Geometrical Application of Pigeonhole Principle

Eric Kim

Seoul International School

Abstract

The pigeonhole principle in theory, is a simple concept that many people can understand and is applicable to a variety of problems. However, in order to apply the pigeonhole principle, certain conditions must be met within the given parameter, greatly increasing the difficulty. Nonetheless, with proper set-up, the pigeonhole principle can be used to eliminate extensive calculation and simplify the process.

Problem Statement

The following problem from 2019 SUMaC application problems set requires solid understanding of both geometry and pigeonhole principle.

"An equilateral triangle has sides of length 1 cm.

(a) Show that for any configuration of five points on this triangle (on the sides or in the interior), there is at least one pair of from these five points such that the distance between the two points in the pair is less than or equal to .5 cm.

(b) Show that .5 (in part (a)) cannot be replaced by a smaller number even if there are 6 points.

(c) If there are eight points, can .5 be replace by a smaller number? Prove your answer."

Outline

Part (a) of the problem can be proved by applying the pigeonhole principle. The given triangle needs to be divided into number of regions in an appropriate manner in order to use the principle. Next, Part (b) can be solved by giving a counter example that ultimately refers back to a similar diagram from part (a). Similar to part (a), part (c) can be solved by dividing the triangles into appropriate regions and algebraically compare the maximum lengths of each region. The objective is to minimize the maximum possible distances between two points in order to replace .5 by a smaller number. After solving it algebraically, the answer can be confirmed by graphing the different equations.

Pigeonhole Principle

If there are n pigeons and m pigeonholes, where n > m, then there must be at least one pigeonhole that contains more than one pigeon. Process Paper

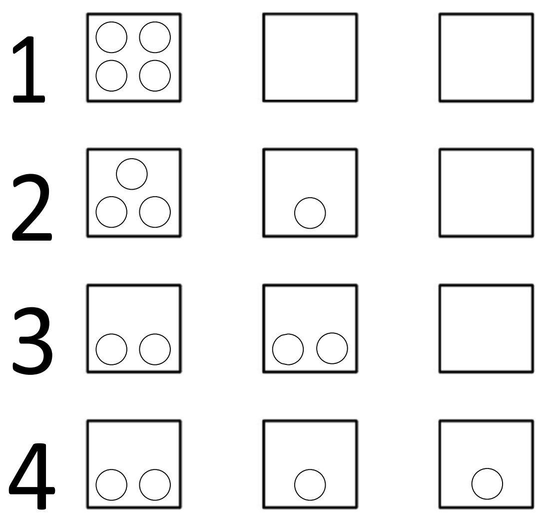

Figure 1

Given n pigeons and m pigeonholes, where n > m, after assigning m pigeons into m pigeonholes such that each pigeonhole contains only one pigeon, then there are n - m pigeons remaining. No matter how the remaining n – m pigeons are distributed, at least one pigoenhole has more than one pigeon, since every pigeonhole is already occupied.

In order to clarify the explanation above, a diagram with all possible configurations is given for four pigeons (balls) and three pigeonholes (boxes) as shown in Figure 1. There are four possible configurations in which the balls can be distributed. In the first configuration, there are four balls in the first box. In the second configuration, there are three balls in the first box and one ball in the second box. In the third configuration, there are two balls in the first box and the other two in the second box. In the last configuration, there are two balls in the first box, and the remaining boxes have one ball each. In each of the configurations, there is at least one box with two or more balls. Thus, the pigeonhole principle remains true, as each of the configuration has at least one pigeonhole with more than one pigeon.

Proof of Part (a)

"An equilateral triangle has sides of length 1 cm. Show that for any configuration of five points on this triangle (on the sides or in the interior), there is at least one pair of from these five points such that the distance between the two points in the pair is less than or equal to .5 cm."

Showing that there is a pair of points with distance less than 0.5 cm is equivalent to showing that the maximum possible distance between two points is 0.5 cm.

If a triangle is divided into four or less regions, at least two out of the five points must be in the same region when distributed according to the pigeonhole principle. Using this fact, it must be proved that the maximum distance between any two points within one of the regions is 0.5 cm.

In order to apply the pigeonhole principle, there must be corresponding "box" and "balls" within the problem. Since the goal is to prove that there is at least one pair of points with distance less than or equal to 0.5 cm, the triangle must be divided into four or less regions, which will function as the boxes. If the five points are considered as the balls, there must be at least one region with two or more points according to the pigeonhole principle. Therefore, the first step is to divide the triangle into four or less regions as shown in Figure 2, which all have a maximum distance of 0.5 cm between any two points inside the region.

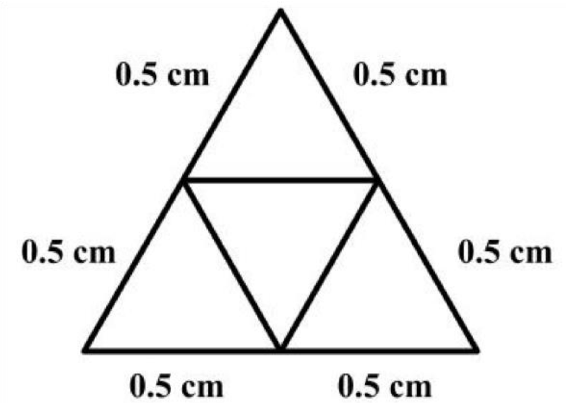

Figure 2

It must be proven that the maximum distance between two points within one of the four regions is 0.5 cm because all four regions are congruent as shown in the diagram above.

Lemma 1

The two points must be on the edge, not the interior of the triangle, in order to have the maximum distance between them.

Proof of Lemma 1

If there are two random points (X, Y) in the interior of the triangle, as shown in Figure 3, a line can be drawn that is passing through these two points as shown in Figure 4. The line intersects two different edges of the triangle at points X′ and Y′ as shown in Figure 5. The distance between X′ and Y′ ($\overline{X'Y'}$) is greater than the distance between X and Y ($\overline{XY}$), meaning there are always two points that are further apart on two edges than two interior points. Therefore, the two points must be on two different edges of the triangle as shown in *Figure 6*

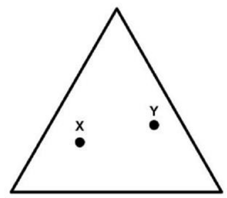

Figure 3

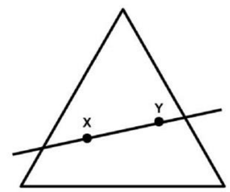

Figure 4

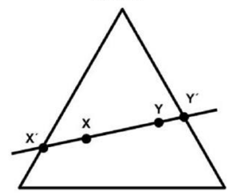

Figure 5

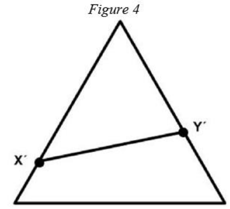

Figure 6

Lemma 2

Two points must be on the vertices, not just on the edges of the triangle.

Proof of Lemma 2

From the two points which are on two different edges, one point can be fixed, while the other point is free to move along the edge of the triangle. The free point should be on the edge of the triangle according to Lemma 1. In this case, the fixed point will be X′ and the free point will be Y′. Draw a perpendicular line from the fixed point to the edge with the free point. The perpendicular distance $\overline{X'P}$ will always be fixed as seen in Figure 7.

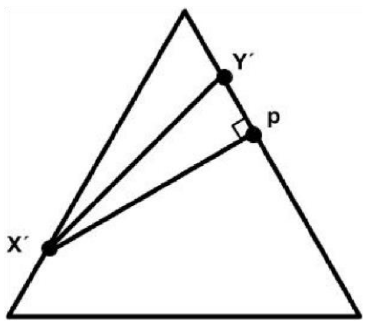

Figure 7.

In order to find the maximum length of must be

$$\overline{X'Y'} = \sqrt{(X'P)^2 + (Y'P)^2}, \quad \overline{Y'P}$$

maximized since $\overline{X'P}$ is a fixed distance. According to Lemma 1, Y' can move along the edge of the triangle, but to maximize $\overline{X'Y'}$, Y' must be placed on a vertex because the length of $\overline{Y'P}$ increases as it approaches a vertex. This leads to two different cases, where Y' can be placed on either the top or the bottom right.

Case 1: Y' fixed to top vertex

Once Y' is fixed, X' can be moved along the edge of the triangle according to lemma 1. The two possible locations for X' are the vertex on the top or the vertex on the bottom left. In order to have the maximum distance, X' must be on the bottom left vertex as shown in Figure 8.

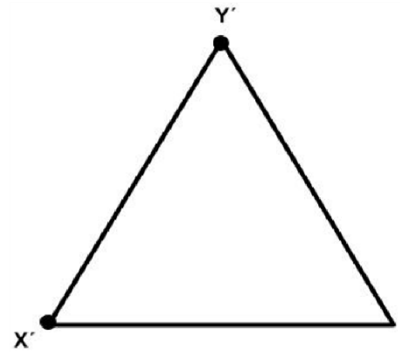

Figure 8

Case 2: Y' fixed to bottom right vertex

Since $\overline{Y'P}$ is a fixed distance, the Pythagorean theorem:

$$\overline{X'Y'} = \sqrt{(X'P)^2 + (Y'P)^2}$$

(Figure 9) requires $\overline{X'P}$ to be maximized in order to find the longest distance for $\overline{X'Y'}$ Therefore, X' must be moved to either one of the remaining two vertices (top or bottom left).

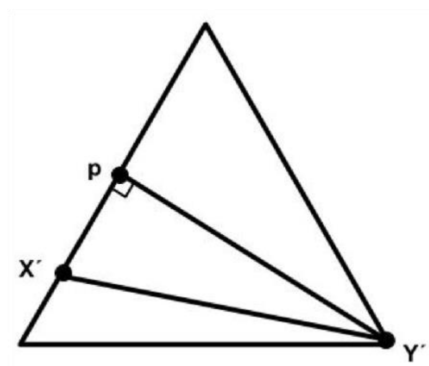

Figure 9

In conclusion, according to both Lemma 1 and 2, it is proven that both points (X' and Y') must be on the vertices of the triangular region in order to maximize the distance between them. Here, the distance between the two points is equal to the side length of the region, which is 0.5 cm. In other words, there exists at least one pair of points whose distance is less than or equal to 0.5 cm.

Proof of Part (b)

"Show that .5 (in part (a)) cannot be replaced by a smaller number even if there are 6 points."

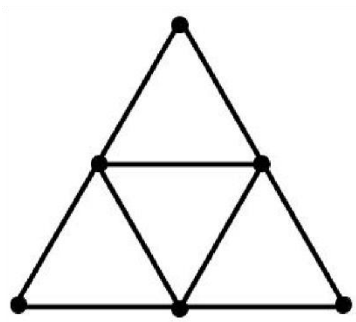

Figure 10

The problem will be proved using a counterexample that shows that 0.5 cm is the maximum distance between any two points of the triangle. In Figure 10, the six points are distributed so that there are three points on each vertices of the triangle and three midpoints on each side of the equilateral triangle. The shortest distance between any two points of the triangle are the two adjacent points, which all have a distance of 0.5 cm.

Proof of Part (c)

"If there are eight points, can 0.5 be replaced by a smaller number? Prove your answer"

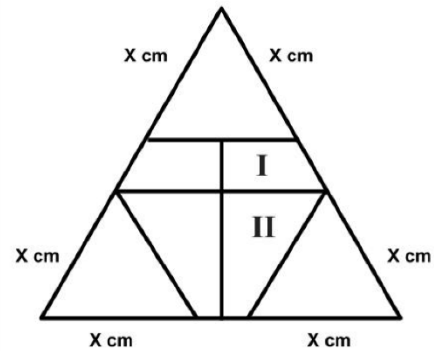

Figure 11

In order to apply the pigeonhole principle with eight points (eight pigeons), there must be less than eight regions (eight pigeonholes). Therefore, the equilateral triangle is divided into seven regions, which consists of three smaller equilateral triangles and two pairs of congruent trapezoids as shown in Figure 11.

Let each side of the smaller equilateral triangle equal X cm, where X must be less than 1/2, otherwise the diagram would be same as in part (a) with 4 triangular regions.

According to Lemma 1 and Lemma 2, two points on a triangular region must be on the vertices of the region in order to have the maximum distance between them, which is X cm. Next, the regions that must be considered are the two remaining types of trapezoidal regions. Since there are two types of trapezoidal regions (I & II), the maximum distance between each of the trapezoids will be analyzed separately.

Analysis of Region I

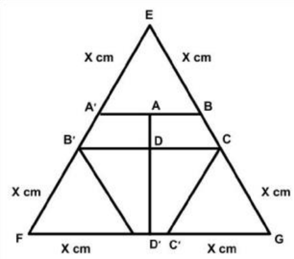

Figure 12

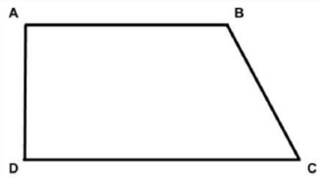

Figure 13

In Figure 12 and 13, the distances that must be considered are lengths $\overline{AB, AC, AD, BC, BD, CD}$. The Pythagorean Theorem will be used in order to find the pair of points that has the maximum distance.

Since ∠BAD is a right angle, ΔBAD is a right triangle and BD is the hypotenuse as shown in Figure 14. According to the Pythagorean theorem: $\overline{BD}^2 = \overline{AB}^2 + \overline{AD}^2$, AB and AD are shorter than BD. Therefore, the lengths AB and AD can be eliminated from the possible options of the maximum length. Likewise, the Pythagorean theorem, $\overline{AC}^2 = \overline{AD}^2 + \overline{DC}^2$ implies that AD and CD are shorter than AC as shown in Figure 15. As a result, CD can be eliminated from the possible options for the maximum length.

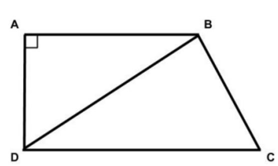

Figure 14

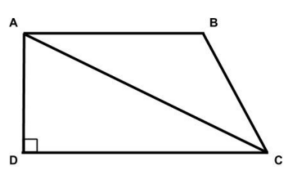

Figure 15

The remaining options are BD, AC, and BC.

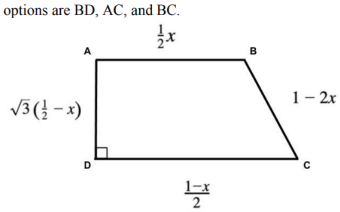

Figure 16

Every side of the trapezoidal region I can be computed as the following, and the results are shown in Figure 16.

$$\overline{AB} = \frac{\overline{A'B}}{2} = \frac{\overline{EB}}{2} = \frac{x}{2}$$

$$\overline{BC} = 1 - \overline{EB} - \overline{CG} = 1 - x - x = 1 - 2x$$

$$\overline{CD} = \frac{\overline{B'C}}{2} = \frac{\overline{EC}}{2} = \frac{\overline{EB} + \overline{BC}}{2} = \frac{x + (1-2x)}{2} = \frac{1-x}{2}$$

$$\overline{AD} = \overline{ED} - \overline{EA} = \sqrt{3}\overline{CD} - \sqrt{3}\overline{AB} = \sqrt{3}(\overline{CD} - \overline{AB}) = \sqrt{3}\left(\frac{1-x}{2} - \frac{x}{2}\right) = \sqrt{3}\left(\frac{1}{2} - x\right)$$

In order to compare the lengths for $\overline{AC}$ and $\overline{BD}$, the Pythagorean theorem will be used. Since length $\overline{AD}$ is common, only lengths $\overline{DC}$ and $\overline{AB}$ need to be checked. Subtracting the lengths of $\overline{DC}$ and $\overline{AB}$ will determine which length is longer.

$$\overline{DC} - \overline{AB} = \frac{1}{2} - \frac{x}{2} - \frac{x}{2} = \frac{1}{2} - x$$

Since the maximum of x is 1/2 or 0.5 cm, ½-X is greater than zero, meaning that DC – AB > 0 and DC > AB. Using the Pythagorean theorem for lengths AC and BD, the following system of equations can be derived.

$$\begin{cases} \overline{BD}^2 = \overline{AB}^2 + \overline{AD}^2 \\ \overline{AC}^2 = \overline{DC}^2 + \overline{AD}^2 \end{cases}$$

$$\overline{AC}^2 - \overline{BD}^2 = \overline{DC}^2 + \overline{AD}^2 - (\overline{AB}^2 + \overline{AD}^2) = \overline{DC}^2 - \overline{AB}^2$$

It is proven that DC > AB, so DC – AB will be greater than zero. Rearranging the equation as an inequality, it can be seen that AC is greater than BD.

$$\overline{AC}^2 - \overline{BD}^2 > 0$$
$$\overline{AC}^2 > \overline{BD}^2$$
$$\overline{AC} > \overline{BD}$$

After eliminating $\overline{BD}$ from the list of possible options, the remaining lengths to consider are $\overline{AC}$ and $\overline{BC}$. ∠ ABC is 120 degrees, as △A'BC is equilateral and ∠CBE is a straight angle. Subtracting ∠A'BE, one angle of an equilateral triangle, from 180° gives 120°. △ABC is an obtuse triangle, making ∠ABC the largest angle of the triangle. Therefore, AC, the opposite side of ∠ABC, is the maximum length for Region I.

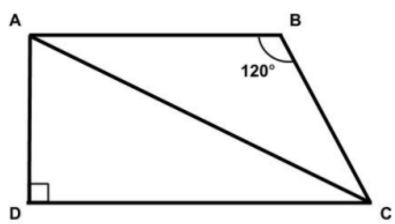

Figure 17

The law of cosines can be used to find the length of AC, since the lengths of AB, AC and the angle between them are known, as seen in Figure 17.

$$\overline{AB} = \frac{x}{2}$$

$$\overline{BC} = 1 - 2x$$

$$\overline{AC}^2 = \overline{AB}^2 + \overline{BC}^2 - 2 \cdot \overline{AB} \cdot \overline{BC} \cdot \cos 120° \text{ (Law of cosines)}$$

$$\overline{AC}^2 = \left(\frac{x}{2}\right)^2 + (1-2x)^2 - \left\{2 \cdot \frac{x}{2} \cdot (1-2x) \cdot \cos 120°\right\}$$

$$\overline{AC}^2 = \frac{13x^2}{4} - \frac{7x}{2} + 1$$

$$\overline{AC} = \sqrt{\frac{13x^2}{4} - \frac{7x}{2} + 1}$$

$$\overline{AC} = \frac{1}{2}\sqrt{13x^2 - 14x + 4}$$

Next, the lengths in Region II (Figure 18) must be analyzed. Considering that the trapezoids in both Region I and Region II have two perpendicular distances, the same approach for finding the maximum length can be used. The maximum distance for Region II will then be C'D

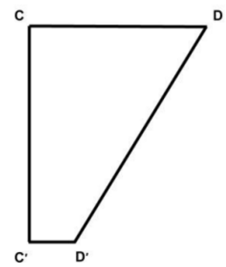

Figure 18

The lengths of CD and DD' are needed to use the Pythagorean theorem for finding the length of CD'. Only the length for DD' needs to be found since the length for CD has been found previously

$$\overline{CD} = \frac{1-x}{2}$$

$$\overline{DD'} = \overline{ED'} - \overline{ED} = \sqrt{3}\overline{D'G} - \sqrt{3}\overline{CD} = \sqrt{3}(\overline{D'G} - \overline{CD}) = \sqrt{3}\left(\frac{1}{2} - \frac{1-x}{2}\right) = \frac{\sqrt{3}}{2}x$$

$$\overline{CD'}^2 = \overline{CD}^2 + \overline{DD'}^2 \text{ (Pythagorean theorem)}$$

$$\overline{CD'}^2 = \left(\frac{1-x}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}x\right)^2$$

$$= \frac{1-2x+x^2}{4} + \frac{3x^2}{4}$$

$$= \frac{1-2x+4x^2}{4}$$

$$\overline{CD'} = \sqrt{\frac{1-2x+4x^2}{4}}$$

$$= \frac{1}{2}\sqrt{4x^2 - 2x + 1}$$

The maximum distances for the three different regions: triangular, trapezoidal I, trapezoidal II are found in terms of x. The goal is to replace and improve the original bound of 0.5cm to a smaller value. This will be done by finding the supremum and minimizing the maximum of these three values within the interval 0 < x < 0.5.

The maximum distance for the triangular region will be denoted as curve 1, trapezoidal region I as curve 2, and trapezoidal region II as curve 3.

Curve 1: $y = x$ (one side length of equilateral triangle)

Curve 2: $y = \frac{1}{2}\sqrt{13x^2 - 14x + 4}$ ($\overline{AC}$)

Curve 3: $y = \frac{1}{2}\sqrt{4x^2 - 2x + 1}$ ($\overline{CD'}$)

The intersection points between each of the curves will be first found to find the maximum y value. Each of the curves are first set equal to each other and the three equalities will be solved to find the corresponding x and y values.

Curve 1 & Curve 2 Intersection point:

$$x = \frac{1}{2}\sqrt{13x^2 - 14x + 4}$$
$$2x = \sqrt{13x^2 - 14x + 4}$$
$$4x^2 = 13x^2 - 14x + 4$$
$$9x^2 - 14x + 4 = 0$$

The quadratic formula will be used since the expression cannot be factored any further.

$$x = \frac{14 \pm \sqrt{14^2 - 4\cdot 9\cdot 4}}{9\cdot 2} = \frac{7 \pm \sqrt{13}}{9}$$

Since x needs to be less than 0.5, $\frac{7+\sqrt{13}}{9}$ cannot be a solution for the expression because $\frac{7+\sqrt{13}}{9} \approx 1.16$, which is greater than 0.5. Therefore, $x = \frac{7-\sqrt{13}}{9} \approx 0.38$. Substituting the x value in the first curve (y = x) gives $y = \frac{7-\sqrt{13}}{9}$. The intersection point between curve 1 and curve 2 is then $(\frac{7-\sqrt{13}}{9}, \frac{7-\sqrt{13}}{9})$ which is approximately $(0.38, 0.38)$. The corresponding y-value for curve 3 when x is 0.38 is $\frac{1}{2}\sqrt{4 \cdot 0.38^2 - 2 \cdot 0.38 + 1} \approx 0.45$. Therefore, the maximum y value between the three curves is 0.45 when x is equal to 0.38.

Curve 1 & Curve 3 Intersection Point:

$$x = \frac{1}{2}\sqrt{4x^2 - 2x + 1}$$
$$2x = \sqrt{4x^2 - 2x + 1}$$
$$4x^2 = 4x^2 - 2x + 1$$
$$2x = 1$$
$$x = \frac{1}{2}$$

Here, x exceeds the original boundary because it is not less than 0.5. Therefore, the intersection point between curve 1 and curve 3 cannot be used as a solution.

Curve 2 & Curve 3 Intersection Point:

$$\frac{1}{2}\sqrt{13x^2 - 14x + 4} = \frac{1}{2}\sqrt{4x^2 - 2x + 1}$$
$$13x^2 - 14x + 4 = 4x^2 - 2x + 1$$
$$9x^2 - 12x + 3 = 0$$
$$3x^2 - 4x + 1 = 0$$
$$(3x - 1)(x - 1) = 0$$

$(3x - 1) = 0$ gives $x = 1/3$ and $(x - 1) = 0$ gives $x = 1$. However, since x needs to be less than 0.5, 1 cannot be a possible solution for the expression, leaving $x=1/3$ as the only possible option. Substituting the x value to the second curve gives

$$y = \frac{1}{2}\sqrt{4\left(\frac{1}{3}\right)^2 - 2\cdot\frac{1}{3} + 1} = \frac{\sqrt{7}}{6} \approx 0.44.$$

The intersection point between curve 2 and curve 3 is then $(1/3, \sqrt{7}/6)$ which is approximately (0.33,0.44). The corresponding y-value for curve 1 when x is equal to 0.33. Therefore, the maximum y value between the three curves is 0.44 when x is equal to 0.33. The goal is to find the minimum value among the maximum values as indicated in Figure 19.

The minimum value 0.44 occurs when x value is equal to 0.33. This implies that the maximum distance between two points is less than or equal to 0.44 when x is 0.33 no matter which two points are chosen. This shows an absolute improvement of 0.06cm or a percent improvement of 12% compared to the original bound, which is 0.5cm.

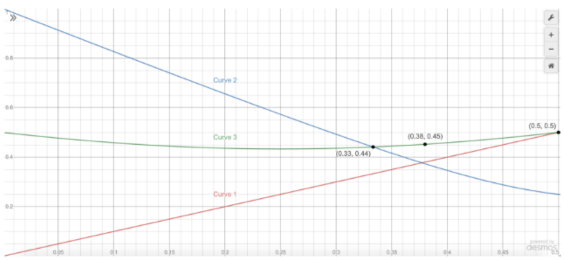

Figure 19