



On Landmark Distances in Polygons

C. Gotsman¹  and K. Hormann² 

¹New Jersey Institute of Technology, Newark, New Jersey, USA

²Università della Svizzera italiana, Lugano, Switzerland

Abstract

We study the landmark distance function between two points in a simply connected planar polygon. We show that if the polygon vertices are used as landmarks, then the resulting landmark distance function to any given point in the polygon has a maximum principle and also does not contain local minima. The latter implies that a path between any two points in the polygon may be generated by steepest descent on this distance without getting “stuck” at a local minimum. Furthermore, if landmarks are increasingly added along polygon edges, the steepest descent path converges to the minimal geodesic path. Therefore, the landmark distance can be used, on the one hand in robotic navigation for routing autonomous agents along close-to-shortest paths and on the other for efficiently computing approximate geodesic distances between any two domain points, a property which may be useful in an extension of our work to surfaces in 3D. In the discrete setting, the steepest descent strategy becomes a greedy routing algorithm along the edges of a triangulation of the interior of the polygon, and our experiments indicate that this discrete landmark routing always delivers (i.e., does not get stuck) on “nice” triangulations.

CCS Concepts

• **Mathematics of computing** → **Paths and connectivity problems; Graph algorithms;** • **Theory of computation** → **Routing and network design problems;**

1. Introduction

Designing routes (or paths) between points in a given planar domain is an important problem arising in many contexts, for example, network routing and robotic navigation. It is especially important to be able to easily generate routes on the fly which have desirable properties, such as short length or moderate curvature and rely only on local knowledge of the domain [SBD*13]. One way to do this is to associate with each domain point a set of *coordinates*, and to define a distance function between points based on those coordinates. A route between a source point and a target point may then be generated by moving from the source to the target along a path in the domain which always decreases this distance from the target. The main challenge is to design a distance function with a global minimum at the target and no other local minima in the domain. This guarantees that a steepest descent strategy always succeeds, namely, to never get “stuck” until the target is reached. We derive a distance function with this property, which additionally approximates the geodesic distance function. It can be evaluated knowing only a small number of precomputed geodesic distances from the target and the current point to certain *landmarks*, but without global knowledge of the domain. The latter would instead be needed for computing the exact geodesic distance, which is not difficult *per se*, using the visibility graph of the domain, and part of our precomputation step.

In this paper, we consider the case where the domain is a simple

planar polygon P with $n \geq 3$ vertices p_1, \dots, p_n , and we study the use of the landmark distance, based on the classical geodesic distance, for routing. For any two points $x, y \in P$, the *geodesic distance* $g(x, y)$ is the length of the shortest path between x and y within P . Given a set of m *landmark* points $\ell_1, \dots, \ell_m \in P$, we denote the geodesic distance between any $x \in P$ and ℓ_i by $g_i(x) = g(\ell_i, x)$. For a fixed *target* point $t \in P$, we define the *landmark functions*

$$\lambda_i(x) = |g_i(x) - g_t(x)|, \quad i = 1, \dots, m \quad (1)$$

and the *landmark distance* between t and x as

$$\mathcal{L}_t(x) = \max_{1 \leq i \leq m} \lambda_i(x). \quad (2)$$

Note that the landmark functions (see Figure 6) and the landmark distance (see Figure 5) are non-negative and $\mathcal{L}_t(t) = 0$. If $\mathcal{L}_t(x) = \lambda_i(x)$, then we say that $\mathcal{L}_t(x)$ is *determined* by the landmark ℓ_i . If this is true also for the landmark ℓ_j , then we say that $\mathcal{L}_t(x)$ is *co-determined* by ℓ_i and ℓ_j . For example, $\mathcal{L}_t(t)$ is trivially co-determined by *all* landmarks.

After introducing the concept of generalized hyperbolas based on the geodesic distance and deriving an intimate connection to the landmark distance (Section 3), we primarily deal with the case where the landmarks ℓ_i are naturally chosen to be the vertices of P and the target t is an arbitrary point within P . Our main objective is to show that, in this case, $\mathcal{L}_t(x)$ is free of local minima within P , except for the unique global minimum at $x = t$. Therefore, the

steepest descent path, starting at any *source* $s \in P$ leads to the target t without getting stuck at a local minimum (Section 4).

The landmark distance induces a partition of P into regions, where a region is characterized by a common set of co-determining landmarks. We discuss the geometry of this partition, and observe that the steepest descent paths are piecewise conics that have an initial straight segment and then follow the boundaries of these regions (Section 5).

While the steepest descent path is not identical to the geodesic path, in practice it is quite similar, and it reproduces the latter if the landmarks are dense along the boundary of P (Section 6.2). We further observe that taking only a subset of the polygon vertices as landmarks is usually not sufficient to preserve the landmark distance (Section 6.1) and that the absence of local minima property does not hold in the case of multiply connected polygons (Section 6.3). Despite this, the landmark distance is still a good approximation of the geodesic distance when landmarks are dense along the polygon edges. Finally, we discuss possible extensions of our method to the computation of geodesic distances on 2-manifold surfaces with boundaries in 3D (Section 6.4) and to discretizations of the continuous domain by well-formed triangulations (Section 6.5). Our experiments demonstrate that the discrete analogue of the steepest descent path – so-called “greedy routing” – always “delivers”, specifically, it finds a path along the edges of a triangulation T of the interior of P , between any two vertices of T without getting stuck, as long as T is “nice”.

A concrete application of our results can be found in robotic navigation. Instead of searching for the globally shortest path from source to target with the classical Dijkstra or the more sophisticated A^* algorithm, greedy routing with our landmark distance function is completely local and efficiently determines which edge to follow in each step, so as to eventually arrive at the target. In fact, after precomputing the geodesic distances from each vertex of T to the vertices of P , the landmark distance (2) between any vertex of T and the target can be determined in $O(m)$ time, and greedy routing simply follows the edge to the neighbor with the smallest landmark distance. In contrast to previous work in this direction [CGH18a, CGH18c] our theory indicates that these paths converge to the globally shortest geodesic paths as the resolution of the triangulation and the density of vertices on the domain boundary increase.

2. Previous Work

A number of coordinate-based distance functions have been proposed, in particular for discrete graphs, which model many real-world networks. In this scenario, each graph vertex is assigned a vector of coordinates, or *embedding*, carefully designed to support *greedy routing* on the graph, namely, that for each vertex there exists a neighbor which decreases some distance in embedding space to any target vertex. These are sometimes called *greedy embeddings*. While a celebrated result due to Bose and Morin [BM04] shows that the simple Euclidean coordinates support greedy routing on Delaunay triangulations, this does not generalize to arbitrary plane graphs, especially those discretizing a non-convex domain.

In the continuous case, given a polygon P , it may be possible

to move from $s \in P$ towards $t \in P$ by following a direction that decreases the Euclidean distance to t . Obviously, such a direction always exists if s is an interior point and, in this case, the decrease is maximized in the (negative) gradient direction of the Euclidean distance. It also exists in many cases if s is a boundary point, typically by following the boundary of P . But there will be boundary points at which no such direction exists within P and the routing procedure will get stuck at this local minimum. Subregions of P within which this type of greedy routing is possible (independent of the target point t) are called *greedy-routable regions* (GRR) and can be identified, and a routing procedure must be able to transition between the regions when stuck. Optimal decomposition of polygons into a minimal number of GRR’s (which may be non-convex) is a difficult problem which has been treated by Tan and Kermarrec [TK12] and by Nöllenburg et al. [NPR17]

Going back to the discrete case of graphs, the landmark distance (2) is particularly attractive as each (graph-theoretic) landmark distance function λ_i may be viewed as a precomputed “coordinate” of a graph vertex, and the landmark distance as the l_∞ (or max) norm on those coordinates. The distance is used to guide the routing process, essentially by trying to reduce this distance to a given target vertex t when moving from a vertex to one of its neighbors. Keeping the number of coordinates small is crucial for effective storage of this “embedding” of the graph. While we study the l_∞ norm of the landmark functions, other norms, in particular the l_1 and l_2 norms, have been considered [FGG*05, FRZ*05] for greedy routing on graphs. In those papers, landmarks are typically chosen as random vertices in the graph. Unfortunately, these distance functions do not support greedy routing, in the sense that local minima exist, and the router can get stuck there. Despite a number of “fixes”, there does not seem to be a complete understanding of how to make this work.

The landmark distance has also been used [Cho05, GH05] as a heuristic function in the context of the A^* algorithm for shortest-path computation on graphs. This is a well-known admissible and consistent heuristic and sometimes quite effective if the landmarks are chosen judiciously.

Our work considers the *continuous* version of the landmark distance within a simply connected planar polygonal domain P . In a nutshell, each point $x \in P$ is assigned the coordinate vector $\mathbf{g}(x) = (g_1(x), \dots, g_m(x))$ of geodesic distances to the vertices of the polygon. We will show that this suffices to generate a steepest descent path from any source point to any target point, albeit the path will not be the geodesic path, that is, the shortest possible. However, as more “virtual” vertices are added along the polygon boundary, the coordinate vectors become longer but the path becomes shorter. In the limit, each point is assigned a continuous boundary function and the steepest descent path converges to the geodesic path. This is similar to the continuous “divergence distance” coordinates [CGH18c] of a point associated with the planar polygonal boundary, which also support the generation of a steepest descent path from source to target. However, those coordinates are derived from conformal theory and generate paths of a hyperbolic nature, which are not geodesic paths. The continuous coordinate vectors may be discretized to “reduced divergence coordinates” [CGH18a] containing a finite number of coordinates, at the

expense of further distortion of the resulting gradient descent paths. It seems that these coordinates support steepest descent routing also in multiply connected polygonal domains [CGH18b].

A concept similar in spirit to landmark routing, called “beacon routing”, was introduced by Biro et al. [BIKM13]. In this method, a finite number of “beacon” points are positioned within a polygon and routing proceeds towards the closest beacon until a boundary is reached or the Euclidean distance to the beacon is identical to that of the target from the beacon. If a boundary is reached, it is followed as long as the distance to the beacon decreases. At that point the beacon is replaced with another one and the process continues until the target is reached. The main challenge is, given a polygon, to position a finite number of beacons in the polygon which can support beacon routing between any two points without getting stuck.

The main goal of this paper is to show that the simple and natural choice of using the vertices of P as landmarks, coupled with the l_∞ landmark distance, avoids all the problems outlined above and provides a simple method to route between any two points within a simple polygon.

3. Preliminaries

3.1. The Geodesic Distance

The fundamental distance function that we rely on to define landmark distances in a polygon P is the geodesic distance, which is the length of the geodesic curve between two points $x, y \in P$. The geodesic curve $\gamma(x, y)$ is the shortest path between two points within P , which is well-known to be a *polyline* between x and y , bending at a subset of the reflex (i.e., non-convex) vertices of P , and relatively easy to compute based on the *visibility graph* of P [dBCvKO08]. The last vertex of P on $\gamma(x, y)$ before y is called the *anchor* vertex of $\gamma(x, y)$. Obviously, if P is convex, then $\gamma(x, y)$ is the straight line between x and y and its anchor is x , otherwise the anchor is a reflex vertex. The shortest-path tree from any point x to the vertices of P has leaves at a subset of the vertices, called the *extreme* vertices of P relative to x . If z lies on $\gamma(x, y)$, then we say that x, y , and z are *co-geodesic*. Denote by $g(x, y) = |\gamma(x, y)|$ the *geodesic distance* between x and y , and by $e(y, x)$ and $e(x, y)$ the geodesic straight-line extensions of $\gamma(x, y)$ beyond x and y towards the boundary ∂P of P (see Figure 1), which may degenerate to a single point, if x or y are boundary points.

Geodesic distances within polygons have been studied extensively. We state here some basic but useful properties.

Proposition 1 (see [Aro89, Note 3.4]) The geodesic distance $g(x, y)$ is a metric and continuous in both x and y .

Proposition 2 (see [Aro89, Note 3.12]) The geodesic distance $g(x, y)$ is a continuously differentiable function of y , and its gradient with respect to y is the unit vector in the direction of $e(x, y)$, that is, $\nabla g(x, y) = (y - a) / \|y - a\|$, where a is the anchor of $\gamma(x, y)$.

Lemma 1 (see [PSR89, Lemma 1]) Let $a, b, c \in P$. Then, as x varies along $\gamma(b, c)$, $g(a, x)$ is a convex function of $g(b, x)$ and $g(a, x) \leq \max\{g(a, b), g(a, c)\}$.

Corollary 1 If $a \in P$, then $g(a, x)$ is a convex function of x within P .

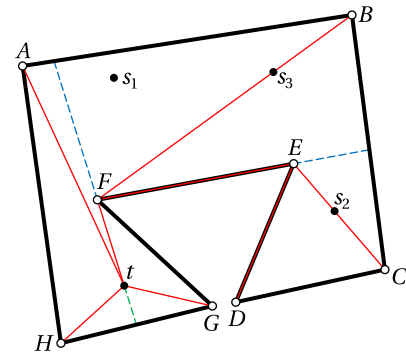


Figure 1: Shortest-path tree (red) from target point t to vertices of a non-convex polygon P . Here, the tree leaves are the convex vertices of P . The other vertices (E and F) are reflex vertices. Vertex F is the anchor vertex for B and vertex E is the anchor for C and D . Points t, F, B and points t, F, E, C are co-geodesic, thus $\mathcal{L}_t(s_3) = g(s_3, t)$, co-determined by landmarks F and B , and $\mathcal{L}_t(s_2) = g(s_2, t)$, co-determined by landmarks F, E , and C . However, $\mathcal{L}_t(s_1) < g(s_1, t)$ and is determined by landmark G . Dashed lines mark the geodesic extensions $e(t, E)$ and $e(t, F)$ (blue) and $e(E, t) = e(F, t)$ (green).

Proof A bivariate function over P is convex, if and only if it is convex on any line segment in P . Let b and c be two points in P , such that the line segment $[b, c]$ is also in P . This line segment is therefore also the shortest path $\gamma(b, c)$. Lemma 1 implies that $g(a, x)$ is a convex function of $g(b, x)$, which is a uniform parameterization of $[b, c]$. So, $g(a, x)$ is convex over $[b, c]$, hence also over all P . \square

3.2. Generalized Hyperbolas

The landmark distance is closely related to the definition of hyperbolas with respect to the geodesic distance instead of the usual Euclidean distance. To this end, let $s, t \in P$ be fixed and consider the function

$$D(x) = g(x, s) - g(x, t)$$

for $x \in P$. Obviously, D is continuous in P , and it turns out that D is bounded and has the following properties.

Proposition 3 Let $\Delta = g(s, t)$. Then $-\Delta \leq D(x) \leq \Delta$ for any $x \in P$. Moreover, $D(x) = -\Delta$, if $x \in e(t, s)$ and $D(x) = \Delta$, if $x \in e(s, t)$. In particular, $D(s) = -\Delta$ and $D(t) = \Delta$, and D increases linearly by arc length along $\gamma(s, t)$.

Proof By the triangle inequality, $g(x, t) \leq g(x, s) + g(s, t)$, which is equivalent to the lower bound $-\Delta \leq D(x)$, with equality if and only if s lies on $\gamma(x, t)$, which includes the case $x \in e(t, s)$. The upper bound follows similarly from $g(x, s) \leq g(x, t) + g(t, s)$. Finally, if $x \in \gamma(s, t)$ splits $\gamma(s, t)$ in the ratio $\lambda : 1 - \lambda$ for some $\lambda \in [0, 1]$, then it follows from the subpath property of shortest paths that $g(x, s) = \lambda g(s, t)$ and $g(x, t) = (1 - \lambda)g(s, t)$, hence $D(x) = (2\lambda - 1)\Delta$. \square

We denote the contour line of D for iso-value $\delta \in [-\Delta, \Delta]$ by

$$B_\delta = \{x \in P : D(x) = \delta\}.$$

The contour line B_0 contains all points that are equidistant from s

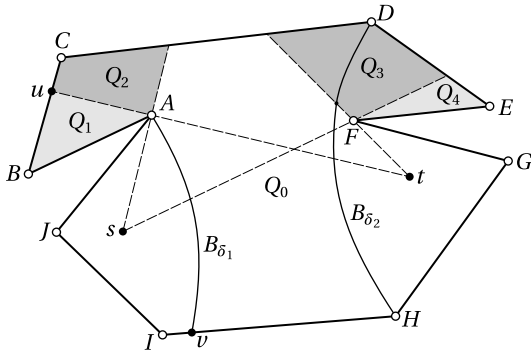


Figure 2: Partitioning of a non-convex polygon P into regions Q_0, \dots, Q_5 with anchor vertices $\hat{s}_1 = \hat{t}_1 = \hat{s}_2 = A, \hat{t}_3 = \hat{s}_4 = \hat{t}_4 = F, \hat{s}_0 = \hat{s}_3 = s, \text{ and } \hat{t}_0 = \hat{t}_2 = t$. The δ -bisector B_{δ_1} connects the regular endpoint v with the degenerate endpoint A and contains the entire shadow Q_1 of A , including the convex vertex B . The δ -bisector B_{δ_2} has two regular endpoints and consists of two hyperbolic arcs, one in the region Q_0 with foci s and t , the other in the region Q_3 with foci s and F .

and t and is called the *bisector* of s and t . The bisector is a straight line if P is convex, and for general simple polygons, we recall a result from [Aro89, Lemma 3.22].

Lemma 2 If s and t are in *general position*, namely, B_0 does not contain a vertex of P , then B_0 is a smooth curve connecting two points on ∂P and having no other points in common with ∂P . It is the concatenation of $O(n)$ straight lines and hyperbolic arcs.

The proof of this lemma uses a partition of P into polygonal regions, where each region Q contains two vertices \hat{s} and \hat{t} , such that $g(s, x) = g(s, \hat{s}) + \|\hat{s} - x\|$ and $g(t, x) = g(t, \hat{t}) + \|\hat{t} - x\|$ for all $x \in Q$ (see Figure 2). Apart from special cases that may occur at the boundary of Q , this means that \hat{s} and \hat{t} are the common anchors of the geodesic paths from s and t to all points in Q . The restriction of B_0 to Q is either empty, the straight-line bisector of \hat{s} and \hat{t} , or a hyperbolic arc with foci \hat{s} and \hat{t} . If the two anchor vertices coincide for some Q and $\hat{s} = \hat{t}$ is equidistant from s and t , then s and t are not in general position and B_0 contains the entire region Q .

For $\delta \neq 0$, the contour line B_δ , which we call the δ -bisector of s and t , behaves similarly. If P is convex, then B_δ is a hyperbolic arc with foci s and t , and essentially a piecewise hyperbolic curve, otherwise.

Lemma 3 If s and t are in *general position* with respect to $\delta \in (-\Delta, \Delta)$, namely, B_δ does not contain a vertex of P , then B_δ is a smooth curve connecting two points on ∂P and having no other points in common with ∂P . It is the concatenation of $O(n)$ straight lines and hyperbolic arcs.

Proof As in the proof of Lemma 2, let Q be a region from the partition of P with anchor vertices $\hat{s} \neq \hat{t}$. The condition $x \in B_\delta \cap Q$ is then equivalent to $\|\hat{s} - x\| - \|\hat{t} - x\| = \delta - g(s, \hat{s}) + g(t, \hat{t})$, which implies that $B_\delta \cap Q$ is either empty or a hyperbolic arc with foci \hat{s} and \hat{t} , which degenerates to a straight line, if and only if $\delta = g(s, \hat{s}) - g(t, \hat{t})$. The rest of the proof is exactly as in [Aro89]. \square

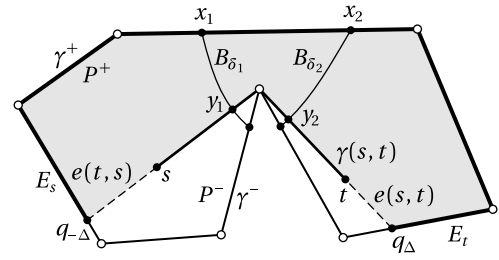


Figure 3: Notation used in the proof of Theorem 1.

As for the bisector B_0 , it can happen that the δ -bisector B_δ contains an entire region Q , if the two anchor vertices of Q coincide and $\hat{s} = \hat{t} \in B_\delta$ (see Figure 2). Such a region, which can also be the union of several regions Q , may occur at either end of B_δ , in the shadow of a reflex vertex p of P (with respect to s and t), defined as the set of all points $x \in P$, for which $\gamma(x, p) = \gamma(x, s) \cap \gamma(x, t)$. In this case, we call p a *degenerate endpoint*. Apart from this degeneracy, B_δ is always as described in Lemma 3 (see Figure 4).

The special case $|\delta| = \Delta$ is not covered by Lemma 3, but it follows from Proposition 3, that the contour line $B_{-\Delta}$ is the straight line segment $e(t, s)$, which connects s with some point $q_{-\Delta} \in \partial P$, and B_Δ is the straight line segment $e(s, t)$, connecting t with some point $q_\Delta \in \partial P$. As before, these boundary points may be degenerate.

We now partition ∂P at $q_{-\Delta}$ and q_Δ into two polygonal paths γ^+ and γ^- (see Figure 3), and we analyze the behavior of $D(x)$ on ∂P .

Theorem 1 The function D increases monotonically along the paths γ^+ and γ^- from $q_{-\Delta}$ to q_Δ .

Proof As D is continuous, with $D(q_{-\Delta}) = -\Delta$ and $D(q_\Delta) = \Delta$, it follows from the intermediate value theorem that there exist for any $\delta \in (-\Delta, \Delta)$ two boundary points $q_\delta^+ \in \gamma^+$ and $q_\delta^- \in \gamma^-$ with $D(q_\delta^+) = D(q_\delta^-) = \delta$. By Lemma 3, these must be the endpoints (or lie in the shadow of possibly degenerate endpoints) of the contour line B_δ . Moreover, it follows from Proposition 3, that there exists a unique point $q_\delta \in \gamma(s, t) \cap B_\delta$ with $D(q_\delta) = \delta$.

We now partition P by the extended geodesic $e(t, s) \cup \gamma(s, t) \cup e(s, t) = \gamma(q_{-\Delta}, q_\Delta)$ through s and t into two regions and focus on the region P^+ that contains γ^+ . Let $x_1, x_2 \in \gamma^+$ be any two points ordered in the sense that x_1 is encountered first when travelling from $q_{-\Delta}$ to q_Δ along γ^+ , and assume that $\delta_1 = D(x_1) > D(x_2) = \delta_2$. Further let $y_1 = q_{\delta_1}$ and $y_2 = q_{\delta_2}$ be the unique points on $\gamma(s, t)$ with $D(y_1) = \delta_1$ and $D(y_2) = \delta_2$. By the linearity of D along $\gamma(s, t)$, the point y_2 is encountered first when travelling from s to t along $\gamma(s, t)$. This, however, implies that the contour line B_{δ_1} , which connects x_1 and y_1 , and the contour line B_{δ_2} from x_2 to y_2 must intersect, which is impossible. By contradiction, we must therefore have $\delta_1 \leq \delta_2$, which asserts that D is monotonically increasing along γ^+ . The same argument works for γ^- . \square

In analogy to the case when the geodesic distance is identical to the Euclidean distance, we call

$$H_\delta = B_{-\delta} \cup B_\delta = \{x \in P : |g(x, s) - g(x, t)| = \delta\}$$

for $\delta \in (0, \Delta)$ a (generalized) *hyperbola* with foci s and t and semi-major axis $\delta/2$. We further refer to $B_{-\delta}$ and B_δ as the *s-branch* and

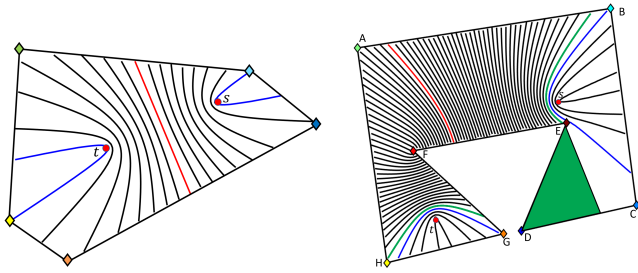


Figure 4: Hyperbolas H_δ with foci s and t in a convex (left) and a non-convex polygon (right). The branches of H_δ are the contour lines $B_{-\delta}$ and B_δ of D . The bisector B_0 is shown in red and the tightest hyperbola in blue. The green hyperbola has a degenerate endpoint (E) and contains all points in the shadow of E (green).

the t -branch of H_δ , because they bend around s and t , respectively. The vertices of H_δ are the two points $q_{-\delta}$ and q_δ where the two branches intersect $\gamma(s, t)$. As $\delta \rightarrow 0$, both branches converge to the bisector B_0 , and we get the degenerate hyperbola $H_0 = B_0$. As $\delta \rightarrow \Delta$, the two arms of $B_{-\delta}$ and of B_δ both converge to $e(t, s)$ and $e(s, t)$, respectively, and we obtain the degenerate hyperbola $H_\Delta = e(t, s) \cup e(s, t)$ with infinite curvature at its vertices s and t . Consequently, we say that H_{δ_1} is tighter than H_{δ_2} , if $\delta_1 > \delta_2$.

We are now ready to establish the connection between the landmark distance and generalized hyperbolas.

Proposition 4 Let $\delta = \mathcal{L}_t(s)$ be the landmark distance between t and s . Then H_δ is the tightest hyperbola that contains at least one landmark and $\mathcal{L}_t(s)$ is co-determined by all landmarks on H_δ .

Proof By the definition of the landmark function, the landmark ℓ_i lies on the hyperbola H_{δ_i} with $\delta_i = \lambda_i(s)$, and the statement then follows directly from the definition of the landmark distance. \square

Proposition 4 gives rise to a geometric procedure for determining the landmark distance between t and s and the landmarks that co-determine $\mathcal{L}_t(s)$. Start with the degenerate hyperbola H_Δ for $\delta = \Delta$ and “open” it by continuously decreasing δ until H_δ “hits” a landmark. This and all other landmarks that are “hit” by this tightest hyperbola H_δ , including landmarks that lie in the shadow of a degenerate endpoint, co-determine the landmark distance at s (see Figure 4), and its value $\mathcal{L}_t(s) = \delta$ is the geodesic distance between the vertices of H_δ .

4. The Landmark Distance

4.1. Basic Properties

Using the landmark distance \mathcal{L}_t in (2) for routing from some source $s \in P$ to the target $t \in P$ by steepest descent is possible only if \mathcal{L}_t has a global minimum at the target and no other local minima in P . If the landmarks are the vertices of P , that is, $m = n \geq 3$ and $\ell_i = p_i$ for $i = 1, \dots, n$, then the first condition is not hard to show.

Proposition 5 The landmark distance has a unique global minimum at t .

Proof By definition, $\mathcal{L}_t(x) \geq \mathcal{L}_t(t) = 0$. Suppose there exists some

$x \neq t$ with $\mathcal{L}_t(x) = 0$. Then, again by definition, $\lambda_i(x) = 0$, hence $\ell_i \in H_0$ for $i = 1, \dots, n$. But this contradicts Lemma 2, by which the bisector $B_0 = H_0$ contains at most two landmarks. If B_0 has degenerate endpoints, then there might be more co-determining landmarks in the shadow of those endpoints, but never all of them. \square

For this choice of landmarks, the landmark distance turns out to induce a metric on P , which justifies calling \mathcal{L}_t a distance function.

Theorem 2 The bivariate function $\mathcal{L}(x, y) = \mathcal{L}_y(x)$ is a metric.

Proof By Proposition 5, $\mathcal{L}(x, y) = 0 \Leftrightarrow x = y$, and the symmetry $\mathcal{L}(x, y) = \mathcal{L}(y, x)$ follows directly from the definition. To show the triangle inequality, assume that $\mathcal{L}_y(x)$ is determined by landmark ℓ_i , $\mathcal{L}_z(x)$ by ℓ_j , and $\mathcal{L}_y(z)$ by ℓ_k . Then,

$$\begin{aligned} \mathcal{L}(x, y) &= |g_i(x) - g_i(y)| \\ &\leq |g_i(x) - g_i(z)| + |g_i(z) - g_i(y)| \\ &\leq |g_j(x) - g_j(z)| + |g_k(z) - g_k(y)| = \mathcal{L}(x, z) + \mathcal{L}(z, y). \end{aligned}$$

\square

Note that Proposition 5 and Theorem 2 are not true for arbitrarily chosen landmarks, especially if they are not constrained to lie on the boundary of P , as it may then happen that all landmarks lie on the bisector B_0 of some s and t with $s \neq t$ and then $\mathcal{L}_t(s) = 0$.

The following three observations, however, are true for any choice of landmarks. We first show that the landmark distance is a lower bound on the geodesic distance.

Lemma 4 For any $x \in P$,

$$\mathcal{L}_t(x) \leq g(x, t),$$

with equality if and only if $\ell_i \in e(x, t) \cup e(t, x)$ for some i .

Proof By the triangle inequality,

$$\begin{aligned} g_i(x) &\leq g(x, t) + g_i(t) \Rightarrow g_i(x) - g_i(t) \leq g(x, t), \\ g_i(t) &\leq g(t, x) + g_i(x) \Rightarrow g_i(t) - g_i(x) \leq g(t, x), \end{aligned}$$

hence $\lambda_i(x) = |g_i(x) - g_i(t)| \leq g(x, t)$, with equality if either $\ell_i \in e(x, t)$ or $\ell_i \in e(t, x)$. Since this is true for all $i = 1, \dots, m$, the statement follows from the definition of $\mathcal{L}_t(x)$ in (2). \square

An immediate consequence of Lemma 4 is that the landmark distance \mathcal{L}_t at the landmark ℓ_i is determined by ℓ_i itself.

Corollary 2 For any landmark ℓ_i ,

$$\mathcal{L}_t(\ell_i) = \lambda_i(\ell_i) = g(\ell_i, t).$$

Moreover, the landmark distance satisfies a maximum principle on the geodesic between any two landmarks.

Lemma 5 For any landmarks ℓ_i and ℓ_j and any $x \in \gamma(\ell_i, \ell_j)$,

$$\mathcal{L}_t(x) \leq \max\{\mathcal{L}_t(\ell_i), \mathcal{L}_t(\ell_j)\}.$$

Proof To prove the statement, note that

$$\mathcal{L}_t(x) \leq g(x, t) \leq \max\{g_i(t), g_j(t)\} = \max\{\mathcal{L}_t(\ell_i), \mathcal{L}_t(\ell_j)\},$$

where the first inequality follows from Lemma 4, the second from Lemma 1, and the last equality from Corollary 2. \square

A more general maximum principle holds if all landmarks are located on the boundary of P , but not necessarily at the vertices.

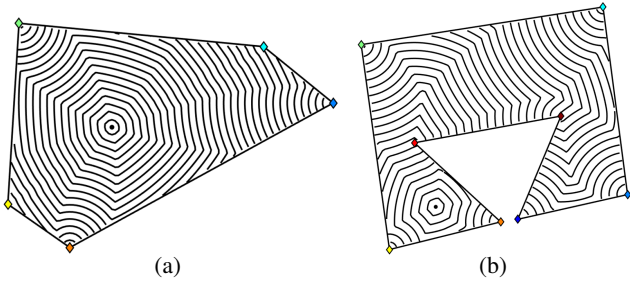


Figure 5: Contours of the landmark distance from t (black point).

Theorem 3 For any region $O \subset P$ and any $x \in O$,

$$\mathcal{L}_t(x) \leq \max\{\mathcal{L}_t(y) : y \in \partial O\}.$$

Proof Assume that $\mathcal{L}_t(x)$ is determined by ℓ_i , and consider the geodesic $\gamma(\ell_i, x)$. Since $\ell_i \in \partial P$ and $x \in O$, this geodesic must intersect ∂O at some y_1 . Likewise, let y_2 be the intersection of ∂O with the extension $e(\ell_i, x)$ of $\gamma(\ell_i, x)$. By Lemma 1, $g_i(z)$ increases from 0 as z varies along $\gamma(\ell_i, x) \cup e(\ell_i, x)$, hence $g_i(y_1) \leq g_i(x) \leq g_i(y_2)$. We now distinguish two cases. If $g_i(x) \leq g_i(t)$, then

$$\mathcal{L}_t(x) = g_i(t) - g_i(x) \leq g_i(t) - g_i(y_1) = \lambda_i(y_1) \leq \mathcal{L}_t(y_1).$$

Otherwise, if $g_i(x) > g_i(t)$, then

$$\mathcal{L}_t(x) = g_i(x) - g_i(t) \leq g_i(y_2) - g_i(t) = \lambda_i(y_2) \leq \mathcal{L}_t(y_2).$$

In both cases the statement follows, because $y_1, y_2 \in \partial O$. \square

The global maximum of \mathcal{L}_t can then be identified easily by using Corollary 2, Theorem 3, and Lemma 5.

Corollary 3 The global maximum of the landmark distance is obtained at the landmarks ℓ_i that are farthest from t , by geodesic distance, and equal to the distance of these landmarks from t ,

$$\max_{x \in P} \mathcal{L}_t(x) = \mathcal{L}_t(\ell_i) = g(\ell_i, t) \iff i \in \arg \max_{j=1, \dots, m} g(\ell_j, t).$$

Going back to the case where the landmarks are the vertices of P , it further follows that the global maximum of \mathcal{L}_t is obtained at one (or several) of the *convex* vertices of P . Often, but not always, local maxima occur at the other vertices (see Figure 5).

4.2. Absence of Local Minima

We now turn to the central result of this paper, namely that the landmark distance \mathcal{L}_t has no local minima inside the polygon P , except at the target t , in the case when the landmarks are the vertices of P . The key insight for proving this property is that for any point $s \in \text{Int} P \setminus \{t\}$, the number of *different* (negative) gradients of the landmark functions corresponding to the landmarks that co-determine $\mathcal{L}_t(s)$ is at most four, even though there can be more than four such landmarks, and that they all point in a common direction.

Recalling the definition of the landmark function $\lambda_i(x)$ in (1), it follows directly from Proposition 2 that its gradient is the unit vector in the direction of the extension $e(\ell_i, x)$ of $\gamma(\ell_i, x)$.

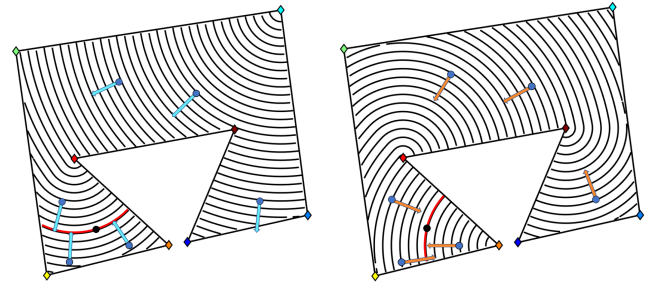


Figure 6: Contours and examples of (negative) gradients of the landmark functions for the cyan and orange landmarks of the polygon in Figure 5(b). The zero contour passing through t is marked in red.

Corollary 4 For any landmark ℓ_i and any $x \in P \setminus \{\ell_i\}$, let $u = (x - a) / \|x - a\|$, where a is the anchor of $\gamma(\ell_i, x)$. Then (see Figure 6),

$$\nabla \lambda_i(x) = \begin{cases} u, & \text{if } g_i(x) > g_i(t), \\ -u, & \text{if } g_i(x) < g_i(t). \end{cases}$$

If $g_i(x) = g_i(t)$, then $\lambda_i(x) = 0$ and its gradient is undefined, and likewise at $x = \ell_i$, where λ_i has a local maximum.

We now observe that the co-determining landmarks of $\mathcal{L}_t(s)$ are in very specific places. To this end, let E_s and E_t be the edges of P that intersect with the extended geodesic between s and t , that is, $q_{-\Delta} \in E_s$ and $q_{\Delta} \in E_t$ (see Figure 3). If $q_{-\Delta}$ happens to be a vertex of P , then E_s is defined as the degenerate edge that consists of only this vertex, and likewise for q_{Δ} and E_t .

Theorem 4 If the landmark ℓ_i co-determines $\mathcal{L}_t(s)$, then ℓ_i is an endpoint of E_s or E_t , or ℓ_i lies in the shadow of one of these endpoints. In the latter case, the corresponding edge endpoint is itself a co-determining landmark ℓ_j of $\mathcal{L}_t(s)$ and a degenerate endpoint of the hyperbola H_{δ} for $\delta = \mathcal{L}_t(s)$. Moreover, $\nabla \lambda_i(s) = \nabla \lambda_j(s)$.

Proof It follows from Lemma 3 that for any $\delta \in (-\Delta, \Delta)$, the hyperbola H_{δ} has exactly four (possibly degenerate) endpoints on ∂P . By Theorem 1, these four points move monotonically along ∂P as δ decreases from Δ to 0, two starting at $q_{-\Delta}$, moving along γ^+ and γ^- , and likewise for the other two starting at q_{Δ} , until they meet at q_0^+ and q_0^- , the endpoints of the bisector B_0 . Following the geometric interpretation of Proposition 4, the opening process will therefore terminate as soon as an endpoint of E_s or E_t is “hit”. Hence, these endpoints are the only candidates for co-determining landmarks of $\mathcal{L}_t(s)$, with one exception: if a co-determining landmark ℓ_j is a degenerate endpoint of the hyperbola H_{δ} for $\delta = \mathcal{L}_t(s)$, then all landmarks in the shadow of ℓ_j are co-determining landmarks, too. In fact, if ℓ_i is one of these landmarks, then $g_i(s) = g_i(\ell_j) + g_j(s)$ and $g_i(t) = g_i(\ell_j) + g_j(t)$, because ℓ_j is a point on both $\gamma(\ell_i, s)$ and $\gamma(\ell_i, t)$, and therefore $\lambda_i(s) = \lambda_j(s)$. Moreover, $e(\ell_i, s) = e(\ell_j, s)$, which implies $\nabla \lambda_i(s) = \nabla \lambda_j(s)$ by Corollary 4. \square

For $\delta = \mathcal{L}_t(s)$, let a_1 and a_2 be the anchors of the paths $\gamma(q_{-\delta}^-, s)$ and $\gamma(q_{-\delta}^+, s)$ from the endpoints $q_{-\delta}^-, q_{-\delta}^+ \in E_s$ of the s -branch of the tightest hyperbola H_{δ} to s . Similarly, let a_3 and a_4 be the anchors of the paths from the endpoints $q_{\delta}^+, q_{\delta}^- \in E_t$ of the t -branch

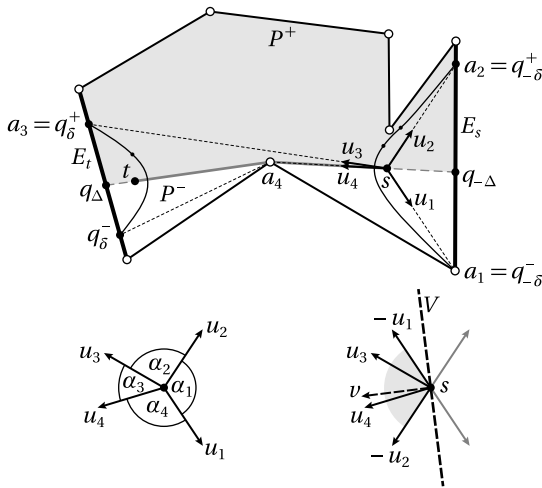


Figure 7: Notation used in the proof of Lemma 6.

of H_δ to s . Moreover, let $u_j = (a_j - s) / \|a_j - s\|$ for $j = 1, \dots, 4$ (see Figure 7). Corollary 4 and Theorem 4 then imply that $-\nabla \lambda_i(s) \in \{-u_1, -u_2, u_3, u_4\}$ for all co-determining landmarks ℓ_i , and it turns out that these four vectors all point in a common direction.

Lemma 6 There exists a vector $v \in \mathbb{R}^2$, such that $v \cdot u_j < 0$ for $j = 1, 2$ and $v \cdot u_j > 0$ for $j = 3, 4$.

Proof First note that the u_j , since they point in the direction of the first segment of the geodesics from s to the corresponding endpoints of H_δ , have the same order by angle around s as those endpoints along ∂P , which, by construction, is the order given by their indices. Next consider the angles $\alpha_j = \angle(u_j, u_{j+1})$ for $j = 1, \dots, 4$, where u_5 is identified with u_1 . Since $q_{-\delta}^+$ and q_δ^+ lie in P^+ , the region “above” $\gamma(q_{-\Delta}, q_\Delta)$, then so do the geodesics between s and these endpoints, and in particular a_2 and a_3 . As the boundary of P^+ is locally straight at s , we conclude that $\alpha_2 < \pi$, and likewise $\alpha_4 < \pi$. To see that $\alpha_1 < \pi$, imagine first the case where $s = q_{-\Delta}$, so that $\alpha_1 = \pi$. Then observe that α_1 decreases first strictly, then monotonically, as we move s along $\gamma(q_{-\Delta}, q_\Delta)$ to its actual position, and a similar reasoning shows that $\alpha_3 < \pi$.

Since all $\alpha_j < \pi$, there exists some line V through s that separates u_1 and u_2 from u_3 and u_4 (see Figure 7). The unit vector v perpendicular to V , in the direction of u_3 and u_4 then satisfies the conditions in the statement. \square

Theorem 5 The landmark distance $\mathcal{L}_t(x)$ does not have any local minima in $\text{Int}P$, except at $x = t$.

Proof We have already established in Proposition 5 a unique global minimum of 0 at $x = t$. For any other $x \in \text{Int}P$, we know from Lemma 6 that there exists a common direction v , in which all landmark functions λ_i that co-determine $\mathcal{L}_t(x)$ decrease. By construction, \mathcal{L}_t must therefore also decrease in this direction and hence cannot have a local minimum at x . \square

In Section 5, we state the actual direction of steepest descent of the landmark distance \mathcal{L}_t , but let us first take a closer look at the behavior of \mathcal{L}_t on the boundary of P .

Lemma 7 The landmark distance \mathcal{L}_t is convex along any edge of P .

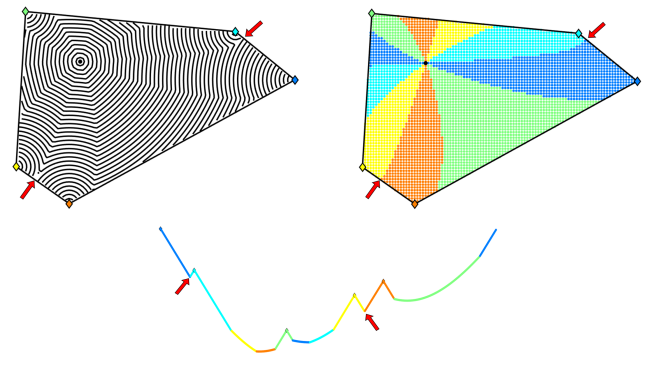


Figure 8: Landmark distance from t (black point) for a convex polygon: contours (top left), color-coded by the determining landmark (top right), and behavior on the boundary with the same color-coding (bottom). Note the local maximum at and the piecewise-linear behavior in the neighborhood of each landmark, and the two points (marked by red arrows), where two such regions meet. At these points, the landmark distance has a saddle and decreases (only) in the inward direction perpendicular to the edge.

Proof Let x_1 and x_2 be two points on an edge E of the polygon, and let $s = \alpha x_1 + (1 - \alpha)x_2$ for some $\alpha \in (0, 1)$ be a point on the same edge between them. Assume that the landmark distance at s is determined by landmark ℓ_k , that is,

$$\mathcal{L}_t(s) = |g_k(s) - g_k(t)| \geq |g_i(s) - g_i(t)|, \quad i = 1, \dots, n,$$

and distinguish two cases. On the one hand, if $g_k(s) \geq g_k(t)$, then, by the convexity of g_k (see Lemma 1),

$$\begin{aligned} \mathcal{L}_t(s) &= g_k(s) - g_k(t) = g_k(\alpha x_1 + (1 - \alpha)x_2) - g_k(t) \\ &\leq \alpha g_k(x_1) + (1 - \alpha)g_k(x_2) - \alpha g_k(t) - (1 - \alpha)g_k(t) \\ &\leq \alpha |g_k(x_1) - g_k(t)| + (1 - \alpha) |g_k(x_2) - g_k(t)| \\ &\leq \alpha \mathcal{L}_t(x_1) + (1 - \alpha) \mathcal{L}_t(x_2). \end{aligned}$$

On the other hand, if $g_k(s) < g_k(t)$, then, following the arguments in the proof of Theorem 4, ℓ_k is one of the endpoints of $E = E_s$, and

$$\begin{aligned} \mathcal{L}_t(s) &= g_k(t) - g_k(s) = g_k(t) - \alpha g_k(x_1) - (1 - \alpha)g_k(x_2) \\ &\leq \alpha |g_k(t) - g_k(x_1)| + (1 - \alpha) |g_k(t) - g_k(x_2)| \\ &\leq \alpha \mathcal{L}_t(x_1) + (1 - \alpha) \mathcal{L}_t(x_2), \end{aligned}$$

where the second identity follows because g_k is linear on E . \square

Figure 8 provides an example of this boundary behaviour and also shows the linear behavior that \mathcal{L}_t inherits from λ_i in the vicinity of each ℓ_i along the boundary. Note that the proof of Lemma 6 also extends to the open edges of P , with one exception.

If $s \in \text{Int}E$ for some edge $E = [\ell_1, \ell_2]$ of P , it may happen that the landmark distance $\mathcal{L}_t(s)$ is co-determined (at least) by the two endpoints of E . In that case, the two corresponding landmark functions λ_1 and λ_2 have opposing gradient vectors, hence $\alpha_1 = \pi$. Taking as v the unit vector in the inward direction perpendicular to E , we get $v \cdot u_j = 0$ for $j = 1, 2$ and $v \cdot u_j > 0$ for $j = 3, 4$. As λ_1 and λ_2 decrease linearly in the distance to ℓ_1 and ℓ_2 , locally at s , this suffices to conclude that the λ_i of all co-determining landmarks ℓ_i , and

therefore also \mathcal{L}_t , decrease in the direction v . In fact, each such s is a saddle point of \mathcal{L}_t , but not a local minimum (see Figure 8).

If $s = \ell_i$ is a vertex of P , then $\mathcal{L}_t(s)$ is co-determined by ℓ_i itself (by Corollary 2), but $\nabla\lambda_i$ is undefined at s (by Corollary 4) and the reasoning in the proof of Lemma 6 cannot be applied. However, it follows from Lemma 4 that $\mathcal{L}_t(x) = \lambda_i(x) = g(x, t)$ for any $x \in \gamma(s, t)$, so the landmark distance is guaranteed to decrease in the direction v pointing along $\gamma(s, t)$. Overall, we conclude that $\mathcal{L}_t(x)$ is free of local minima in P , except for the global minimum at $x = t$.

5. Geometry of the Landmark Partition

For a given target point $t \in P$, it is useful to partition the interior of the polygon as follows. For any $x \in P$, let $I_x = \{i : \mathcal{L}_t(x) = \lambda_i(x)\}$ be the index set of all co-determining landmarks and denote by $\eta(x) = \#\{\nabla\lambda_i(x) : i \in I_x\}$ the number of well-defined and different gradients of the corresponding landmark functions. By Corollary 4, $\eta(x) = 0$ for $x = t$ and also for $x = \ell_i$, unless ℓ_i and t happen to be co-geodesic with some other landmark ℓ_j . By Theorem 4, $\eta(x) \leq 4$, and we call x a *double*, *triple*, or *quadruple* point, if $\eta(x)$ is 2, 3, or 4.

A region of the *landmark partition* is then defined as the closure of the set of all points x with a common index set I_x and $\eta(x) = 1$. Note that in some special cases (e.g., when t is co-geodesic with two landmarks), a region may be degenerate, namely, have no interior. For a non-degenerate region R , Theorem 4 implies that the co-determining landmarks at any $x \in \text{Int}R$ (i.e., the ℓ_i with $i \in I_x$) are the vertices of a common subtree T of the shortest-path trees rooted at s and t . We label R by the index of the landmark at the root of T and call the latter the *common anchor landmark* of all co-determining landmarks.

By definition, \mathcal{L}_t has a continuous gradient at any $x \in P$ with $\eta(x) = 1$, in particular at the interior points of each region. More precisely, $\nabla\mathcal{L}_t(x) = \nabla\lambda_i(x)$ for any $i \in I_x$. Although $\nabla\mathcal{L}_t(x)$ is undefined at all points $x \in P$ with $\eta(x) \neq 1$, we know from Section 4.2, that a direction of steepest descent does exist, except at $x = t$. In practice, a path generation routine will “probe” in a small neighborhood (e.g., a circle) around x to find the point/direction which decreases the landmark distance the most.

We now provide some observations related to the geometry of the landmark partition and the structure of the steepest descent paths.

5.1. Convex Polygons

Let us first study the basic case when P is a *convex* polygon. In this case, the geodesic distance $g(x, y)$ is equal to the Euclidean distance $\|x - y\|$. Denote $r_i = \|t - \ell_i\|$ for $i = 1, \dots, n$. See Figure 9 for a detailed example of the convex case.

Each region of the landmark partition is determined by a single landmark. The boundary between adjacent regions are double points s at which the distance is co-determined by two landmarks ℓ_i and ℓ_j satisfying $\|s - \ell_i\| - r_i = \|\|s - \ell_j\| - r_j\|$. There are four possible cases, depending on the relationship between the distances of s and t to ℓ_i and ℓ_j . We distinguish between two of these cases, and the other two are essentially the same as these: 1) s is closer to

ℓ_i and to ℓ_j than t is, so $\|s - \ell_i\| - \|s - \ell_j\| = r_i - r_j$; 2) s is further from ℓ_i than t , but closer to ℓ_j than t , so $\|s - \ell_i\| + \|s - \ell_j\| = r_i + r_j$. In the first case s and t lie on a branch of the *hyperbola* whose foci are ℓ_i and ℓ_j with semi-major axis length $|r_i - r_j|/2$ and in the second case s and t lie on the *ellipse* whose foci are ℓ_i and ℓ_j with semi-major axis length $(r_i + r_j)/2$. The triple points s at the intersection of three regions (co-determined by three landmarks) are at the intersection of two of the ellipses and one of the hyperbolas mentioned above. They are also the centers of *Apollonius circles* [Wik21] defined by the set of circles centered at ℓ_i with radii r_i , namely circles tangent (either interior or exterior) to these three circles. It is easy to see that a triangle contains no triple points for any target $t \in P$, and that if s is a triple point for a target t , then t is a triple point for s . Quadruple points are those at the intersection of four regions and co-determined by four landmarks. These are at the intersection of four ellipses and two hyperbolas and are very rare, usually occurring only when the polygon possesses some symmetry, for example, a regular polygon (see Figure 10).

A path of steepest descent from s to t proceeds in a straight line from s away from the determining landmark of $\mathcal{L}_t(s)$ until a boundary between regions is encountered. Then it continues along region boundaries until t is reached. Note that since the boundaries are conics, the direction of steepest descent, which is tangent to this boundary, is well-known to be the bisector of the angle between this point and the two foci of the conic (see Figure 9(f)). It is easy to see that this is also the bisector of the angle between the two extreme (of the up to four) gradients defined at this point.

If s is one of the vertices of P , say $s = \ell_i$, and $\mathcal{L}_t(s)$ is co-determined not only by ℓ_i , but also by another landmark ℓ_j , then ℓ_j must be the endpoint q_Δ of the extended geodesic from s to t , by Proposition 4 and Theorem 4, that is, s , t , and ℓ_j are co-geodesic. In this case, the unique direction of steepest descent points towards t , and the steepest descent path is identical to the geodesic path $\gamma(s, t)$. If ℓ_i is the only determining landmark of $\mathcal{L}_t(s)$, then any inward direction can be taken as the initial direction, because \mathcal{L}_t inherits the local maximum at s from λ_i in this case and decreases with the same rate in all directions. This includes the globally best direction towards t that yields $\gamma(s, t)$ as the steepest descent path, but a local probing procedure is not able to find it without being given additional information.

5.2. Non-convex Polygons

The geometry of the landmark partition for non-convex polygons is more complicated. See Figure 11 for a detailed example of this case. Each region may be co-determined by *multiple* landmarks, but the corresponding landmark functions all have the *same* gradient.

As in the convex case, the partition region boundaries are conics, and the double points s are located on the boundary between two adjacent regions, on an ellipse or hyperbola that depends on t and the co-determining landmarks of these regions. However, unlike in the convex case, the foci of these conics are not necessarily the co-determining landmarks themselves, but more generally their anchor landmarks, or even the anchors of the geodesics from these anchor landmarks to s , and the semi-major axis length depends on the geodesic distances between t , the anchor landmarks, and the

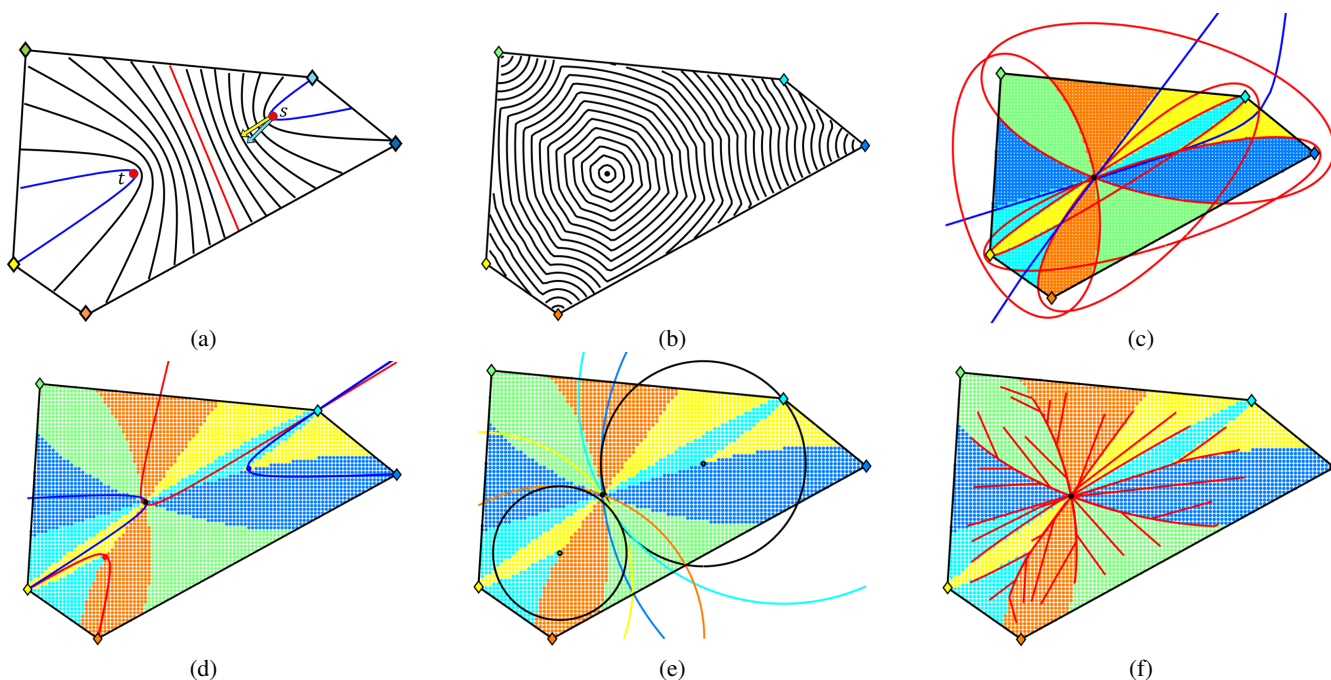


Figure 9: (a) Convex polygon P and hyperbolic δ -bisectors between two (red) interior points s and t . The red curve is the 0-bisector. The tightest blue hyperbola passes through two polygon vertices, which are the two co-determining landmarks of $\mathcal{L}_t(s)$. The two colored vectors are the negative gradients of the corresponding landmark functions at s , one away from the cyan vertex and one towards the yellow vertex, thus s is a double point. (b) Contours of the landmark distance from t (black point). The global maximum is obtained at the landmark farthest from t and the global minimum at t . There are no local minima in P . (c) The five landmarks are color-coded and interior points $s \in P$ are colored by the determining landmark for $\mathcal{L}_t(s)$, thus partitioning the interior of P . Boundaries between the regions are (red) ellipses or (blue) hyperbola with foci at two landmarks. The two triple points where three regions meet are the intersection of two ellipses and one hyperbola. (d) The tightest hyperbola with foci at t and one of the two triple points intersects the three co-determining landmarks. (e) The circles centered at each landmark and intersecting t . The triple points are at the centers of Apollonius circles tangent to the three circles centered at the landmarks defining the triple point. (f) Sample steepest descent paths to the target t . These are straight lines until a region boundary is reached, at which point only boundaries are followed to t . At any double point s , the bisector of the negative gradients of the landmark functions associated with the two co-determining landmarks is exactly the tangent to the ellipse or hyperbola defining the boundary. For triple points, the tangent to the hyperbola is always between the two tangents of the two ellipses.

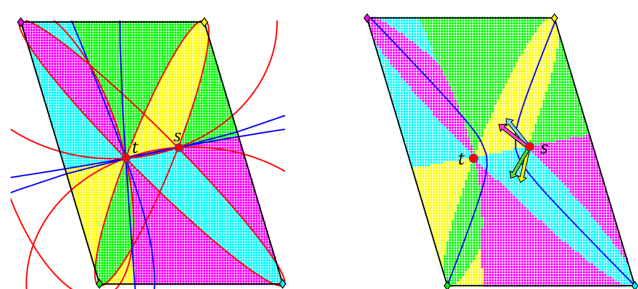


Figure 10: A quadruple point s involves four ellipses and two hyperbolas (left). The tightest hyperbola passes through four landmarks, whose landmark functions induce four aligned negative gradients (right).

landmarks in common. In this case, the boundary is a degenerate hyperbola and thus a straight line segment. However, the points on these boundaries are not double points, because the gradients inherited from both regions are identical.

The characterization of triple points carries over similarly, and they are located at the intersection of two ellipses and a hyperbola. A “virtual” t , which is the second intersection of these three conics, may be associated with this triple point (in the convex case this is exactly t), and similarly for quadruple points.

As before, a path of steepest descent from s to t begins with a straight line segment and then follows the region boundaries towards t . This path is uniquely determined, unless it starts at or passes through a vertex of P at which \mathcal{L}_t is co-determined by a single anchor landmark. At such points, similar to the convex case, \mathcal{L}_t decreases with the same rate in any direction towards the interior of the region associated with this common anchor landmark.

geodesic anchors. Moreover, for non-convex polygons it may happen that two adjacent regions have two or more co-determining

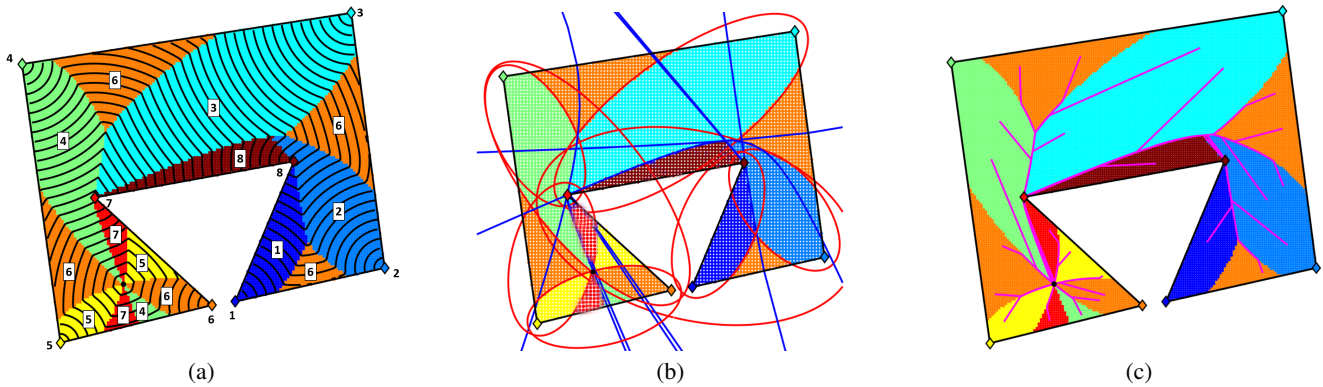


Figure 11: (a) Non-convex polygon P with contours of landmark distance from t (black point), partitioned, labeled, and color-coded into regions according to the common anchor landmark of the co-determining landmarks of the region. The global maximum is obtained at a convex vertex and the global minimum at t . There are no local minima in P . The red region associated with reflex vertex number 7 is co-determined also by vertices 8, 1, 2, and 3, and the brown region associated with reflex vertex number 8 is co-determined also by vertices 1 and 2. (b) Double points are on the boundaries between two regions, which are (red) ellipses or (blue) hyperbola with foci at two anchor landmarks. The triple points, where three regions meet, are the intersection of two ellipses and one hyperbola. The boundary between regions 8 and 2 is a degenerate hyperbola with foci at vertices 8 and 2 and does not consist of double points. (c) Sample steepest descent paths to the target t . These are straight lines until a region boundary is reached, at which point only boundaries are followed to t . These paths were generated numerically by steepest descent, thus they may have slight precision issues. For the paths passing through vertex 8, our implementation chose the “right-most” of all possible continuing directions out of vertex 8 and towards the interior of region 8.

6. Discussion

6.1. Fewer Landmarks

In the previous section we saw that some landmarks may play a more significant role than others, for example, those at the leaves of shortest path trees to the polygon vertices, which always includes the subset of convex vertices. This begs the question as to whether there is a subset of the polygon vertices which, when used as landmarks, reproduces the same landmark distance function as when all the polygon vertices are used. A related question is whether diluting the set of vertices to a subset (e.g., just the convex vertices) perhaps does not exactly reproduce the full landmark distance, but still yields a distance function which is free of local minima. We suspect that in many cases this is true, and, if not, local minima will be present only on the boundary of the polygon.

Figure 12 compares the contours of the landmark distance from two target points in the “man” polygon when all polygon vertices are used as landmarks vs. when only the convex vertices are used. We observe that when some of the reflex vertices are extreme relative to the target (i.e., are leaves of the shortest path tree from the target to the vertices), then the contours are different (see right leg), albeit still with no local minima. If none of the reflex vertices are extreme, it seems that their removal does not change the landmark distance.

6.2. More Landmarks

Our work studies the landmark distance \mathcal{L}_t in the case when the landmarks are positioned at all the polygon vertices. But what happens for other positionings of the landmarks, for example, if there are additional landmarks in the interior of the polygon or on the edges?

It is easy to see that Theorem 5 is not necessarily true when P has interior landmarks, as there may be local maxima of \mathcal{L}_t at the interior landmarks. Typically, each local maximum introduces an additional local minimum and two saddle points (see Figure 13).

However, if we stay with landmarks on the edges of the polygon, the situation only improves if more and more landmarks are added (see Figure 14), as the following theorem predicts.

Theorem 6 If P is a simple polygon and the landmarks are dense along the boundary of P , then \mathcal{L}_t delivers for all $t \in P$ and the steepest descent path reproduces $\gamma(s, t)$.

Proof Given $s, t \in P$, consider the extensions $e(t, s)$ and $e(s, t)$ of $\gamma(s, t)$ towards $q_{-\Delta} \in \partial P$ and $q_{\Delta} \in \partial P$ (cf. Section 3.2). If there happens to be a landmark ℓ_i at $q_{-\Delta}$ or q_{Δ} then, by Lemma 4, $\mathcal{L}_t(s) = g(s, t)$, so \mathcal{L}_t will deliver and reproduce $\gamma(s, t)$. This will happen as more landmarks are added along the edges of P . \square

6.3. Multiply Connected Polygons

Our results have been obtained for simply connected polygons. It would be nice if they could be generalized to multiply connected polygonal domains, that is, those with “holes”. Alas, experiments show that local minima cannot be avoided in this case, even at the limit where landmarks are placed densely along all edges of the polygon. The fundamental reason seems to be that a geodesic between two interior points cannot always be extended to the boundaries of the polygon while preserving the geodesic property, or alternatively, the geodesic diameter of such a polygon is not always determined by two vertices of the polygon [BKO13]. The top row of Figure 15 shows some examples of local minima of the landmark

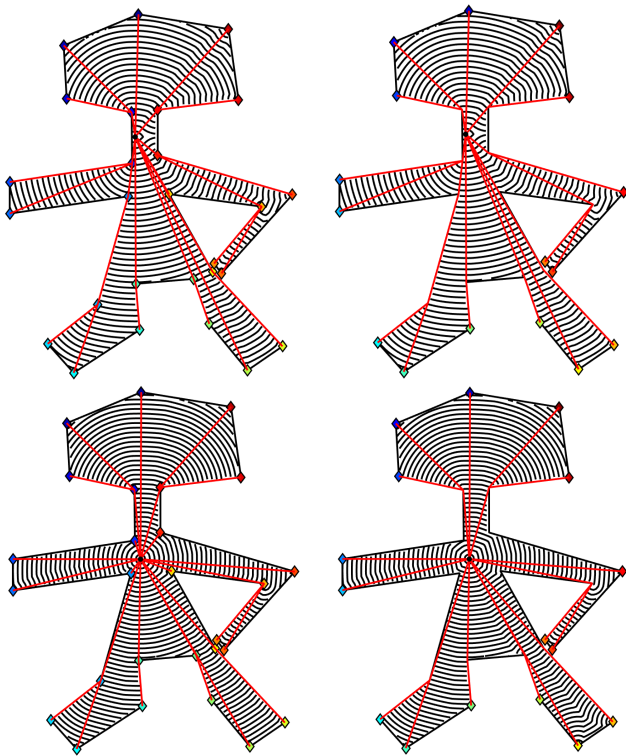


Figure 12: Comparison of the landmark distance when using all 27 vertices (left) of the non-convex “man” polygon or just the subset of 16 convex vertices (right) as landmarks. If the (red) shortest path tree to the polygon vertices rooted at t (black point) has no leaves at the reflex vertices, then the landmark distance seems to be unaffected by their removal (top). If the shortest path tree rooted at t has leaves at some of the reflex vertices (shoulder and armpits), then their removal may change the landmark distance, which can be seen in the right leg (bottom).

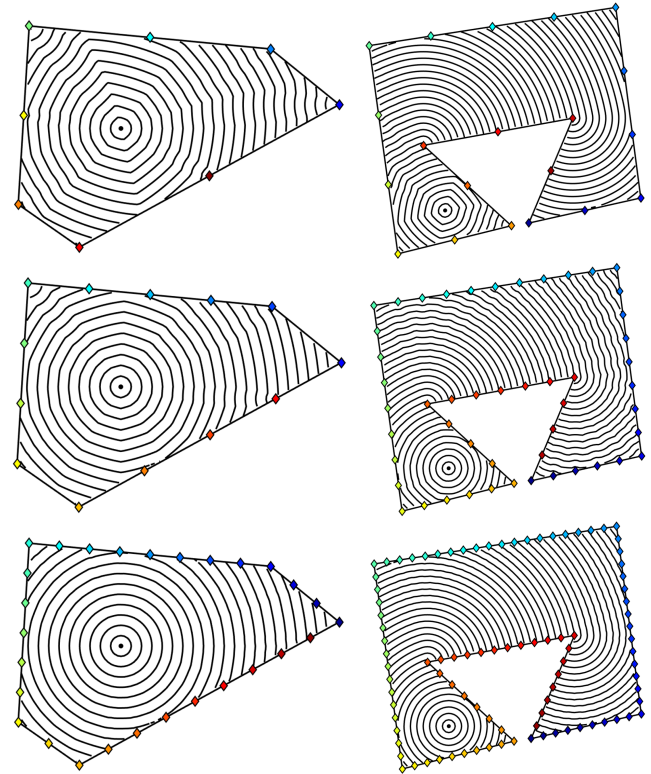


Figure 14: Adding successively more landmarks along the polygon edges (from top to bottom) makes the landmark distance closer to the geodesic distance.

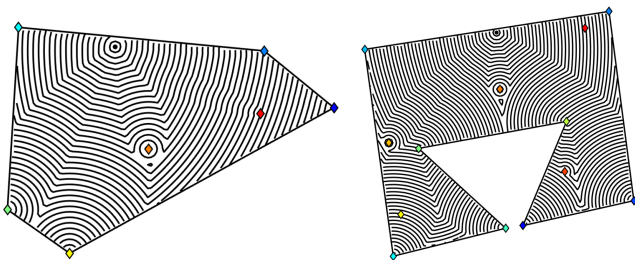


Figure 13: Interior landmarks may introduce local minima and maxima into the landmark distance.

distance in two multiply connected polygons. However, this problem seems to be mitigated somewhat as more landmarks are positioned on the boundaries, in which case the resulting landmark distance, as in the simply connected case, approaches the true geodesic distance. This is demonstrated in the middle and bottom rows of Figure 15.

6.4. 2-Manifold Surfaces in 3D

We have demonstrated that the landmark distance may be an excellent approximation for the geodesic distance in simple planar polygons, and even in multiply connected polygons, by positioning landmarks densely along the boundaries. A natural question is whether this generalizes to 2-manifolds with boundaries in three dimensions. These types of questions have been treated by mathematicians interested in solving “inverse problems” where the metric structure of a manifold with boundary is to be derived from information relating the boundary to the interior. This seems to be possible under suitable conditions on the surface (e.g., that its curvature is bounded) [KKL07], so there is good reason to believe that our method may be applied also to manifolds.

Efficiently computing geodesic distances and their contours over 3D surfaces has been the focus of much attention in the geometry processing community (see, e.g., [KS98, SSK⁺05, CWW13] and the survey [CLPQ20]). The objective of recent work [CWW13] is to preprocess the surface (or some discretization of it), so that “queries” for the geodesic contours from arbitrary target points may be computed very quickly based on that preprocessing. Our work may follow this approach, where once a surface is discretized into, say, a triangle mesh, the geodesic distances of each interior vertex to all boundary vertices are precomputed and stored. Assuming N vertices in the mesh, the boundary will typically con-

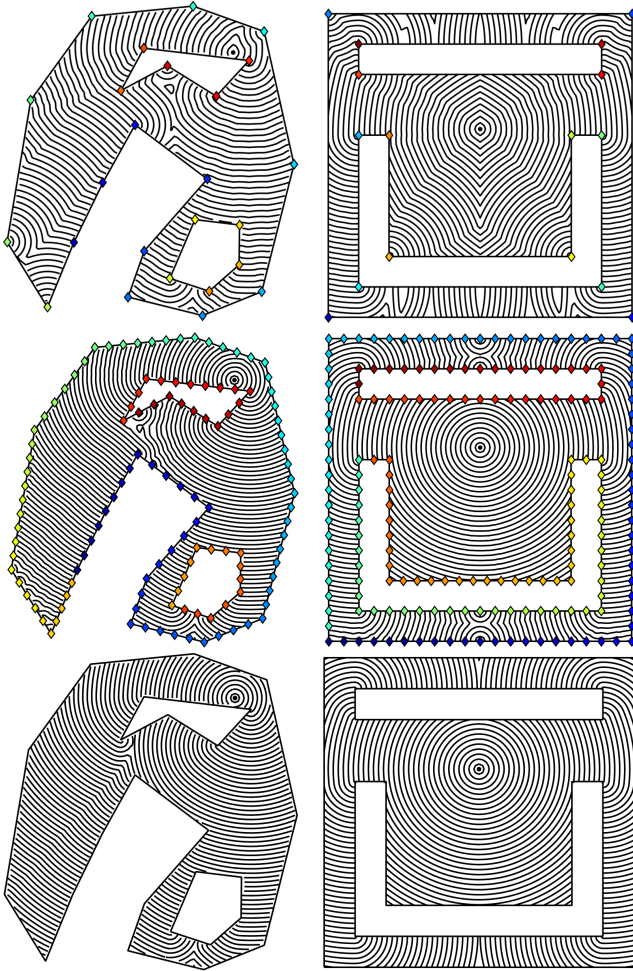


Figure 15: Landmark distance on multiply connected polygons. With landmarks on the polygon vertices, the landmark distance from the black point is not necessarily free of local maxima or minima (top). Adding landmarks along the boundary (middle) brings the landmark distance closer to the true geodesic distance (bottom), although there are a few inaccuracies where geodesic “fronts” meet.

tain $n = O(\sqrt{N})$ vertices, so the size of the resulting data structure would be $O(N\sqrt{N})$. Computing all landmark distances as an approximation to the geodesic distances in a naive serial manner will cost $O(\sqrt{N})$ time per triangulation vertex for the max operator in (2), in total $O(N\sqrt{N})$ for the entire mesh, which is quite expensive compared to other methods. However, the computation is “embarrassingly parallel”, meaning, given enough processors, the task can be trivially parallelized down to almost constant time, something which is far from obvious for the competing methods. Moreover, the landmark distance allows us to query a single vertex-to-vertex distance in $O(\sqrt{N})$ time, while some other methods, like [CWW13], can only compute the distance between a vertex and all other vertices at the same time. Beyond that, it may be possible to reduce the $O(\sqrt{N})$ time per mesh vertex by a more

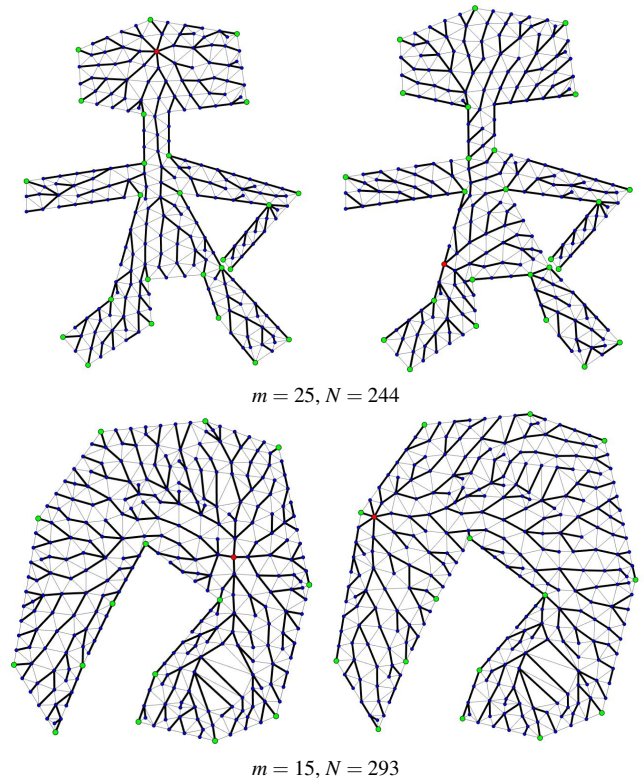


Figure 16: Landmark distance routing trees for two (red) target vertices in a triangulation of the interior of a non-convex polygon with N (blue) vertices and m (green) landmarks. We have checked (exhaustively) that the landmark distance delivers to all possible target vertices of these triangulations.

efficient search procedure to quickly identify the boundary vertex that determines the landmark distance for that vertex. We leave this very interesting direction of using our results to efficiently compute geodesic distances on surfaces for future work.

6.5. The Discrete Case

Given that the continuous landmark distance \mathcal{L}_t is void of local minima at the interior of a simple polygon P , it is natural to ask what happens when the interior of P is discretized by a triangulation T , and routing is attempted along the edges of this plane graph. Will the landmark distance, now formulated using the (weighted) graph theoretic distance between vertices of the graph in lieu of the continuous geodesic distance, support greedy routing, namely guarantee delivery? In other words, for each vertex v of T , will there always exist a neighboring vertex w of v such that $\mathcal{L}_t(w) < \mathcal{L}_t(v)$? It seems that for a general triangulation the answer is no. However, we have experimentally observed that for “well-behaved” triangulations, namely a “nice” (uniform) distribution of the triangulation vertices with a “nice” triangle structure (e.g., constrained Delaunay), seems to always result in delivery (see Figure 16), without the need to augment T with additional edges, as in [CGH18a]. However, proving this observation requires more investigation, as does

the question of the resulting *stretch* of the generated paths, namely, the ratio between their lengths and the length of the true shortest path. Ideally, this should not be too large. In fact, in the examples of Figure 16, the stretch was no more than 1.03 on average.

References

- [Aro89] ARONOV B.: On the geodesic Voronoi diagram of point sites in a simple polygon. *Algorithmica* 4, 1 (1989), 109–140. doi:10.1007/BF01553882. 3, 4
- [BIKM13] BIRO M., IWERKS J., KOSTITSYNA I., MITCHELL J. S. B.: Beacon-based algorithms for geometric routing. In *Algorithms and Data Structures*, Dehne F., Solis-Oba R., Sack J.-R., (Eds.), vol. 8037 of *Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, 2013, pp. 158–169. doi:10.1007/978-3-642-40104-6_14. 3
- [BKO13] BAE S. W., KORMAN M., OKAMATO Y.: The geodesic diameter of polygonal domains. *Discrete & Computational Geometry* 50, 2 (Sept. 2013), 306–329. doi:10.1007/s00454-013-9527-8. 10
- [BM04] BOSE P., MORIN P.: Online routing in triangulations. *SIAM Journal on Computing* 33, 4 (Aug. 2004), 937–951. doi:10.1137/S0097539700369387. 2
- [CGH18a] CHEN R., GOTSMAN C., HORMANN K.: Efficient path generation with reduced coordinates. *Computer Graphics Forum* 37, 5 (Aug. 2018), 37–48. Proceedings of SGP. doi:10.1111/cgfm.13489. 2, 12
- [CGH18b] CHEN R., GOTSMAN C., HORMANN K.: On divergence-based distance functions for multiply-connected domains, Jan. 2018. arXiv:1801.07099. 3
- [CGH18c] CHEN R., GOTSMAN C., HORMANN K.: Path planning with divergence-based distance functions. *Computer Aided Geometric Design* 66 (Nov. 2018), 52–74. doi:10.1016/j.cagd.2018.09.002. 2
- [Cho05] CHOW E.: *A graph search heuristic for shortest distance paths*. Tech. Rep. UCRL-CONF-210878, Lawrence Livermore National Laboratory, Mar. 2005. URL: <https://www.osti.gov/biblio/862393>. 2
- [CLPQ20] CRANE K., LIVESU M., PUPPO E., QIN Y.: A survey of algorithms for geodesic paths and distances, July 2020. arXiv:2007.10430. 11
- [CWW13] CRANE K., WEISCHEDEL C., WARDETZKY M.: Geodesics in heat: A new approach to computing distance based on heat flow. *ACM Transactions on Graphics* 32, 5 (Sept. 2013), Article 152, 11 pages. doi:10.1145/2516971.2516977. 11, 12
- [dBCvKO08] DE BERG M., CHEONG O., VAN KREVELD M., OVERMARS M.: *Computational Geometry: Algorithms and Applications*, 3rd ed. Springer, Berlin, 2008. doi:10.1007/978-3-540-77974-2. 3
- [FGG*05] FANG Q., GAO J., GUIBAS L. J., DE SILVA V., ZHANG L.: GLIDER: Gradient landmark-based distributed routing for sensor networks. In *Proceedings IEEE Infocom 2005* (Miami, Mar. 2005), Makki K., Knightly E., (Eds.), vol. 1, IEEE, pp. 339–350. doi:10.1109/INFCOM.2005.1497904. 2
- [FRZ*05] FONSECA R., RATNASAMY S., ZHAO J., EE C. T., CULLER D., SHENKER S., STOICA I.: Beacon vector routing: Scalable point-to-point routing in wireless sensor networks. In *Proceedings of the 2nd Symposium on Networked Systems Design & Implementation* (Boston, May 2005), NSDI'05, USENIX Association, pp. 329–342. 2
- [GH05] GOLDBERG A. V., HARRELSON C.: Computing the shortest path: A* search meets graph theory. In *Proceedings of the Sixteenth Annual ACM–SIAM Symposium on Discrete Algorithms* (Vancouver, Jan. 2005), SODA'05, Society for Industrial and Applied Mathematics, pp. 156–165. 2
- [KKL07] KATSUDA A., KURYLEV Y., LASSAS M.: Stability of boundary distance representation and reconstruction of Riemannian manifolds. *Inverse Problems & Imaging* 1, 1 (Feb. 2007), 135–157. doi:10.3934/ipi.2007.1.135. 11
- [KS98] KIMMEL R., SETHIAN J. A.: Computing geodesic paths on manifolds. *Proceedings of the National Academy of Sciences* 95, 15 (July 1998), 8431–8435. doi:10.1073/pnas.95.15.8431. 11
- [NPR17] NÖLLENBURG M., PRUTKIN R., RUTTER I.: Partitioning graph drawings and triangulated simple polygons into greedily routable regions. *International Journal of Computational Geometry & Applications* 27, 1 & 2 (Mar. & June 2017), 121–158. doi:10.1142/S0218195917600068. 2
- [PSR89] POLLACK R., SHARIR M., ROTE G.: Computing the geodesic center of a simple polygon. *Discrete & Computational Geometry* 4, 6 (Dec. 1989), 616–626. doi:10.1007/BF02187751. 3
- [SBD*13] SOUSSI O., BENATITALLAH R., DUVIVIER D., ARTIBA A., BELANGER N., FEYZEAU P.: Path planning: A 2013 survey. In *Proceedings of the 5th International Conference on Industrial Engineering and Systems Management* (Rabat, Oct. 2013), IESM'13, IEEE, pp. 849–856. 1
- [SSK*05] SURAZHSKY V., SURAZHSKY T., KIRSANOV D., GORTLER S. J., HOPPE H.: Fast exact and approximate geodesics on meshes. *ACM Transactions on Graphics* 24, 3 (July 2005), 553–560. Proceedings of SIGGRAPH. doi:10.1145/1073204.1073228. 11
- [TK12] TAN G., KERMARREC A.-M.: Greedy geographic routing in large-scale sensor networks: A minimum network decomposition approach. *IEEE/ACM Transactions on Networking* 20, 3 (June 2012), 864–877. doi:10.1109/TNET.2011.2167758. 2
- [Wik21] WIKIPEDIA: Circles of Apollonius. https://en.wikipedia.org/wiki/Circles_of_Apollonius, 2021. [Online; accessed 15-April-2021]. 8