MINIMUM PATH IN A EUCLIDEAN SPACE

JAMES C HATELEY

Abstract. Given a finite number of points and a fixed starting point in $\mathbb{R}^2$, or $\mathbb{R}^3$, a minimum or minimal distance between a finite number of points is desired for many applications in the sciences. Using the property that a respective sphere is classified by its surface, linear maps and geodesic curves can be used to order a set of points to contain the minimum or a minimal distance.

1. Introduction

In this paper only the case of $\mathbb{R}^2$ will be examined thoroughly. $\mathbb{R}^3$ will be addressed, and cases greater than $\mathbb{R}^3$ need a more rigorous treatment that will not be provided here. The following algorithm produces a minimal distance, with insight on how to obtain the minimum path between a given set of points. A short explanation of this algorithm with notation is introduced before the case of $\mathbb{R}^2$.

The shortest distance between a given finite set of points (denoted S) in a respective space is a curve tracing some edges of some polyhedron with the vertices as a subset of S. In order to find a minimal path, the arithmetic mean of the points is taken (μ_x). From this mean the maximum distance (ρ) is computed from the set of points and (μ_x). A sphere is drawn around the set with radius ρ. Each element of S is then mapped (τ) to the boundary of the constructed sphere. Since each sphere of concern is classified by its surface, a uniform expansion map (which will be defined) on the surface will hit every element in the image of τ. τ Might not necessarily be bijective, so uniqueness of a curve will be dealt with in the spaces of $\mathbb{R}^2$ and $\mathbb{R}^3$. A new map (σ) will be defined by the extending the image of τ back into its pre-image. So $\sigma \circ \tau$ map elements from S to itself with re-indexing ($\overline{S}$) with respect to the expansion map on the surface of the sphere. A simple observation and a calculation will solve the problem of uniqueness. Then the set $\overline{S}$ will be a reordering of S with a minimal distance, if not the minimal distance.

2. Notation

For the purposes of this paper all general notation used will be defined here. A finite set of points in $\mathbb{R}^n$ will be defined by the set S.

$$S = \{x_i\}_{i=1}^{k} \in \mathbb{R}^n, x_i = (x_{i_1}, x_{i_2}, \dots x_{i_k})$$

The arithmetic mean of the set will be denoted as μ_x.

$$\mu_x := \sum_{i=1}^{k} \frac{x_i}{k}$$

The radius for a respective n-dimensional sphere is denoted by ρ and is defined as follows.

$$\rho := \max_{x_i \in \mathbb{R}^n} \|\mu_x - x_i\|$$

Where $\|.\|$ is the standard Euclidean distance. The above gives the ability to construct a minimum sized sphere where all of S is contained the closure of the sphere. The size of the sphere is not important, although it must contain all of S.

$$B(\mu_x, \rho) := (x_1 - \mu_{x_1})^2 + (x_2 - \mu_{x_2})^2 + \cdots + (x_n - \mu_{x_k}) = \rho^2$$

Finally, let the set T consist of points that are on the boundary of $B(\mu_x, \rho)$ and define a linear map τ such that $\tau : S \to T$.

$$\tau(x_i) = \rho \frac{x_i - \mu_x}{\|x_i - \mu_x\|}$$

So, $T = \{\tau(x_i)\}_{i=1}^n \in \partial B(\mu_x, \rho)$ It is easy to see that this map is surjective, but not necessarily injective. Not being injective creates the possibility that the curve being produced is not unique.

3. The case of $\mathbb{R}^2$

When τ is applied to the set S, each point is mapped to the boundary of a circle, this is commonly known as the circumference. Given a fixed starting point, and traversing the circle counter-clockwise until we return to our starting point we will have come across all points in the image of τ. Consider the pre-image of τ denoted by σ, $\sigma : T \to \overline{S}$

$$\sigma \circ \tau : \{x_i\}_{i=1}^n \to \{x_j\}_{j=1}^n$$

Note that if $i = j$ this does not imply that $x_i = x_j$. Without loss of generality suppose x_1 is our starting point.

$$(\sigma \circ \tau)\, x_1 = x_1$$

Denote the angle between μ_x and two points in T by α

$$\alpha = \cos^{-1}\left(\frac{\langle \tau(x_i), \tau(x_j)\rangle}{\rho^2}\right)$$

After a parameterization of $B(\mu_x, \rho)$ into polar coordinates, a minimal path can be calculated by the boundary of $B(\mu_x, \rho)$.

$$(\sigma \circ \tau)\, x_{i+1} = \left\{x_j : \min\left\{\int_0^\alpha \rho\, dt > 0\right\}\right\}$$

The map σ re-indexes T into $\overline{S}$ in a counter clockwise direction. A minimal curve, if not the minimum will be indexed in ascending order in $\overline{S}$. This might be a little much for going around the circumference of a circle but the intuition is similar for the more complicated case of $\mathbb{R}^3$.

4. Uniqueness of a curve in $\mathbb{R}^2$

If two points lie on the same radius ρ the map sigma is not well defined. Another restriction is needed to correct this. If $(\sigma \circ \tau)x_i$ is equal to $(\sigma \circ \tau)x_{i+1}$, then let $(\sigma \circ \tau)x_{i+1}$ equal to the closet element x_j of S such that $\tau(x_i)$ is equal to $\tau(x_j)$. This assignment assigns the next element down or up ρ of S to $\overline{S}$, guaranteeing that the map $(\sigma \circ \tau)$ is injective. There is one final concern, and that is when $(\sigma \circ \tau)x_i$ approaches such a line ρ containing 3 or more points. To resolve ρ in this case consider 3 points on ρ, or when $\tau(x_j) = \tau(x_k) = \tau(x_l)$. Without loss of generality suppose that $x_j < x_k < x_l$ on ρ starting from μ_x. If $(\sigma \circ \tau)x_i$ is equal to x_k then the curve will pass through x_k twice. To avoid this the minimum distance from set of $\tau(x_{i+1})$ needs to be calculated. The resolution to this problem is the following algorithm.

Definition 1. *A radius ρ is called complicated if it has two or more points of S on it from μ_x to the boundary of $B\,(\mu_x, \rho)$.*

Complicated ρ Algorithm. *Let the radii $\rho_i, \ldots, \rho_{i+k}$ be complicated, so $\tau(x_j) = \tau(x_k) = \cdots = \tau(x_l)$. Without loss of generality suppose that $x_j < x_k < \cdots < x_l$ on ρ starting from μ_x. If $(\sigma \circ \tau)x_i$ is not equal to x_j or x_l then the curve will pass through x_k twice. To avoid this, the minimum distance from set of $\{\tau(x_i), \cdots, \tau(x_{i+k+1})\}$ needs to be calculated. Define M and M_i in the following recursive manner.*

$$M_i = \min\{\|x_{i-1} - x_{\delta^i}\| + \|x_{\delta^i} - x_{\delta^i+1}\|\}$$
$$M = \{x_i : \min\{M_i + M_{i+1} + \cdots + M_{i+k+1}\}\}$$

Where δ^i is equal to either j or l, i is either equal to l or j, $x_{\delta^j} < x_{\delta^l}$ and x_{δ^j}, x_{δ^l} are the extremes on each respective ρ. Then let $(\sigma \circ \tau)x_i$ equal the chain of M_i of M such that the chain M_i connects all points of ρ_i.

So if there is one complicated ρ with 3 points and $M_i = \|x_{i-1} - x_j\| + \|x_l - x_{j+1}\|$, then $(\sigma\circ\tau)x_i = x_j$, $(\sigma\circ\tau)x_{i+1} = x_k$, $(\sigma\circ\tau)x_{i+2} = x_l$, $(\sigma\circ\tau)x_{i+3} = x_{j+1}$. When $\sigma\circ\tau$ approaches any radius ρ with two or more points this algorithm is implemented. With this construction a unique minimal curve is produced. If there are a series of radii ρ with multiple points, then the *complicated ρ algorithm* calculates the shortest distance. Notice that this algorithm is just a permutation algorithm. The computational speed can become lengthy if there is a large series of complicated ρ. In fact, the *complicated ρ algorithm* does 2^n calculations for a sequence of n complicated ρ.

Example 1.

$$S = \{(5,-5), (1,4), (-2,-1), (7,-5), (-4,2), (-8,6), (5,5), (-4,-6)\}$$

It is easily calculated that $\mu_x = (0,0)$ *and that* $\rho = 10$*, so* $B(\mu_x, \rho) = x^2 + y^2 = 100$*. Notice that* τ *is bijective and by construction, the point* $(-8,6) \in \partial B(\mu_x, \rho)$*.*

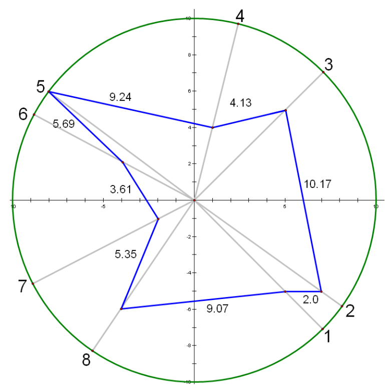

Figure 1

In this example an unambiguous path colored blue. For following Figures, when a smaller distance is available it will be colored by light blue and the longer distance will be colored brown.

Since τ is bijective, the map σ is well defined for *Figure 1*. From the starting point (labeled 1) the traversed path can easily be identified and the minimum distance can be calculated, which is approximately 49.26. This is nice, but what happens when two more points are added along radius two. Adding points can cause μ_x and ρ to change, but for purposes of this example suppose it does not. There are two important observations to make as points are added, one is that all points must be contained inside the $B(\mu_x, \rho)$.

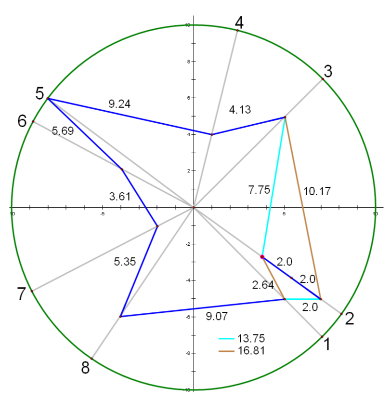

Figure 2

Observe two points are added onto ρ_2, the distance between each point is of length 2. Sigma determines the minimum distance after ρ and takes the value of the compliment of $\tau(x_2)$. After $(\sigma \circ \tau)x_2$ is chosen, so is the rest of the path. In this case, $\sigma \circ \tau$ traverses the minimum distance. What happens when the point on ρ_3 is moved? Originally the point was at $(5, 5)$. Observe how $\sigma \circ \tau$ acts on it when the point is moved to $(9, -1)$ then $(7, 1)$.

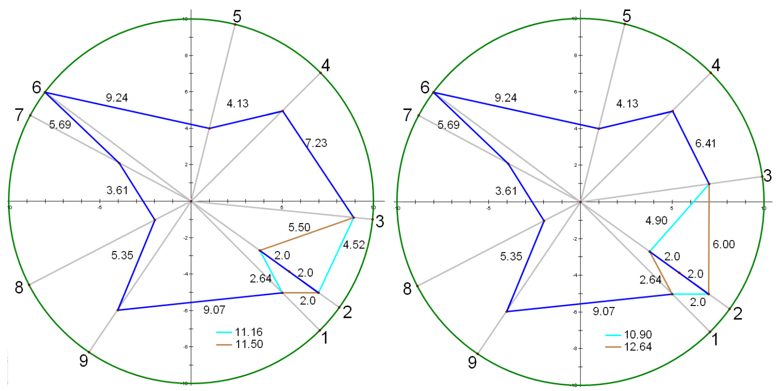

FIGURE 3

$\sigma \circ \tau$ in both cases have mapped the minimum distance. This is great, but expected. Take notice on how the path of $\sigma \circ \tau$ change as the minimum distance between ρ_2 and ρ_3 change. Now onward to a little more complicated situation. Another point on ρ_3 is added and the point on ρ_4 is moved. Observe how the change in ρ_4 effects the minimum curve, and the path of $\sigma \circ \tau$.

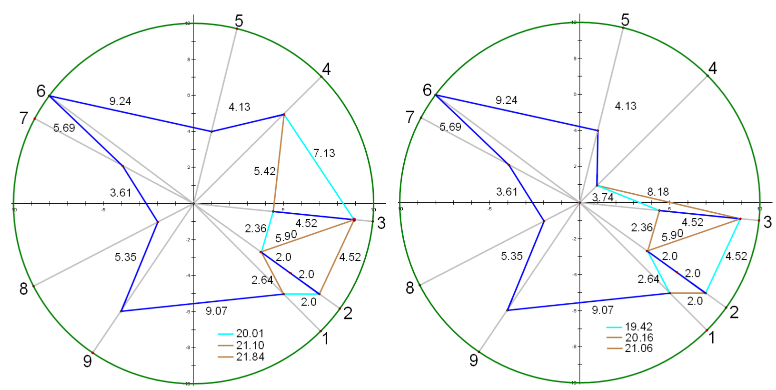

FIGURE 4

In both cases of *Figure 4* $\sigma \circ \tau$ calculates the minimum curve. It might not be obvious yet, but $\sigma \circ \tau$ does not always map the minimum curve. However, $\sigma \circ \tau$ will calculate the minimum distance between a series of complicated ρ. This is because when $\sigma \circ \tau$ approaches a complicated ρ, the *complicated ρ algorithm* is initialized. Before dismissing this example, keep in mind that a recalculation of μ_x and ρ should be made. Before recalculating the values of μ_x and ρ observe the following.

In the right part of *Figure 5* μ_x is translated down and length of ρ is increased. The image of τ is bijective, so there is only one path for $\sigma \circ \tau$ to map. $\sigma \circ \tau$ does not travel the minimum curve, but it is still a minimal distance. This is a quick way

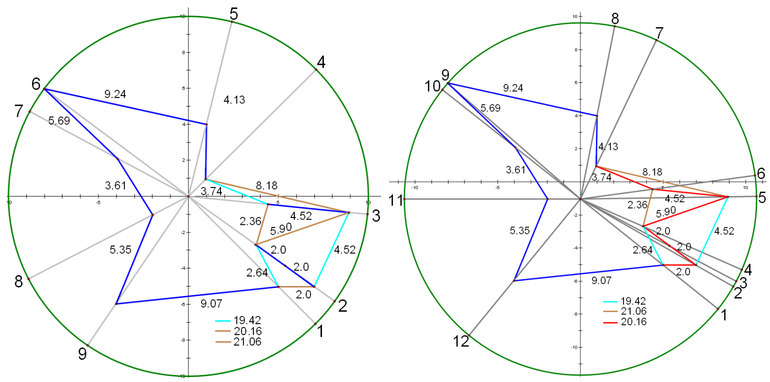

FIGURE 5

to calculate a minimal curve, if not the minimum curve, when all that is needed is a minimal distance. If the number of complicated ρ_i, from i to $i+k$, is large the *complicated ρ algorithm* might some time to compute. Shifting μ_x and recalculating $\sigma \circ \tau$ is greatly reduces computation time. The other important observation that needs to be made is the following definition.

Definition 2. *A set S is not cluttered if there exists a $\epsilon > 0$ such that $\frac{1}{\epsilon} < \rho$ and*

$$\min \|x_i - x_j\| + \epsilon < \min \|\mu_x - x_i\|,$$

OR S is not cluttered if there exists a $\epsilon > 0$ such that $\frac{1}{\epsilon} < \rho$ and

$$\min \|x_i - x_j\| = \min \|\mu_x - x_i\|,$$

and $\rho < 2 * \min \|\mu_x - x_i\| + \frac{1}{\epsilon}$, otherwise S is cluttered.

The idea is that x_i and x_j cannot be to close together, while x_i cannot be to far from μ_x.

Example 2. *Now let $\hat{S}$ equal the set of points from Figure 5.*

$$\hat{S} = (S - (5,5)) \cup \{(3.75, -2.67), (5.38, -3.84), (1,1), (4.3, -0.5), (9, -1)\}$$

Again, it is easily calculated that $\mu_x \approx (1.53, -1.00)$ and that $\rho \approx 11.84$, so $B(\mu_x, \rho) = (x - 1.53)^2 + (y+1)^2 = 140.19$. τ is bijective and the point $(-8, 6) \in \partial B(\mu_x, \rho)$.

The right part of *Figure 6* $\sigma \circ \tau$ has mapped the minimum curve onto $\overline{S}$, when μ_x was either moved or recalculated, $\overline{S}$ does not produce the minimal curve. This leads to the following claim.

Claim 1. *let $\tilde{\rho} = \min \|\mu_x - x_i\|$. If S is not cluttered then $\exists \tilde{\mu}_x \in B(\mu_x, \tilde{\rho})$ such that τ is bijective and $\sigma \circ \tau$ on $B(\tilde{\mu}_x, r)$, for some $r > 0$, will reorder S so that $\overline{S}$ has, in ascending order, the minimum path.*

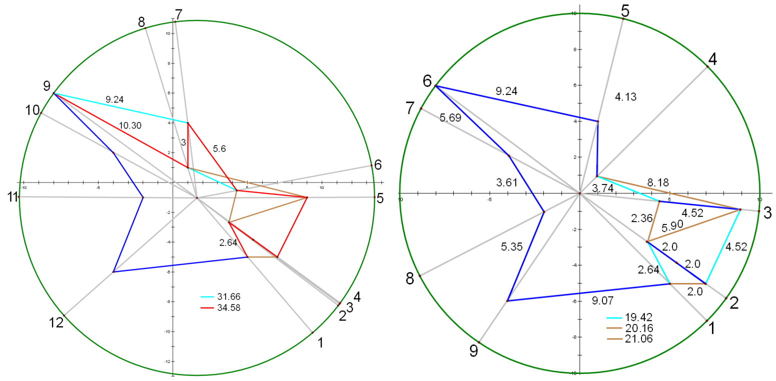

FIGURE 6

Looking at *Example 2* $\min \|x_j - x_i\| = 2$, let $\epsilon = .2$, $\min \|\mu_x - x_i\| = \sqrt{5}$, so $5 < \rho$ and $2 + \epsilon = 2.2 < \sqrt{5}$

S is not cluttered and $\tilde{\rho} = \min \|\mu_x - x_i\| \approx 3.22$. So $(0,0) \in B(\tilde{\mu}_x, r)$ and if we let $\tilde{\mu}_x = (0,0)$ then $\sigma \circ \tau$ constructs the minimum curve from $B(\tilde{\mu}_x, 10)$. Note that $B(\tilde{\mu}_x, 10)$ was the same circle from *Example 1*, and this result can be seen in the right part of *Figure 6*. But there is a better location for μ_x, by the claim. Looking at *Figure 7*, when μ_x is moved to $(0.75, -0.4)$ τ is bijective and orders $\overline{S}$ to contain the minimum path.

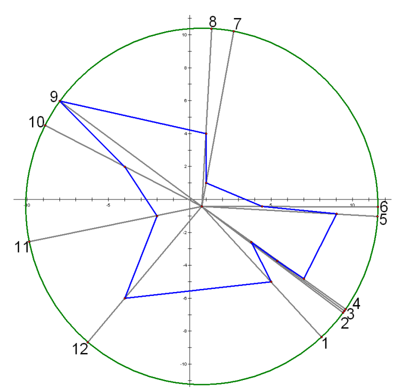

FIGURE 7

5. THE CASE OF $\mathbb{R}^3$

As in the case of $\mathbb{R}^2$, when τ is applied to a set S in $\mathbb{R}^3$, τ maps S to the boundary of $B(\mu_x, \rho)$. The surface of $B(\mu_x, \rho)$ is much more complex. So $\sigma \circ \tau : S \to \overline{S}$.

$$\sigma \circ \tau : \{x_i\}_{i=1}^n \to \{x_j\}_{j=1}^n$$

where $i = j$ does not imply that $x_i = x_j$. When τ is applied to s the image of τ lies on the surface of a sphere. A uniform expansion map will touch the entire image of τ. If possible, the easiest way to determine how σ acts on T is visually, but most of the time this is not possible. However, it is possible to calculate the minimum distance between two arcs on a sphere, since a sphere has constant curvature.

Expansion map. *Starting from the projection of a fixed initial point $\tau(x_1)$ a continuous uniform expansion of the point on the boundary of $B(\mu_x, \rho)$ will tough all of the image of τ. Furthermore, the map will contract onto a single point $\tilde{\tau}(x_1)$ opposite of $\tau(x_1)$. So, any arc connecting $\tilde{\tau}(x_1)$ and $\tau(x_1)$ will be an arc of a great circle.*

Before any calculations are done, a parameterization for $B(\mu_x, \rho)$ is needed.

$$x(u, v) = (\mu_{x_1} + \rho \cos(u) \cos(v), \mu_{x_2} + \rho \sin(u) \cos(v), \mu_{x_3} + \rho \sin(v))$$

The minimum distance between two points on a surface is a geodesic. In the case of a sphere, a geodesic is an arc that travels parallel to the tangent vectors of $\tau(x_i)$ and $\tau(x_{i+1})$. The distance of this arc can be calculated by the geodesic equation, which is a 2nd order differential equation.

$$\frac{d^2 x^\lambda}{dt^2} + \Gamma^\lambda_{uv} \frac{dx^u}{dt} \frac{dx^v}{dt} = 0$$

Where Γ^λ_{uv} is the Christoffel symbol. Lucky, neither the Christoffel symbol nor tensor notation is needed here. Define the the metric $\{E, F, G\}$ for the surface as

$$E = \langle x_u, x_u \rangle, F = \langle x_u, x_v \rangle, G = \langle x_v, x_v \rangle$$

For $B(\mu_x, \rho)$, $F = 0$, so the geodesic equation can be reduced to a system of 2nd order differential equations.

$$\begin{aligned} u'' + \frac{E_u}{2E}(u')^2 + \frac{E_v}{E} u' v' - \frac{G_u}{2E}(v')^2 = 0 \\ v'' - \frac{E_v}{2E}(u')^2 + \frac{G_u}{G} u' v' - \frac{G_v}{2G}(v')^2 = 0 \end{aligned}$$

The metric $\{E, F, G\}$ for the surface of $B(\mu_x, \rho)$ is then defined as the following

$$E = \langle x_u, x_u \rangle = \rho^2 \cos^2(v), F = \langle x_u, x_v \rangle = 0, G = \langle x_v, x_v \rangle = \rho^2$$

For further simplification of calculations, consider $B(0, 1)$, then the system of 2nd order differential equations for the distance on the surface of $B(0, 1)$ reduces to

$$\begin{aligned} u'' - 2\tan(v) u' v' = 0 \\ v'' + \sin(v) \cos(v) (u')^2 = 0 \end{aligned}$$

With the aid of maple, the solution to the system of concern is

$$\sin(u - c_2) = \frac{c_1 \tan(v)}{\sqrt{1 - c_1^2}}$$

Using the angle sum formula for sin and converting back to Euclidian coordinates, a nicer looking representation for the solution is obtained.

$$-\sin(c_2)x_1 + \cos(c_2)x_2 - \frac{c_1 x_3}{\sqrt{1 - c_1^2}} = 0$$

Let, the normal vector $\dot{n}$ be defined as the following

$$\dot{n} = \left\langle -\sin(c_2), \cos(c_2), \frac{c_1}{\sqrt{1 - c_1^2}} \right\rangle$$

This implies that the curves of interest are on the intersection of $B\left(\mu_x, \rho\right)$ and the plane $\dot{n} \cdot (x - \mu_x)$, so in fact the shortest arc does lie on a great circle. In order to calculate the arc of a great circle denote the angle between $\tau(x_i)$, μ_x, $\tau(x_j)$ by α

$$\alpha = \cos^{-1}\left(\frac{\langle \tau\left(x_i\right), \tau\left(x_j\right)\rangle}{\rho^2}\right)$$

Applying the same parameterization for $\tau\left(x_i\right)$ and $\tau\left(x_j\right)$ then simplifying, a general formula for any angle between $\tau\left(x_i\right)$ and $\tau\left(x_j\right)$ is obtained.

$$\alpha = \cos^{-1}\left(\cos(u_i) cos(u_j)(\sin(v_i)\sin(v_j) + \cos(v_i) cos(v_j)) + \sin(v_i)\sin(v_j)\right)$$
$$\alpha = \cos^{-1}\left(\cos(u_i) cos(u_j)\cos(v_i - v_j) + \sin(v_i)\sin(v_j)\right)$$

The distance between any to points on a the surface of $B\left(\mu_x, \rho\right)$ is then $\alpha\rho$. Assuming that x_1 is our starting point, $(\sigma \circ \tau)x_1 = x_1$

$$\left(\sigma \circ \tau\right) x_{i+1} = \left\{x_j : \min\left\{\int_0^\alpha \rho \; dt > 0\right\}\right\}$$

The end result is the map $\sigma \circ \tau$ re-indexes S into $\overline{S}$. Unlike the case for $\mathbb{R}^2$, if $\left(\sigma \circ \tau\right) x_i = \left(\sigma \circ \tau\right) x_{i+1}$ then $\tau(x_i)$ and $\tau(x_{i+1})$ lie on the same circle with the radius as an arc with length $\alpha\rho$. From the uniform expansion map and $\sigma \circ \tau$, a minimal curve will be easy to map, which is pretty neat. Of course, this is after the problem of uniqueness is addressed.

6. Uniqueness of a curve in $\mathbb{R}^3$

If ρ is complicated then the complicated ρ-algorithm can be implemented to find the shortest distance of $\tau(x_i)$ and $\tau(x_{i+1})$. This solve the issue for complicated ρ. There is another problem that arises in $\mathbb{R}^3$

Definition 3. *A circle on the surface of $B(\mu_x, \rho)$ is complicated if $\tau(x_i)$, $\tau(x_j)$ lie on the circle and the circle can be generated by the uniform expansion map.*

If two points, $\tau(x_i)$ and $\tau(x_j)$ have the same arc length from $\tau(x_1)$ then the circle on the surface of $B(\mu_x, \rho)$ that $\tau(x_i)$ and $\tau(x_j)$ belong to is complicated. When this situation arises, the ordering for $\overline{S}$ is not unique. In these cases another 2^n algorithm is needed to resolve the minimum distance of the system of complicated circles.

This structure for σ on T generates a spiraling order for $\overline{S}$ towards the anti-pole of $\tau(x_1)$ on $B(\mu_x, \rho)$. This distance is unique but most likely nowhere near the minimum. To start to alter the path of σ consider the normal plane at $\tau(x_1)$, the tangent and binormal vectors span this plane. The osculating plane splits the sphere into two hemispheres, then let the expansion map traverse the positive half of this plane when it reaches the anti-pole of $\tau(x_1)$, the map traverses the negative half of the plane in the negative direction. It is easy to see that this construction will produce a curve of small distance compared to the set $\overline{S}$ in a spiraling order.

The reordering of S with respect to arc length and osculating plane contains a minimal path. Since the polyhedron is homeomorphic to $B(\mu_x, \rho)$, by claim 1, there exists a point $\tilde{\mu}_x$ near μ_x such that τ is bijective on $B(\tilde{\mu}_x, r)$, for some $r > 0$, and $\overline{S}$ contains the minimum path.

7. Conclusion

This algorithm, under certain conditions, finds the minimum path for a set of points. Furthermore, it can be seen easily that this algorithm can be calculated in polynomial time. The minimum path of a finite number of points in $\mathbb{R}^2$ form a closed polygon, and this polygon is homeomorphic to a circle. This does not mean that traversing around the circle will give light to the minimum path and it does not necessarily mean that the path is minimal. To guarantee a minimal path, the center of the circle that is homeomorphic to this polygon must be contained inside the polygon. And if the set of points is not cluttered, then it can be guaranteed that a minimum path can be found from the circumference of the circle.

In $\mathbb{R}^3$ the minimum path is some edges of some polyhedron, and this polyhedron is homeomorphic to a sphere. Again, this does not mean that calculating arc lengths on this sphere will give light to the minimum path, or even a minimal path. The center of the sphere must be contained inside the polyhedron and the set of points cannot be cluttered. If both of these conditions are satisfied the minimum path between the points can be found without doing permutations of the points and calculating distance.

8. Further areas of Study

The distance of a arc between $\tau(x_i)$ and $\tau(x_{i+1})$ is proportional to $\|\tau(x_i) - x_i\|$ and $\|\tau(x_{i+1}) - x_{i+1}\|$. If the conditions discussed above are satisfied then there is a homotopy between the polygon and a circle. Then the length of the curve can then be found by the circumference. Similarly in $\mathbb{R}^3$, there is a homotopy between the polyhedron and the sphere, and the length of the minimum curve can be found as arcs on great circles on the surface of the sphere.

This algorithm was only considered for a set of points in Euclidean 2 and 3 space. There might be interesting results in higher Euclidean spaces. If we consider a circle on a hyperbolic surface, the algorithm will be similar and will give light to a minimal path given certain restraints. A Torus would be another surface with peculiar results. Another interesting consideration is the algorithm on surfaces without constant curvature. These surfaces are not as nice a sphere. Calculating geodesics on such surfaces can be very troublesome and if a surface does not have constant speed, then it is impossible to generalize the shape an expansion map from a point.

If a set of points wants to be considered in $\mathbb{R}^2$ or $\mathbb{R}^3$ where the ending is not the same as starting point. A similar algorithm can be thought about. However, many cases arise for a situation of this type and similar cases would have to be classified into categories. If there is such a classification for a set of points then there should exist an algorithm that can be calculated in polynomial time.

References

[1] Kreyszig, Erwin, *Differential Geometry* (New York: Dover, 1991), Chapters 5, 6
[2] Roman, Steven, *Advanced Linear Algebra* (New York: Springer, 2008), Chapter 2
[3] Bishop, R. L., Goldberg, S. I., *Tensor Analysis on Manifolds* (New York: Dover, 1980), Chapter 5
[4] Chern S. S., Shen Z., *Riemann-Finsler Geometry* (Massachusetts: World Scientific, 2005), Chapters 5, 6

E-mail address, James C Hateley: `jhateley@uci.edu`