RMPh solutions

April 2021

1 Solution to "Steiner's symmetry breaking highway"

Part I. Before we start designing the highway network for each set of cities we can make several observations that apply to all triangle configurations. Because the minimal path between two points A and O on the plane, is given by a straight line, the highway network should only consist of straight lines and intersections. Secondly, since every intersection includes at least 3 edges (or 2 in the degenerate case where the intersection is the same as one of the vertices of the triangle), there should never be more than one intersection if we wanted to minimize the total length of the highway network connecting three cities (which we will denote by A, B and C).

a) *(1 point)* The symmetric configurations are those that are invariant under 120 degrees rotations around the center of the equilateral triangle and under reflections around its median axes. There is only one such symmetric configuration which involves only one intersection: the intersection is in the center of the triangle and the edges of the network connect this intersection point to the vertices of the triangle. The total length of this configuration is $D = 3(2/3)(\sqrt{3}L/2) = \sqrt{3}L$.

b) i) *(1 points)* To prove the identity let's consider the vector connecting A to O to be $\vec{d}$ and A to O' to be $\vec{d'}$, with $\vec{d'} = \vec{d} + \vec{\delta}$ and $d = \sqrt{\vec{d}\cdot\vec{d}}$. Then the norm of the vector $\vec{d'}$ is given by

$$d' = d + \delta d = \sqrt{\vec{d'}\cdot\vec{d'}} = \sqrt{d^2 + 2\vec{\delta}\cdot\vec{d} + \delta^2} = d + \frac{\vec{\delta}\cdot\vec{d}}{d} + O(\delta^2)\,. \tag{1.1}$$

Then, for $d \neq 0$,

$$\delta d = \vec{\delta}\cdot\vec{e} + O(\delta^2)\,, \tag{1.2}$$

with $\vec{e} = \vec{d}/d$ is the unit vector pointing from A to O. Importantly, for b) iii), the formula above applies as long $d \neq 0$.

b) ii) *(1.2 points)* To minimize the minimum length path we wish to find an intersection point such that the derivative of the length with respect to the infinitesimal translation vector $\vec{\delta}$ is zero. Using (1.2), we find that if the intersection point is translated by δ, then the total length of the network changes as

$$\delta D = \delta\cdot(\vec{e}_{OA} + \vec{e}_{OB} + \vec{e}_{OC}) + O(\delta^2)\,, \tag{1.3}$$

where $\vec{e}_{OA}$, $\vec{e}_{OB}$, $\vec{e}_{OC}$ are the normal vectors pointing from the intersection point O to A, B or C respectively. Since we want δD to be zero to leading order in δ, for any direction of the vector $\vec{\delta}$, then $\vec{e}_{OA} + \vec{e}_{OB} + \vec{e}_{OC} = 0$. This, in turn, implies

$$(\vec{e}_{OA} + \vec{e}_{OB})^2 = 2 + 2\vec{e}_{OA}\cdot\vec{e}_{OB} = 1 \tag{1.4}$$

which implies that $\vec{e}_{OA}\cdot\vec{e}_{OB} = -1/2$ and, consequently the angle between OA and OB is 120 degrees. Similarly, starting from the same relation we can prove that the angle between OB and OC, and OC and OA are all 120 degrees.

There is one corner case that we should also be concerned with. The formula (1.2) is valid as long as $d \neq 0$, i.e. if the point O is not the same as A, B, or C. We need to make sure that if O is the same as A, B, or C, then it is not the network of minimal length. Let's assume without loss of generality that $O = A$. Then, translating O by an infinitesimal vector δ yield a length change

$$\delta D = \delta + \vec{\delta} \cdot (\vec{e}_{AB} + \vec{e}_{AC}) \tag{1.5}$$

We can note that the minimum of $\vec{\delta} \cdot (\vec{e}_{AB} + \vec{e}_{AC})$ is given by $\vec{\delta}$ pointing in the opposite direction as $(\vec{e}_{AB} + \vec{e}_{AC})$ For such a vector $\vec{\delta}$, the length change becomes

$$\delta D = \delta(1 - |\vec{e}_{AB} + \vec{e}_{AC}|)\,. \tag{1.6}$$

When the angle between AB and AC is less than 120 degrees, then $|\vec{e}_{AB} + \vec{e}_{AC}| = \sqrt{2 + 2\vec{e}_{AB} \cdot \vec{e}_{AC}} > 1$, and the total length of the network can be decreased by going along a vector $\vec{\delta}$ opposite to $(\vec{e}_{AB} + \vec{e}_{AC})$.

Therefore, for a triangle that has all angles smaller than 120 degrees, the minimal length network has an intersection point within the interior of the triangle triangle, that has the angles between any two straight edges equal to 120 degrees. Such an intersection always exists for such triangles and is named Toricelli's or Fermat's point.

b) iii) *(0.8 points)* We will call the vertex which has the angle θ_A greater than 120 degrees A. When the triangle has an angle greater than 120 degrees, than it is impossible to have $\vec{e}_{OA} + \vec{e}_{OB} + \vec{e}_{OC} = 0$. Instead the minimum is achieved by taking O to be in the vertex A. To see this, note that according to (1.6), the smallest variation of the length of the network is given by $\delta D = \delta(1 - |\vec{e}_{AB} + \vec{e}_{AC}|) = \delta(1 - \sqrt{2 + 2\cos(\theta_A)})$. Since $\cos(\theta_A) < -1/2$ we have that $\delta D > 0$. Therefore, moving the intersection point O can only increase the length of the network. Consequently, this proves that having the intersection point O in A yields the minimal length highway network.

Part II. For convenience, we will denote the vertices of the square as A, B, C, and D.

c) *(0.7 points)* The symmetric configuration is the network that is invariant when rotated by 90 degrees, or reflected around the horizontal or vertical axis of the square. This is given by a network with one intersection point O placed in the center of the square connected to the vertices A, B, C and D through four straight edges. The total length of this network is given by $2\sqrt{2} \cdot L$, where L is the length of the square's edges.

Let's prove that this configuration has a minimal length among the network with a single intersection point. Using (1.2), the variation if we translated O by $\vec{\delta}$ is given by

$$\delta D = \vec{\delta}(\vec{e}_{OA} + \vec{e}_{OB} + \vec{e}_{OC} + \vec{e}_{OD}) + O(\delta^2) = O(\delta^2)\,, \tag{1.7}$$

since at the center $\vec{e}_{OA} + \vec{e}_{OB} + \vec{e}_{OC} + \vec{e}_{OD} = 0$. Therefore, this implies that the configuration with O in the center of the square is indeed a minimal length configuration amongst networks with one intersection point.

d) *(1.5 points)* Let's consider other possible configurations. Once again, since each intersection has at least three edges entering it (except for the case where an intersection point agrees with a vertex which we will once again check separately), the minimal length network can have at most 2 intersection points, each with 3 edges entering it. We will call these two intersection points O_1 and O_2. Translating O_1 or O_2 independently by any infinitesimal vectors $\vec{\delta}_1$ or $\vec{\delta}_2$, respectively, should not decrease the distance. Therefore, we can use the variation formula (1.3), independently for the points O_1 and O_2. Therefore the angles between all the edges entering O_1 and O_2, respectively, are all equal to 120 degrees. This leads to one of the two configuration shown in figure 1, both of which have paths of equal length.

The length of this configuration is:

$$D = 4\left(\frac{2}{\sqrt{3}}\frac{L}{2}\right) + \left(L - 2\frac{L}{2}\frac{1}{\sqrt{3}}\right) = L(1 + \sqrt{3}) = L \cdot 2.732\ldots \tag{1.8}$$

where the term in the first parenthesis gives the length of the the edges connecting $O_{1,2}$ to the closest vertex of the square and the term in the second parenthesis is the length of the edge connecting O_1 to O_2.

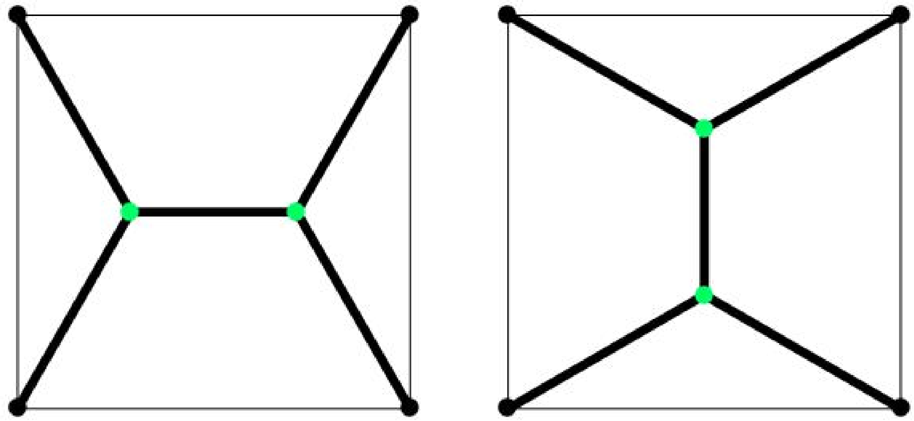

Figure 1: Minimum length configuration for the square.

Finally, we also need to make sure that the configurations that we just discussed are indeed those with the shortest path, not only amongst those that have two intersections. First, we note that $1+\sqrt{3} < 2\sqrt{2}$ and, therefore, the two intersection configuration has a shorter length than the one intersection configuration. We also need to make sure that the degenerate configurations where the intersection points are the same as two of the vertices of the square is longer than the two intersection configuration. The former configuration has length $3L > L(1+\sqrt{3})$.

Therefore, the configurations in figure 1 are those of shortest overall length. Note that each of the two configurations do not obey all the symmetries of the square: for instance, 90 degrees rotations around the center of the square maps one configuration to the other, but not to itself. This is in contrast to the example in part c) where all the symmetry transformations leave the configuration invariant. Therefore, as hinted at in the problem's introduction, this is a basic example of spontaneous symmetry breaking in which the minimal configuration is not fully symmetric.

e) *(0.8 points)* As we get the intersection points O_1 and O_2 closer or further away from each other (we will denote the distance from O_1 to O_2 as $2x$), the total length is given by,[1]

$$D(x) = 4\sqrt{\frac{L^2}{4} + \left(\frac{L}{2} - |x|\right)^2} + 2|x|\,. \tag{1.9}$$

The sketch of $D(x)$ is:

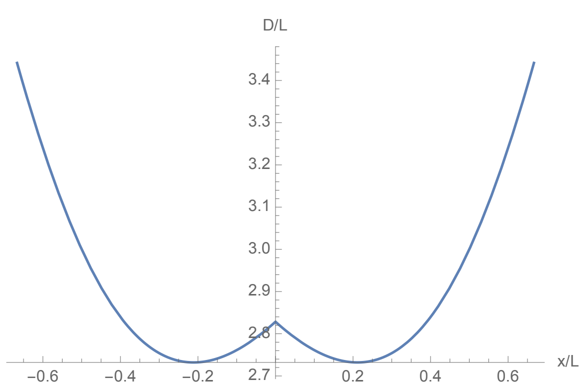

Figure 2: The total length of the 2 intersection configuration as a function of the coordinate of the point O_1. The distance between O_1 and O_2 is $2|x|$.

[1] The absolute value ensures that we connect $O_{1,2}$ to the closest pair of vertices.

where we need to emphasize that the minimal value is achieved for the configurations shown in figure 1.

Part III. An important point to generalizing Steiner's observation regarding the two cities placed on the Equator, is that the choice of North and South pole is arbitrary for a homogeneous sphere. Therefore, with a different choice of North and South pole, any two points A and B on the sphere can be placed on the same equator. This is called the "great circle" that passes through A and B.

f) *(1.0 points)* Since lines of constant longitude are all "great circles" between the North and South pole, the minimal path between the North, South pole and any third city is determined by the longitude line passing through the third city (i.e. the three points are co-linear). In the case of Bucharest, the minimal path is therefore given by the 26.1 degrees East longitude line going from the North to the South pole.

g) *(2.0 points)* We will present a computational and a more abstract solution to finding the minimal path connecting the three cities. Once again, before we start the derivation we need to make several general observations. Just like the situation on the plane, for 3 cities, we are only interested in the case with one intersection, which we will once again denote by O. To achieve the minimal length configuration this intersection point is connected to the each of the three vertices of the triangle along the "great circles". The task is therefore, to find the location of the intersection point O.

Solution I: We will denote the normal unit vectors at the vertex points by $\vec{n}_A$ (which we choose to be at the north pole), $\vec{n}_B$ and $\vec{n}_C$, and the normal unit vector at O by $\vec{n}_O$. Let's translate the point O by an infinitesimal vector $\vec{\delta}$ to the point O'. The normal vector at O' is $\vec{n}_{O'} = \vec{n}_O + \vec{\delta}$.[2] the change in distance between O and one of the vertices (let's say A) is

$$\delta d_{OA} = R\left[\arccos\left(\vec{n}_A \cdot \vec{n}_O + \vec{n}_A \cdot \vec{\delta}\right) - \arccos\left(\vec{n}_A \cdot \vec{n}_O\right)\right] = -R\frac{\vec{n}_A \cdot \vec{\delta}}{\sqrt{1-(\vec{n}_A \cdot \vec{n}_O)^2}}\,. \tag{1.10}$$

Therefore, the variation of the overall distance of the network under an infinitesimal translation of O is

$$\delta D = \delta d_{OA} + \delta d_{OB} + \delta d_{OC} = -R\,\vec{\delta}\cdot\left(\frac{\vec{n}_A}{\sqrt{1-(\vec{n}_A \cdot \vec{n}_O)^2}} + \frac{\vec{n}_A}{\sqrt{1-(\vec{n}_C \cdot \vec{n}_O)^2}} + \frac{\vec{n}_C}{\sqrt{1-(\vec{n}_C \cdot \vec{n}_O)^2}}\right)\,. \tag{1.11}$$

This does not determine the angles between the edges just yet. To do that, we can express the inner product above in terms of the tangent vectors $\vec{e}_{OA}$, $\vec{e}_{OB}$, and $\vec{e}_{OC}$ pointing from O to each vertex of the triangle, A, B or C respectively. To find $\vec{e}_{OA}$, we first observe that this vector is coplanar with $\vec{n}_O$ and $\vec{n}_A$. Therefore, as long as O and A are not at opposite poles of the sphere or are identical to each other, than $\vec{e}_{OA}$ can be expressed as a liner combination of $\vec{n}_O$ and $\vec{n}_A$. Therefore, we need to find α and β such that

$$\vec{e}_{OA} = \alpha\vec{n}_O + \beta\vec{n}_A\,, \qquad \vec{e}_{OA}\cdot\vec{n}_O = 0\,, \qquad \vec{e}_{OA}\cdot\vec{e}_{OA} = 1\,, \tag{1.12}$$

which implies

$$\alpha + \beta\vec{n}_A\cdot\vec{n}_O = 0\,, \qquad \alpha^2 + \beta^2 + 2\alpha\beta\vec{n}_O\cdot\vec{n}_A = 1\,. \tag{1.13}$$

The solutions to this system of equation yields

$$\vec{e}_{OA} = \pm\left(\frac{\vec{n}_O\cdot\vec{n}_A}{\sqrt{1-(\vec{n}_O\cdot\vec{n}_A)^2}}\vec{n}_O - \frac{1}{\sqrt{1-(\vec{n}_O\cdot\vec{n}_A)^2}}\vec{n}_A\right)\,. \tag{1.14}$$

We obtained two solutions only one of which points from O to A along the shortest path. With a bit of thought one can check that this corresponds to the $-$ solution in (1.14). Using (1.14) we find that (1.11) can be expressed as

$$\delta D = R\vec{\delta}\left(\vec{e}_{OA} + \vec{e}_{OB} + \vec{e}_{OC}\right) \tag{1.15}$$

[2]Note that the vector $\vec{n}_{O'}$ is still a normal vector to leading order in δ since $|\vec{n}_{O'}| = 1 + \vec{n}_O\cdot\vec{\delta} + O(\delta^2) = 1 + O(\delta^2)$, since $\vec{\delta}$ and $\vec{n}_O$ are orthogonal.

This is precisely the same formula as the one found on the plane! We will see why this is the case in an intuitive way as part of solution II. Using the same reasoning as in Part 1 we can conclude that the angle between all the edges is again precisely 120 degrees. Finally, as before, we need to check the corner cases in which O coincides with one of the point A, B and C. The angle at A is 60 degrees, the angle at B is 90 degrees and the angle at C is also 90 degrees (note that for a triangle on a sphere the sum of the angles is not 180 degrees). Because all these angles are smaller than 120 degrees we conclude, using the same reasoning as in Part 1, that any of these configurations cannot be a minimum.

Therefore, to summarize, the minimum length configuration is achieved by a network in one intersection, whose edges point along great circles between this intersection point and each vertex of the triangle, and the angle between all vertices is 120 degrees.

Solution II: Let's consider some configuration with one intersection point O which is not necessarily the minimum one. Let's choose a small sphere of radius ϵ, with $\epsilon \ll R$, around the intersection point O. Choose A', B' and C' to be the intersection point between the edges and this small sphere. Since $\epsilon \ll R$ we can approximate well the region of the larger sphere contained within the small sphere by a flat plane. We can therefore use the results in part 1 to find the shortest path between the points A', B' and C'. This is given by the point O', Euler's point in the triangle A', B' and C'. Therefore, as long as initial angle between all the edges coming into O is not 120 degrees, than there is a shorter path by moving O to O' within the $A'B'C'$ triangle. Therefore, if O is indeed the intersection point of the minimal length configuration, it better be that the angle between all its edges is precisely 120 degrees; otherwise, the procedure discussed above will generate an even shorter path. Note that the argument discussed here is independent of the local curvature of the surface and our argument applies equally well on a sphere as it does on a complicated surface, with mountains and complicated valleys.[3] Once again, the only other cases we should worry about are the corner cases in which the point O is identical to one of the vertices of the triangle. Using the same reasoning as in Part I, and the vertex angles discussed in solution I, we however once again conclude that these corner cases do not lead to a shorter configuration.

[3] However, note that in these complicated cases it might be that there are multiple configurations that have one intersection point and 120 degrees between all the edges at the intersection point. We have to compare the overall length of all such configurations and choose the minimal one.

Part I. a)	Identifying all the relevant symmetries Identifying the configuration Computing the total distance	0.3 points 0.4 points 0.3 points
Part I. b) i)	Identifying the relevant inner products Obtaining d' in an expansion in δ Correct answer	0.4 points 0.3 points 0.3 points
Part I. b) ii)	Identifying that the edges are straight lines Identifying the possible number of intersection points Identifying the formula for the variation δD Conclusion about the angles between the edges Covering edge cases	0.2 points 0.3 points 0.3 points 0.3 points 0.1 point
Part I. b) iii)	Identifying the formula for the variation δD Conclusion about the configuration	0.5 points 0.3 points
Part II. c)	Identifying all the relevant symmetries Identifying the configuration Computing the total distance Proof that it is the min. length for 1 intersection	0.2 points 0.2 points 0.2 points 0.1 points
Part II. d)	Identifying the number of intersections Identifying the angles and configurations Computing the total distance Corner cases and comparison	0.35 points 0.6 points 0.35 points 0.2 points
Part II. e)	Qualitative plot With correct minim	0.4 points 0.4 points
Part III. f)	Identifying the type of edges Identifying the correct configuration	0.5 points 0.5 points
Part III. g)	Identifying the type of edges Identifying the number of intersection points Proof that the configuration discussed is the min. one Identifying the configuration Corner cases	0.2 points 0.6 points 0.5 points 0.6 points 0.1 point

Table 1: Point distribution for Problem 1. Alternative solutions exist and are accepted in which case the point distribution might be slightly different since the steps of the proof might be different.