setting $\alpha_2 = 1$, $\alpha = \alpha_1/\alpha_2$, and $(1 - \alpha^2) = \beta$ we get

$$\beta x^4 + 2x_t\beta x^3 + (\beta x_t^2 - \alpha^2 y_t^2 + y_s^2)x^2 + 2x_t y_s^2 x - x_t^2 y_s^2 = 0 \tag{3}$$

Claim. An optimal path is never incident upon an edge at an angle greater than the critical angle.

A path p is said to *pass through* an edge $e = \cap(r_1, r_2)$ at point y, if there exists a point $x \in r_1$ and $x' \in r_2$ such that $\overline{xy} \in int(r_1) \subset p$ and $\overline{yx'} \in int(r_2) \subset p$. If a path p contains a subpath $\overline{uv} \subset e$, then p is called an *edge shared path* and edge e is called a *shared edge*.

Lemmas 1, 2, and 3 are proved in [3] (Lemmas 3.3, 3.6, and 3.7). We state these Lemmas for the convenience of the reader.

Lemma 1. *Let $e = \cap(r_1 r_2)$ be an edge and $\alpha_e = min\{\alpha_1, \alpha_2\}$. An optimal path p passing through the interior of the edge e is incident on e at an angle less than the critical angle θ_c defined by e. If p contains the subpath $\overline{uv}$, where $u, v \in int(e)$, then the angle of incidence at points u and v must be the critical angle θ_c.*

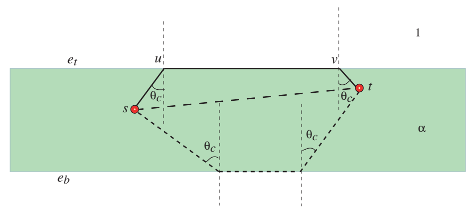


Fig. 2. Optimal path from s to t

Lemma 2. *An optimal path p passing through the edges $(e_1, e_2, \ldots, e_k)$ from point s to a point t with $e_i \neq e_{i+1}$ is the unique optimal path, and it obeys Snell's law at each crossing point of each edge.*

Lemma 3. *Let p be an optimal path. Then either (1) between any two consecutive vertices of p, there is at most one critical point of entry to an edge e, and at most one critical point of exit from an edge e' (possibly equal to e); or (2) p can be modified in such a way that (1) holds without altering the length of the path.*

3 Paths through strips

We call an unbounded region between two parallel lines e_t and e_b a *strip*, as shown in Figure 2. Without loss of generality, we let e_t and e_b be the top and bottom border of the strip, respectively, and we assume that e_t, e_b are horizontal.

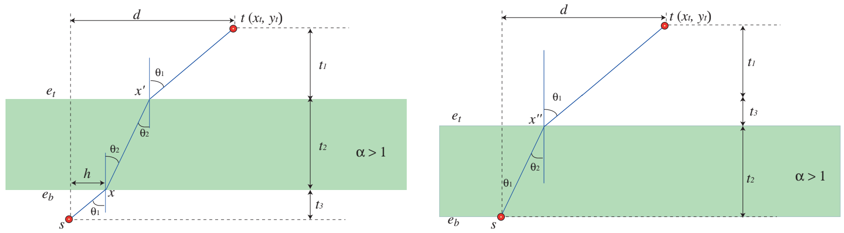
This section presents a solution for a WRP consisting of strips of equal weight. First, consider the optimal path from point s to t through a single strip $\mathcal{S}$. Let the weight of the strip be $\alpha = 1+\epsilon, 0 \leq \epsilon < \infty$. Clearly, if $\epsilon = 0$, then the optimal path is $\overline{st}$. There are three cases to consider: (1) $s \in \mathcal{S}, t \notin \mathcal{S}$, (2) $s, t \in \mathcal{S}$, (3) $s, t \notin \mathcal{S}$.

Case (1) ($s \in \mathcal{S}, t \notin \mathcal{S}$) can be solved using (3).

Case (2) ($s, t \in \mathcal{S}$) The optimal path is either (a) $\overline{st}$, or (b) consists of three segments ($\overline{su}, \overline{uv}, \overline{vt}$), where u and v are points on e_b or e_t. This is shown in Figure 2. By Lemma 1, $\overline{su}$ and $\overline{vt}$ are incident at the critical angle. Therefore, both (a) and (b) can be computed in time $O(1)$. *NOTE*: The optimal path in case (b) need not be unique. Anytime s, t are placed symmetrically with respect to the line $(e_b + e_t)/2$, we have two solutions.

Consider the case where s and t are on different edges, and the critical angle is θ_c, furthermore, angle of incidence of $\overline{st}$ is greater than θ_c. Then, by Claim 2 an optimal path must cross $\mathcal{S}$ at a critical angle and share an edge. Crossing $\mathcal{S}$ can occur at infinitely many points between s and t, thus there are infinitely many paths in such case.

Case (3) ($s, t \notin \mathcal{S}$) If $\overline{st} \cap \mathcal{S} = \emptyset$, then $\overline{st}$ is optimal. Without loss of generality, assume s is below e_b and t is above e_t as shown in Figure 3(a). Claim 3 shows that this case is equivalent to Case(1) and we are done.



(a) (b)

Fig. 3. $s, t \notin \mathcal{S}$ can be solved by Snell's Law by shifting $\mathcal{S}$

Claim. Consider the two cases:

a) Assume $s, t \notin \mathcal{S}$ and $\overline{st} \cap \mathcal{S} \neq \emptyset$ as shown in Figure 3(a).
b) Slide $\mathcal{S}$ from case (a) such that $s \in e_b$ as shown in Figure 3(b).

The angle of incidence θ_1 on e_t in both cases must be equal.

Proof. The optimal solution to case (a) consists of three segments ($\overline{sx}, \overline{xx'}, \overline{x't}$). By Lemma 2, we have

$$\sin\theta_1 = \alpha \sin\theta_2,$$
$$\sin\theta_3 = \alpha \sin\theta_2,$$

$$\therefore \theta_1 = \theta_3, \qquad \text{since } 0 \leq \theta_i \leq \pi. \tag{4}$$

θ_1, θ_2, and θ_3 are the angles of incidence at points x, x', and t respectively. Therefore, we can rearrange segments to obtain a solution to case (b). Clearly, this processes can be reversed, i.e., a solution for case (b) can be split to obtain a solution to case (a). Therefore, both cases are equivalent.

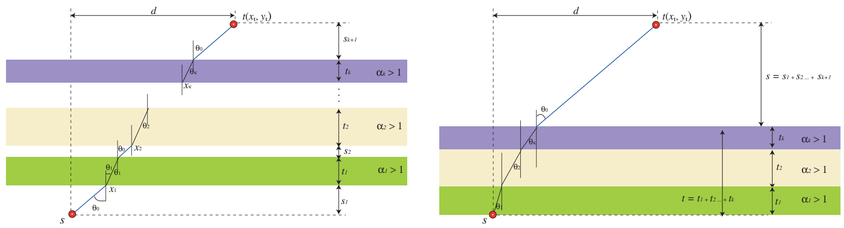
3.1 Multiple strips of different weights

Consider the case of a WRP having more than one strip each having some weight, not necessarily the same. Let $\mathcal{S}_i, t_i, \alpha_i$ denote the strip, thickness, and weight of the i^{th} strip, respectively. Let θ_i', θ_i be the angle of incidence of the optimal path inside the i^{th} strip and the angle in the background preceding the i^{th} strip, respectively. This case is shown in Figure 4(a). Again, by Lemma 2, we have

$$\left.\begin{aligned} \sin\theta_1 &= \alpha_1 \sin\theta_1', \\ \sin\theta_2 &= \alpha_1 \sin\theta_1', \end{aligned}\right\} \text{implies } \theta_1 = \theta_2$$
$$\vdots$$

Since $0 \leq \theta_i \leq \pi$, we have

$$\theta_1 = \theta_2 = \ldots = \theta_{k+1}. \tag{5}$$



(a) (b)

Fig. 4. Multiple strips of varying weights can be grouped

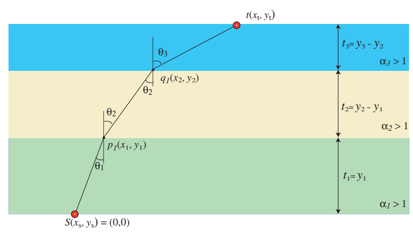


Fig. 5. Three strips of different weights

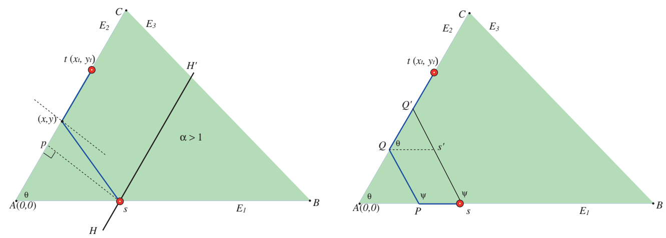
Whenever $\alpha_i = \alpha_j$, we get $\theta_i = \theta_j$. The total weighted length of the optimal path intersecting the regions of equal weights must remain the same regardless of the order in which these strips appear. Henceforth, the complexity of solving the multiple strips case is same regardless of the order in which the strips appear in the input.

Path through three strips In order to answer the question about the complexity of the general WRP, one can employ the following idea: If there are instances of WRP whose exact solutions hinge on solving a (general) polynomial of degree 5 or more, then WRP is not solvable by radicals (see [2]). As we have seen, Snell's law depends on solving a polynomial of degree 4. To obtain a polynomial of degree higher than four, one must consider a more complicated setting than WRP consisting of strips of equal weight.

Therefore, consider a case of a WRP with three strips having different weights, α_1, α_2, and α_3 as shown in Figure 5. We can use Snell's law in order to write the equation for an exact solution to this case.

$$\alpha_1 \sin\theta_1 = \alpha_2 \sin\theta_2 = \alpha_3 \sin\theta_3$$

This gives us the 12^{th} degree polynomial. Unfortunately, we were unable to produce a WRP configuration that yields a polynomial about which we can demonstrate that it is not solvable by radicals, but the above calculation indicates strongly that such a polynomial exists.



(a) Snell's law (b) proof of Observation 1

Fig. 6. Optimal path from s to t

4 Paths through a triangle

In this section, we consider a solution to a WRP consisting of a triangle. Throughout this section, $\triangle ABC$ denotes a triangle ABC having edges $\overline{AB} = E_1$, $\overline{AC} = E_2$, $\overline{BC} = E_3$, $\angle BAC = \theta$ and $\angle ABC = \theta_1$ and $\alpha > 1$ is the weight of the inside of the triangle. Let $s \in E_1$ and $t \in E_2$. If we draw a line $\overline{HH'}$ parallel to $\overline{AC}$ that passes through s, then the case can be viewed as a special case of a single strip when an optimal path cuts through the face. This case is shown in Figure 6(a).

Observation 1 *An optimal path p from $s \in E_1$ to $t \in E_2$ in a triangle with edges E_1, E_2, E_3 (1) does not bend twice as shown in Figure 6(b), i.e., path cannot be incident upon the edge at some point $P \in E_1$ and $A \neq s$ and at some point $Q \in E_2$ and $A \neq t$, or (2) p can be modified such that (1) holds without affecting its cost.*

Proof. Proof follows from similar triangles APQ and AsQ'. Complete proof will be provided in full version of the paper.

Observation 2 (Always Cut) *If there exists two points $x \in E_1$ and $y \in E_2$ such that $\overline{xy}$ is an optimal path, then for every two points $u \in E_1$ and $v \in E_2$, we can find an optimal path that does not go through the vertex A unless $u = A$ or $v = A$.*

Proof. If $\overline{uv}$ is parallel to $\overline{xy}$, by similarity of the triangles $\triangle uAv$ and $\triangle xAy$, we have $\overline{uv}$ which is optimal. Now consider a segment parallel to $\overline{xy}$ passing through v incident on E_1 at u'. Without loss of generality, let u' be closer to A. Take the optimal path to be $\overline{uu'}, \overline{u'v}$

Theorem 1 (Critical Weight). *An optimal path p from $s \in E_1$ to $t \in E_2$ cuts through the face of the triangle if and only if $\alpha \leq \frac{1}{\sin(\theta/2)}$.*

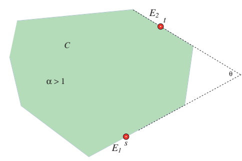


Fig. 7. Optimal path from s to t in a convex polygon

Proof. We want to find α such that the optimal path for some points $s \in E_1$ to $t \in E_2$ is such that:

$$\alpha \cdot |st| > |sA| + |At| \tag{6}$$

Find the minimum value of α, by varying $|At|$, we get $|sA| = |At|$.

$$\alpha > \sqrt{\frac{2}{1-\cos\theta}} = \frac{1}{\sin(\theta/2)}. \tag{7}$$

If e_1 and e_2 are the boundary edges of a region r, a path never passes through r if the weight of r is $\alpha_r > \frac{1}{\sin(\theta/2)}$, where θ is the angle extended by e_1 and e_2. We refer to the maximum of such weight as a *critical weight* defined by the angle extended by e_1 and e_2.

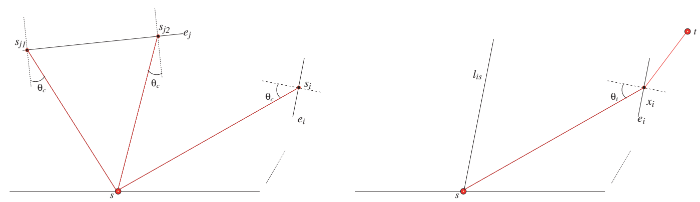
5 Convex polygon

In this section, we present a solution to WRP consisting of a convex polygon $\mathcal{C}$. Our results provide a solution except in the case when s and t are outside of $\mathcal{C}$ and the optimal path consists of three line segments $\overline{sx}$, $\overline{xx'}$, and $\overline{x't}$, where x and x' are points on some edges of $\mathcal{C}$.

Let E_c and V_c denote the set of vertices and edges of $\mathcal{C}$, respectively and $|E_c| = |V_c| = n$. Let α be the weight assigned to the inside of $\mathcal{C}$, $\alpha = 1 + \epsilon, 0 \leq \epsilon < \infty$. The weight of the background (unbounded face) is one. If $\epsilon = 0$, then the optimal path is $\overline{st}$, since $\mathcal{C}$ is convex.

First consider an optimal path from $s \in E_1$ to $t \in E_2$, where E_1 and E_2 are some edges of $\mathcal{C}$ as shown in Figure 7. Let the angle extended by the edges E_1, E_2 be θ. If $\alpha \geq 1/\sin(\theta/2)$, then according to Theorem 1, the optimal path must travel along the border; in this case, we can compute the path traveling along the edges of $\mathcal{C}$ from s to t in $O(n)$ time.

When $\alpha < 1/\sin(\theta/2)$, an optimal path must cut through $\mathcal{C}$ by Theorem 1. Therefore, by Lemma 3, there is at most one shared edge $e_c \in E_c$. Furthermore, the angle of incidence on this edge from s and t must be the critical angle (by Lemma 1). One can employ a binary search on the edges of $\mathcal{C}$ to find e_c in time $O(\log n)$, since the angle of incidence on the edges of a convex polygon is a convex function from a fixed point.



(a) V_s (b) V_a

Fig. 8. V_s and V_a

An optimal path from $s \in E_i$ to $t \in E_j$, where E_i and E_j are some edges of $\mathcal{C}$, must be one of the following:

- straight segment $\overline{st}$,
- $\overline{sx}$, $\overline{xt}$ where x is the point of critical incidence on some edge E_j,
- $\overline{sx_1}, \overline{x_1x_2}, \overline{x_2t}$, where x_1, x_2 are the points of the critical incidence on some edge $e_c \in E_c$ from s and t, respectively,
- a path traveling along the edges of the polygon.

Therefore, when $\alpha < 1/\sin(\theta/2)$, we can find an exact solution in time $O(\log n)$.

Now we give a construction of *weighted exact graph*, or *exact-graph* for short, $G(V, E)$ that contains edges of the optimal path from s and t. Furthermore, the number of edges of G is proportional to n. Therefore, we can find an optimal path in time $O(n \log n)$.

Let $e_i \in E_c, v_i \in V_c$ be the i^{th} edge, vertex, respectively. Let vertices of G be

$$V = s \cup t \cup V_c \cup V_s \cup V_t \cup V_a.$$

For an edge e_i, let s_i be a point on e_i such that $\overline{ss_i}$ is incident on e_i at a critical angle θ_c. See Figure 8(a) for an example. Notice that there are at most two such points on any edge. V_s consists of all points $s_{ij}, i = [1, 2, \ldots, n], j = [1, 2]$.

Similarly, For an edge e_i, let t_i be a point on e_i such that $\overline{tt_i}$ is incident on e_i at a critical angle θ_c. See Figure 8(a) for an example. Notice that there are at most two such points on any edge. V_s consists of all points $t_{ij}, i = [1, 2, \ldots, n], j = [1, 2]$.

Let l_{is} be a line through s parallel to e_i. Let x_i be the point of incidence along edge e_i obtained by applying Snell's law (1). Notice that x_i may not exist. V_a consists of such points x_i for $i = [1, 2, \ldots, n]$. The edge set E consists of:

$$E_p \cup E_s \cup E_t \cup E_a \cup E_v \cup e_{st}$$

$$E_s = \{(s, x) : x \in V_s\}$$

$$E_t = \{(t, x) : x \in V_t\}$$

$$E_a = \{(s, x), (x, t) : x \in V_a\}$$

$$E_v = \{(s, v) : x \in V_c \text{ and } s \notin \mathcal{C}\} \cup \{(t, v) : x \in V_c \text{ and } t \notin \mathcal{C}\}$$

Notice that the points V_s, V_t, and V_a on the edges of $\mathcal{C}$ partitions the edges. E_p consists of all the partitions with their respective end points: $e_{st} = (s, t)$, if $\overline{st}$ does not intersect more than one region except at the end points. Assign a cost to an edge equal to its weighted length.

We can construct an exact graph in time $O(n \log n)$. Since we add a constant number of points on each edge, and these points split an edge into a constant number of partitions, the number of edges and vertices in an exact graph is linearly proportional to the number of edges of $\mathcal{C}$. An optimal path from s to t must be along the edges of exact graph. Therefore, a shortest path algorithm finds the shortest weighted path from s to t in time $O(n \log n)$.

5.1 Paths through regular n-gons

Lemma 4. *If $\alpha \geq 2$, an optimal path never cuts through a regular n-gon. Therefore, such regions can be considerd as a region of ∞ weight (obstacles).*

Proof. Follows directly from Theorem 1.

Observation 3 *A convex polygon has at most 3 internal angles greater than $\pi/3$.*

Observation 4 *A convex polygon having $n > 3$ sides has at most 2 internal angles less than $\pi/3$.*

Lemma 5. *A special case of WRP in which the regions are regular n-gons having a weight $\alpha \geq 2$ can be solved by using the visibility techniques.*

Proof. Consider the special case of WRP in which weights are assigned to polygons from the set {1,2}. [4] conjectures that this problem is as hard as the general WRP problem, in which there are no restrictions on the choice of weights except being nonnegative. By Lemma 4, a WRP consisting of only regular n-gons each of weight $\alpha \geq 2$ can be considered as obstacles, hence WRP can be solved by the technique as given in [4].

Acknowledgments. Authors would like to express thanks to Shayma Alkobaisi and Brandon Haenlein for helpful discussions and feedback.

References

1. Ross L. Finney, Maurice D. Weir, and Frank R. Giordano. *Thomas' CALCULUS Early Transcendentals.* Addison Wesley, 1989.
2. L. Gaal. *Classical Galois Theory.* Markham, 1971. Reprinted by Chelsea, New York, 1979.
3. J. S. Mitchell and C. H. Papadimitriou. The weighted region problem: Finding shortest paths through a weighted planar subdivision. *Jour. On ACM.*, 1990.
4. Gewali L. P., Meng A. C., Mitchell J. S., and Ntafos S. Path planning in $0/1/\infty$ weighted regions with applications. *ORSA Jour. On Comp.*, 1990.