A Hexagon Cutting Problem and its Generalisations

Joel Tan, Matthew Chia, Daniel Leong
NUS High School of Mathematics and Science
January 15, 2016

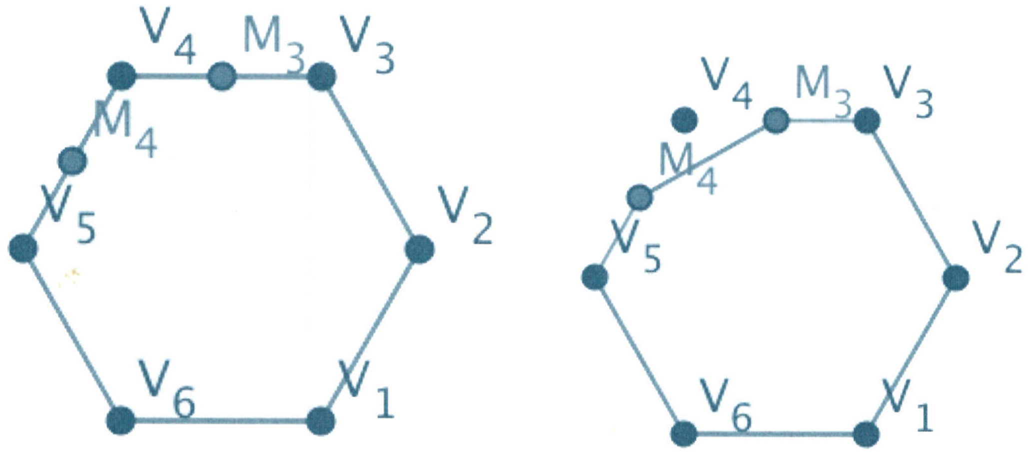
1 Introduction

In this article, we will explore a geometrical problem as follows:

In a convex n-gon, $V_1V_2...V_n$, denote M_i to be the midpoint of V_iV_{i+1} (here $V_1 = V_{n+1}$, $M_1 = M_{n+1}$.). We define the following operation, which we will call a *cut*:
(i) Choose $i \in \{1, 2, ..., n\}$.
(ii) M_iM_{i+1} divides the polygon into two parts: a triangle and an $(n+1)$-gon. We remove (or cut off) the triangle.

The triangle is removed and the $(n+1)$-gon is considered.

For instance, for the hexagon $V_1V_2V_3V_4V_5V_6$ on the following page, a possible cut can be performed by choosing $i = 3$ (shown in (a)), and removing the triangle $\Delta M_3V_4M_4$ (shown in (b)) to obtain a heptagon (shaded brown). If we wish to perform a second operation, we can similarly take the midpoints of any two adjacent sides of this heptagon. Again, the line through these midpoints divide the polygon into a triangle and a octagon. We remove the triangle. Three possible octagons are shown, also shaded brown.



(a) Initial polygon

(b) A possible polygon after an operation

Figure 1: An operation

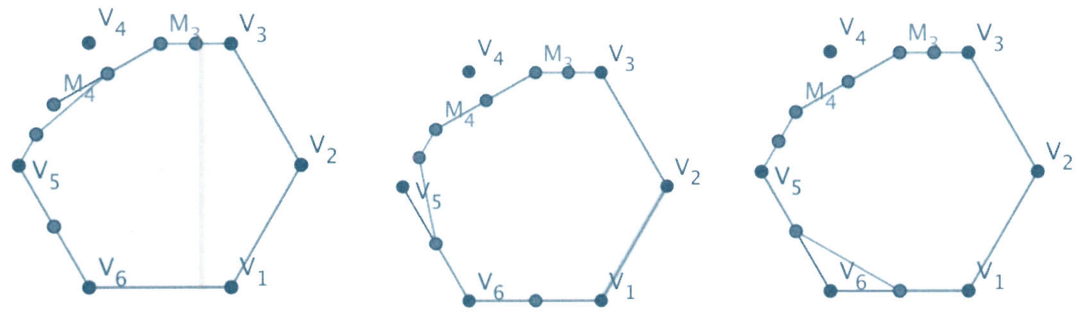


Figure 2: Some of the possible polygons after 2 operations

In Figure 3 on the next page, the cuts shown are not valid: in the figure on the left, the polygon shown is not convex, therefore no operation can be performed. In the figure on the right, it is clearly impossible as this will imply $i = 1$ or $i = 3$. But these would imply that we choose M_1M_2 or M_3M_4, contradiction.

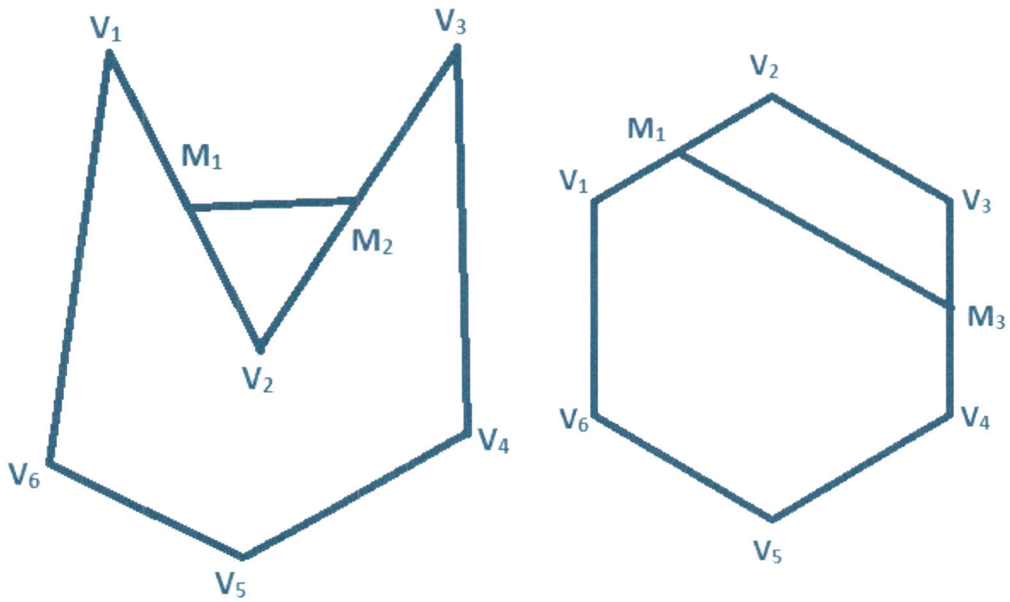


Figure 3: Not Operations

2 Motivation

A similar problem, which first appeared as Problem 4 in the 1997 USA Mathematical Olympiad, asked to prove that regardless of the number of operations done starting from a regular hexagon, the remaining area is equal to at least $\frac{1}{3}$ of the area of the hexagon.

3 Aim of Research

We start from a regular hexagon of side length 1, and area $K_0 = \frac{3\sqrt{3}}{2}$.

After performing a finite number of operations, we wish to determine the minimum possible remaining area, denote it by K.

Afterwards, we will generalise the problem to finding the range of K, starting with a regular n-gon, $n \geq 6$ is an integer.

4 Two Important Lemmas

For convenience, denote $K_0 = \frac{3\sqrt{3}}{2}$.
Denote T_0 to be the initial regular hexagon and T_k to be the polygon formed by an operation on $T_{k-1}, \forall k \in \mathbb{N}$.

Definition 4.1. *For line segments l_1 and l_2, l_1 is part of l_2 if the endpoints of l_1 both lie on l_2.*

Lemma 4.2. *T_k is convex $\forall k \in \mathbb{N} \cup \{0\}$.*

Proof. We proceed by induction. When $k = 0$, it is true.
Suppose it is true for some $k = j \geq 0$, then consider $k = j + 1$. Let $T_j = V_1 V_2 ... V_i$ (here $V_x = V_{x+i}, \forall x \in \mathbb{N}$).

Since T_j is convex, $\angle V_x V_{x+1} V_{x+2} < 180°, \forall x \in \{1, 2, ..., i\}$. Without loss of generality, suppose that in the $(j+1)$th operation, we choose the line l through the midpoints M_1, M_2 of V_1V_2, V_2V_3 respectively. Then $T_{j+1} = V_1 M_1 M_2 V_3 V_4 ... V_i$.

It suffices to consider $\angle V_1 M_1 M_2$ and $\angle V_3 M_2 M_1$ of T_{j+1} to prove T_{j+1} is convex.

$$\angle V_1 M_1 M_2 = 180° - \angle V_2 M_1 M_2 < 180° \tag{1}$$

$$\angle V_3 M_2 M_1 = 180° - \angle V_2 M_2 M_1 < 180° \tag{2}$$

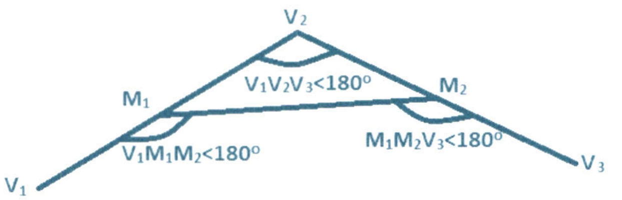


Figure 4: The new polygon is still convex

Hence T_{j+1} is convex and is applicable. By mathematical induction, the lemma is true $\forall k \in \mathbb{N}$. □

Lemma 4.3. *Let $k \in \mathbb{N}$. For each edge E of T_0, there will be an line segment l' such that l' is part of segments E and T_k.*

Proof. We show this by induction on $k \in \mathbb{N}$.

Figure 5: An edge of T_0

Suppose it is true for some $k = l \geq 1$. Consider $k = l + 1$. Consider any edge AB on T_0. Then one of the edges of T_k, name it CD, lies on segment AB. Let P be the midpoint of CD. At most one of $\{CP, PD\}$ is removed when an operation is applied. The remaining segment will be in T_{k+1} and we are done by induction. □

5 An Upper Bound of the Lower Bound

We have written a C++ program to randomly perform operations on a regular n-gon, where $n \geq 3$ is an integer. For a hexagon (which we will be focusing on), the area remaining after cutting randomly and Greedily are 0.81730 and 0.80215 respectively.

6 Geometrical Bounds

6.1 A Lower Bound of $K \geq \frac{1}{2}K_0$

Denote $T_0 = P_1P_2P_3P_4P_5P_6$ and F to be the number of operations. By **Lemma 3.4**, we can choose $Q_i, \forall i \in \{1, 2, ..., 6\}$ such that Q_i is on both P_iP_{i+1} and T_F. (Here $P_{i+6} = P_i, Q_{i+6} = Q_i, \forall i \in \mathbb{N}$.)

Lemma 6.1. *$Q_1Q_2Q_3Q_4Q_5Q_6$ is contained in T_F.*

Proof. This is a consequence of **Lemma 4.2** which states that T_F is convex. □

Denote $[\omega]$ to be the area of a polygon ω.
Let $|P_iQ_i| = a_i, \forall i \in \{1, 2, ..., 6\}$ and let $a_7 = a_1$. The area A cut off from T_0 satisfies

$$A \leq \sum_{i=1}^{6}[\Delta Q_iP_{i+1}Q_{i+1}] = \sum_{i=1}^{6}\frac{1}{2}a_i(1 - a_{i+1})\sin(120^\circ) = \frac{\sqrt{3}}{4}\sum_{i=1}^{6}a_i(1 - a_{i+1}) \tag{3}$$

Denote

$$S = \sum_{i=1}^{6}a_i(1 - a_{i+1}) \tag{4}$$

Then, we have

$$\begin{aligned} -S + \frac{3}{2} = \sum_{i=1}^{6}a_ia_{i+1} - \sum_{i=1}^{6}a_i + \frac{3}{2} &= \sum_{i=1}^{6}(\frac{1}{4} - \frac{1}{2}a_i - \frac{1}{2}a_{i+1} + a_ia_{i+1}) \\ &= \sum_{i=1}^{6}(\frac{1}{2} - a_i)(\frac{1}{2} - a_{i+1}) \end{aligned} \tag{5}$$

However, $a_i, a_{i+1} \in [0, 1]$. Hence

$$|(\frac{1}{2} - a_i)(\frac{1}{2} - a_{i+1})| \leq (\frac{1}{2})(\frac{1}{2}) = \frac{1}{4} \tag{6}$$

and by (6),

$$-S + \frac{3}{2} = \sum_{i=1}^{6}(\frac{1}{2} - a_i)(\frac{1}{2} - a_{i+1}) \geq 6 \times (-\frac{1}{4}) = -\frac{3}{2} \tag{7}$$

with equality occuring iff $\{a_i, a_{i+1}\} = \{0, 1\}, \forall i \in \{1, 2, ..., 6\}$.
Thus

$$S \leq 3 \implies A \leq \frac{3\sqrt{3}}{4} \tag{8}$$

The area of T_0 is $6(\frac{1}{2})(1)(1)(\sin 60^\circ) = \frac{3\sqrt{3}}{2}$ hence the area remaining is

$$K = [T_0] - A = \frac{3\sqrt{3}}{2} - A \geq \frac{3\sqrt{3}}{4} \tag{9}$$

This is at least $\frac{1}{2}$ the area of the initial hexagon.

We know, intuitively, that this is not an optimal bound: The equality cases of (6) state that $\{a_i, a_{i+1}\} = \{0, 1\}, \forall i \in \{1, 2, ..., 6\}$. This is impossible. Thus, we want to find a better bound.

6.2 Tightening the Bound

In this section, we will introduce some terminologies that will be used throughout the rest of this article.

Let the edges of the hexagon be $E_i = P_iP_{i+1}, i \in \{1, 2, ..., 6\}$.

Definition 6.2. *A **dividing segment** is a line segment that divides a polygon into 2 parts during an operation.*

With respect to a polygon $T_k = V_1V_2...V_i$ with $V_{i+1} = V_1$, we assign sets of numbers to dividing segments and edges as follows:

(i) Each edge V_jV_{j+1} of T_k has $\{V_jV_{j+1}\}$ assigned to it.

(ii) Every segment h that is a dividing segment from T_k onwards passes through the midpoints of two edges h_1, h_2 of $T_{k+k'}$ for some $k' \in \mathbb{N} \cup \{0\}$. The set assigned to h is the union of the sets assigned to h_1 and h_2.

(iii) Any part of a segment h has the same set assigned to it as h.

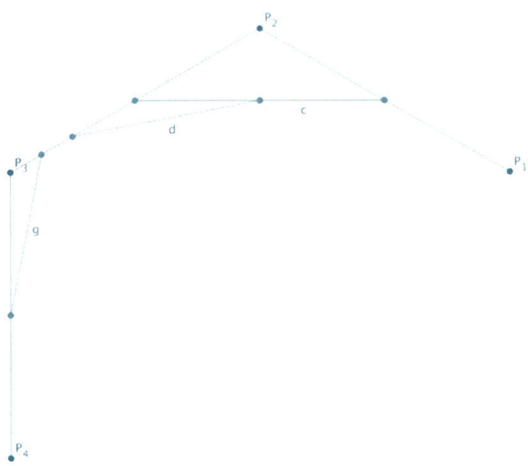


Figure 6: An Example

Take for instance the diagram in Figure 4. c, d, g are dividing segments and have $\{E_1, E_2\}, \{E_1, E_2\}, \{E_2, E_3\}$ assigned to them respectively, with respect to the original hexagon T_0.

Lemma 6.3. *The set assigned to any dividing segment with respect to a polygon $T_k = V_1V_2...V_i$ is in the form $\{V_pV_{p+1}, V_qV_{q+1}\}$ where $p, q \in \{1, 2, ..., i\}, |p-q| \equiv \pm 1 \ (mod\ i)$.*

Lemma 6.4. *A dividing segment with $\{p, q\}$ assigned is only adjacent to dividing segments or edges with $\{p\}, \{q\}$ or $\{p, q\}$ assigned.*

Proof. We prove the above two lemmas simultaneously.

To obtain T_{k+1}, we must join the midpoints of two adjacent sides. Hence the two lemmas are true.

Suppose both statements are true for the $(k+k')th$ dividing segment, $k' \in \mathbb{N}$. Consider the $(k+k'+1)th$ dividing segment. It passes through the midpoints of two adjacent edges of $T_{k+k'}$. Then the union of the sets assigned to these two edges h_1, h_2 have at most 2 elements in the form $\{p, q\}$ where $|p-q| \equiv \pm 1$ (mod i). This follows by the second lemma for the $(k+k')$th dividing segment. Hence the set assigned to the $(k+k'+1)th$ dividing segment has at most 2 elements differing by ± 1 (mod i) assigned to it. However, it cannot have only 1 element unless it is an edge. Thus we have proven the first statement.

Note that the edges adjacent to the $(k+k'+1)th$ dividing segment are parts of l_1 and l_2. Thus the second statement must be true. By mathematical induction, both statements are always true for all dividing segments. □

Definition 6.5. *Define a dividing segment to be (p, q)-originating with respect to T_k if the set assigned to it is $\{p, q\}$ with respect to T_k.*

Lemma 6.6. *An (V_pV_{p+1}, V_qV_{q+1})-originating dividing segment with respect to T_k lies entirely in the triangle t formed by points $V_p, V_{p+1}, V_q, V_{q+1}$.*

Proof. The statement is true for the $(k+1)$th dividing segment.

Suppose it is also true for the $(k+1)$th, $(k+2)$th, ..., $(k+k')$th dividing segment for some $k' \in \mathbb{N}$. The endpoints of the $(k+k'+1)$th dividing segment are on one of the edges V_pV_{p+1}, V_qV_{q+1} or on (V_pV_{p+1}, V_qV_{q+1})-originating dividing segments. Each of these are contained entirely in t by the inductive hypothesis. Since the endpoints of the $(k+k'+1)$th dividing segment are both in t, hence the entire dividing segment is in t. Hence it is true for the $(k+1)$th, $(k+2)$th, ..., $(k+k'+1)$th dividing segment and we are done by mathematical induction. □

6.3 A Lower Bound of $\frac{49}{85}K_0$

Consider J, the midpoint of P_1P_3 and I, the point on segment P_2J such that $|IJ| = 0.4|P_2J|$. We claim that all (E_1, E_2)-originating dividing segments with respect to T_0 lie outside ΔP_3IP_1. By symmetry, this will imply that all dividing segments lie outside the dodecagon, coloured brown, as shown below.

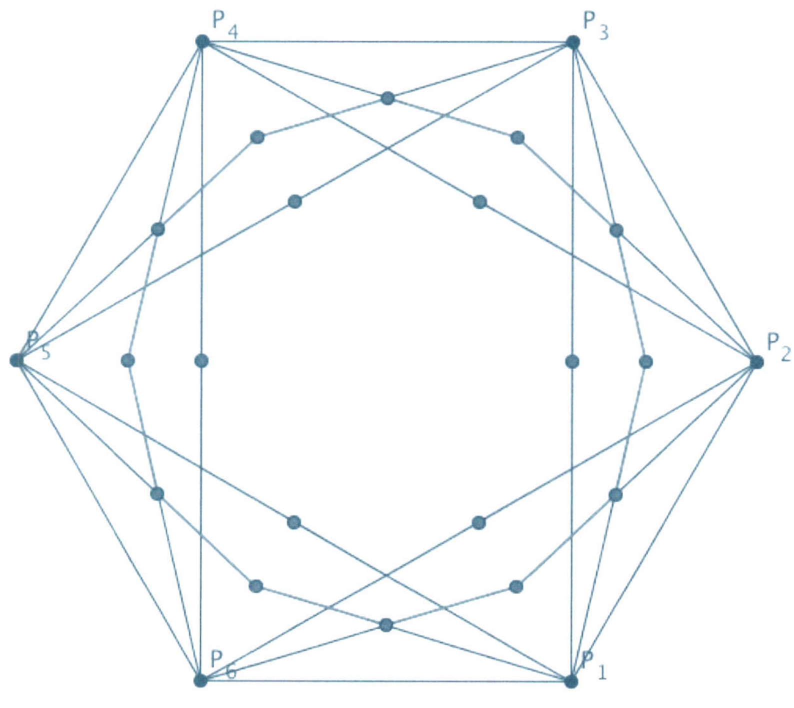


Figure 7: Dodecagon

Suppose otherwise that some part a dividing segment lies inside ΔP_3IP_1. Consider the first time, during the kth operation, that a dividing segment is (E_1, E_2)-originating with respect to T_0.

Suppose the segment is DE, with D on P_1P_2 and E on P_2P_3. Then

$$\frac{DP_2}{P_1P_2} \leq \frac{1}{2}, \frac{EP_2}{P_3P_2} \leq \frac{1}{2} \tag{10}$$

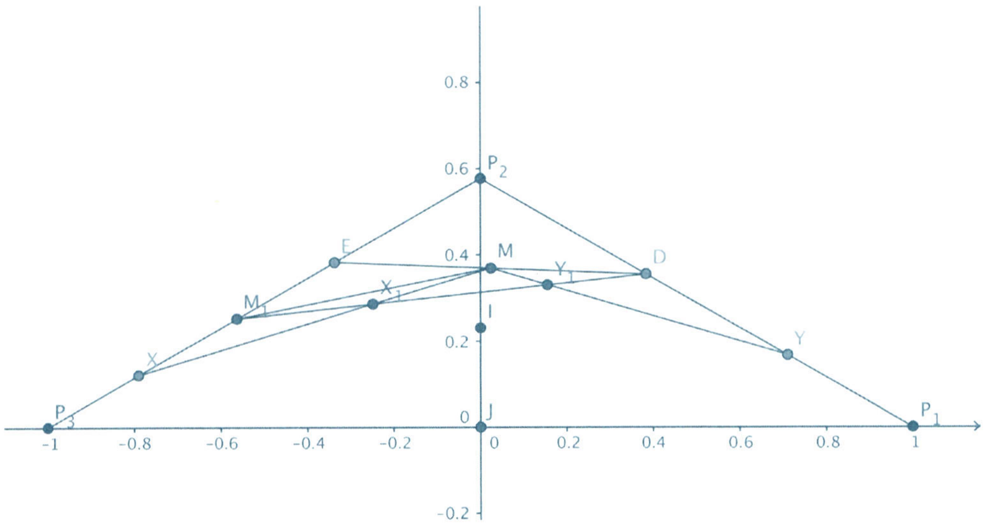


Figure 8

Definition 6.7. *The remaining part of a segment h after a sequence operations is the part of h left after these operations.*

If there are no other (E_1, E_2)-originating segments, we are done. Otherwise, consider the next time a dividing segment is (E_1, E_2)-originating. Let the remaining parts of EP_3 and DP_1 be segments be EX and DY. The midpoint M of DE and the midpoint M_1 of either EX or DY form the dividing segment. Without loss of generality, we assume that M_1 is on EX.

Any other (E_1, E_2)-originating dividing segments must be either (XM_1, M_1M), (M_1M, MD) or (MD, DY)-originating segments with respect to the current polygon.

Thus, all (E_1, E_2)-originating segments with respect to T_0 are in either ΔXM_1M, ΔM_1MD or ΔMDY by **Lemma 4.3**. Let $X_1 = MX \cap M_1D, Y_1 = MY \cap M_1D$.

On the xy-coordinate plane, denote J to be the origin, $P_3 = (-1, 0), P_1 = (1, 0)$.

Consider $T' = P_3XX_1Y_1YP_1$. Note that X, which is on T', is outside of ΔP_3IP_1. Thus T' intersects this triangle at least twice on points not lying on P_1P_3. However, T' cannot intersect an edge of the triangle more than once, hence it intersects P_3I and P_1I once each. Thus I is in the interior of $P_2XX_1Y_1Y$. Let P_2 have y-coordinate y_1.

6.3.1 Case 1: I is inside triangle EP_2D.

E and D have y-coordinates at least $\frac{1}{2}y_1$, hence I, which has y-coordinate $0.4y_1$, must have y-coordinate larger than $\frac{1}{2}y_1$, a contradiction.

6.3.2 Case 2: I is in ΔXEM.

Then I is in ΔP_3EM.

Let $E = (-1 + a, \frac{1}{\sqrt{3}}a), D = (1 - b, \frac{1}{\sqrt{3}}b)$ for some $a, b \in [\frac{1}{2}, 1]$.

$$M = (\frac{a-b}{2}, \frac{1}{\sqrt{3}}\frac{a+b}{2}) \tag{11}$$

I is on the same side of P_3M as E, hence P_3I intersects segment EM. Let the intersection be L.

Line P_3I has equation $y = 0.4y_1x + 0.4y_1$.

L lies on segment DE, hence the y-coordinate is greater than $0.5y_1$ hence the x-coordinate is greater $\frac{0.5y_1 - 0.4y_1}{0.4y_1} = \frac{1}{4}$. This implies that M has x-coordinate greater than $\frac{1}{4}$.

Thus $\frac{a-b}{2} > \frac{1}{4} \implies a - b > \frac{1}{2}$, contradiction as $a, b \in [\frac{1}{2}, 1]$.

6.3.3 Case 3: I is in ΔDMY

By symmetry, from Case 2 we deduce that I also cannot be in ΔDMY.

6.3.4 Case 4: I is in ΔM_1ED

Denote $y(P)$ to be the y-coordinate of any point P.

Let M_2 to be the midpoint of P_2P_3, M_3 to be the midpoint of M_2P_3 and M_4 to be the midpoint of P_2P_1 as shown in the diagram on the next page.

$$y(M_1) = \frac{y(E)+y(X)}{2} \geq \frac{1/2y(P_2)+0}{2} = \frac{1}{4}y(P_2) = y(M_3) \tag{12}$$

hence P_3, M_3, M_1, P_2 are on P_2P_3 in that order.

I and P_2 are on the same side of M_1D and hence are on the same side of M_3D.

P_2, D, M_4, P_1 are on P_2P_1 in that order hence I and P_2 are on the same side of M_3M_4. Thus M_3M_4 intersects the y-axis at a point Q with y-coordinate strictly smaller than $0.4y(P_2)$.

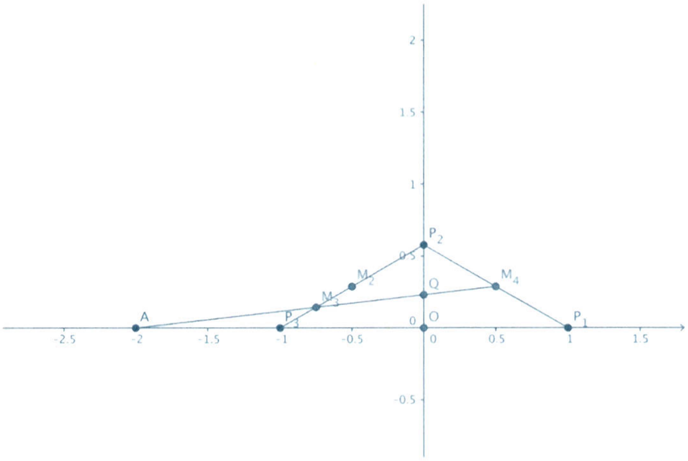


Figure 9

By Menelaus' Theorem [1]

$$1 = \frac{|P_1A|}{|AP_3|}\frac{|P_3M_3|}{|M_3P_2|}\frac{|P_2M_4|}{|M_4P_1|} = \frac{|P_1A|}{|AP_3|}\frac{1}{3} \implies |P_1A| = 3|P_3A| \tag{13}$$

Thus $|AP_3| = |P_3O| = |OP_1|$ where O is the origin, as seen in the above diagram.
By Menelaus' Theorem again,

$$1 = \frac{|OA|}{|AP_3|}\frac{|P_3M_3|}{|M_3P_2|}\frac{|P_2Q|}{|QO|} = \frac{|P_2Q|}{|QO|}\frac{2}{3} \implies |P_2Q| = \frac{3}{2}|QO| \implies |OQ| = \frac{2}{5}|OP_2| \tag{14}$$

Thus $Q = I$, a contradiction.

[1] In a triangle ΔABC and a transversal intersecting AB, BC, CA at D, E, F respectively, $\frac{|AD|}{|DB|}\frac{|BE|}{|EC|}\frac{|CF|}{|FA|} = 1$.

6.3.5 Finding the area of the dodecagon

The claim that $P_2XX_1YY_1$ is always outside the dodecagon has been proven, so we have a lower bound of the area of the dodecagon.

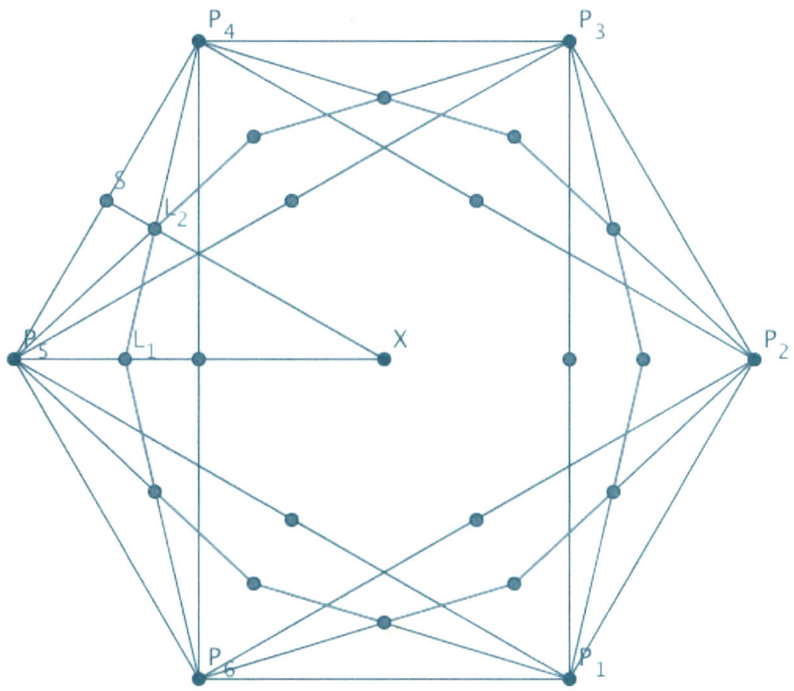


Figure 10

Let the area of the dodecagon be A. We define X, S, M', L_1, L_2 to be the centre of T_0, the midpoint of P_5P_4, the midpoint of P_5X, the point on segment P_5M' with $|L_1M'| = 0.4|P_5M'|$ and $L_1P_4 \cap SX$ respectively.

Note that $|P_5L_1| = \frac{3}{10}|P_5X|$ hence $\frac{|P_5L_1|}{|L_1X|} = \frac{3}{7}$.

Using Menelaus' Theorem once more,

$$1 = \frac{|SL_2|}{|L_2X|}\frac{|XL_1|}{|L_1P_5|}\frac{|P_5P_4|}{|P_4S|} = \frac{|SL_2|}{|L_2X|}\frac{14}{3} \implies |SL_2| = \frac{3}{14}|L_2X| \qquad (15)$$

$$\implies \frac{|L_2X|}{|SX|} = \frac{14}{17}$$

Hence

$$\frac{A}{K_0} = \frac{[\Delta L_2L_1X]}{[\Delta P_5XS]} = \frac{\frac{1}{2}|L_2X||L_1X|\sin(30^\circ)}{\frac{1}{2}|SX||P_5X|\sin(30^\circ)} = \frac{|L_2X|}{|SX|}\frac{|L_1X|}{|P_5X|} = \frac{49}{85} \qquad (16)$$

This is 0.576 to three decimal places.

In fact, we could have considered more individual cases than the 4 given here. However, this will be tedious. Hence, we introduce the *Polygon Array* as shown in the next section.

7 Polygon Arrays and Trees

In this section, we assume the area K_0 of the hexagon T_0 to be 24 for convenience instead.

7.1 Polygon Arrays

Definition 7.1. *$[\Delta V_{x-1}V_xV_{x+1}]$ is the V_x-subarea of the polygon $V_1V_2...V_i$ with indices taken modulo i.*

Definition 7.2. *We define the polygon array J_k of a polygon T_k to be the array as follows:*
Given a polygon $V_1V_2...V_i$ with indices taken modulo i, we have the polygon array to be

$$J_k = [f(\Delta V_iV_1V_2), f(\Delta V_1V_2V_3), f(\Delta V_2V_3V_4), ..., f(\Delta V_{i-1}V_iV_1)] \quad (17)$$

where $f(\Delta V_{x-1}V_xV_{x+1}) = -log_2([\Delta V_{x-1}V_xV_{x+1}]) + 2, \forall x \in \{1, 2, ..., i\}$

Consider T_0 as an example.

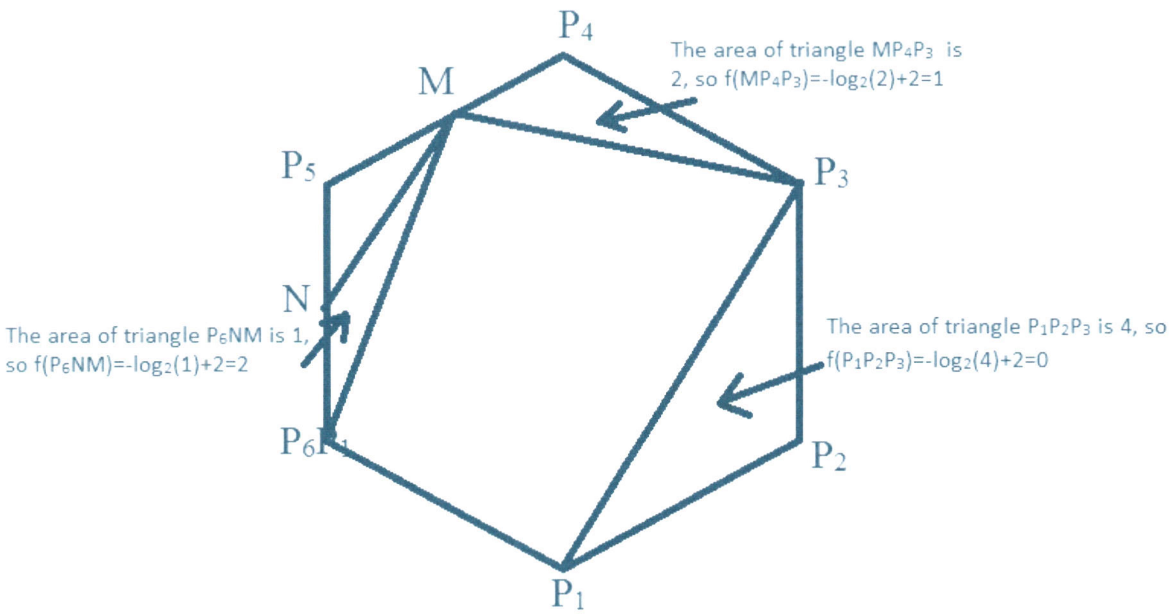


Figure 11: T_0

The initial hexagon's polygon array would thus be $[0, 0, 0, 0, 0, 0]$, representing $V_1, V_2, V_3, V_4, V_5, V_6$-subareas respectively.

If F is the number of operations, we define a sequence of polygon arrays $J_1, J_2, ...J_F$ where J_k is the polygon array of $T_k, k \in \{1, 2, ..., F\}$.

Definition 7.3. *The jth and j'th elements in a polygon array J_k are consecutive iff $|j - j'| = 1$ or $|J_k| - 1$, where $|J_k|$ is the cardinality of J_k.*

Definition 7.4. *A subarray of a polygon array $[a_1, a_2, ..., a_N]$ is $[a_{k_1}, a_{k_1+1}, a_{k_1+2}, ..., a_{k_2}]$ for some $k_1 < k_2$ with $k_1 \le N, k_2 < k_1 + N$. Again, indices are taken modulo n.*

Lemma 7.5. *Let k be any nonnegative integer. J_{k+1} is formed by replacing a three-element subarray $[a, b, c]$ of J_k by the four-element subarray $[a + 1, b + 2, b + 2, c + 1]$.*

Proof. Consider the polygon $T_k = V_1V_2...V_i$. Without loss of generality, suppose that the line segment joining the midpoints M_2, M_3 of V_2V_3 and V_3V_4 was the $(k+1)th$ dividing segment. This means $T_{k+1} = V_1V_2M_2M_3V_4...V_i$.

The $V_1, V_5, V_6, ..., V_i$-subareas of J_k are equal to the $V_1, V_5, V_6, ..., V_i$-subareas of J_{k+1}.

Let the second, third and fourth elements of J_k be $[a, b, c]$.

The V_2, V_4-subareas of J_{k+1} is half the V_2, V_4-subareas J_k respectively and the M_2, M_3-subareas of J_{k+1} are a quarter that of the V_3 subarea of J_k as shown in the figure below.

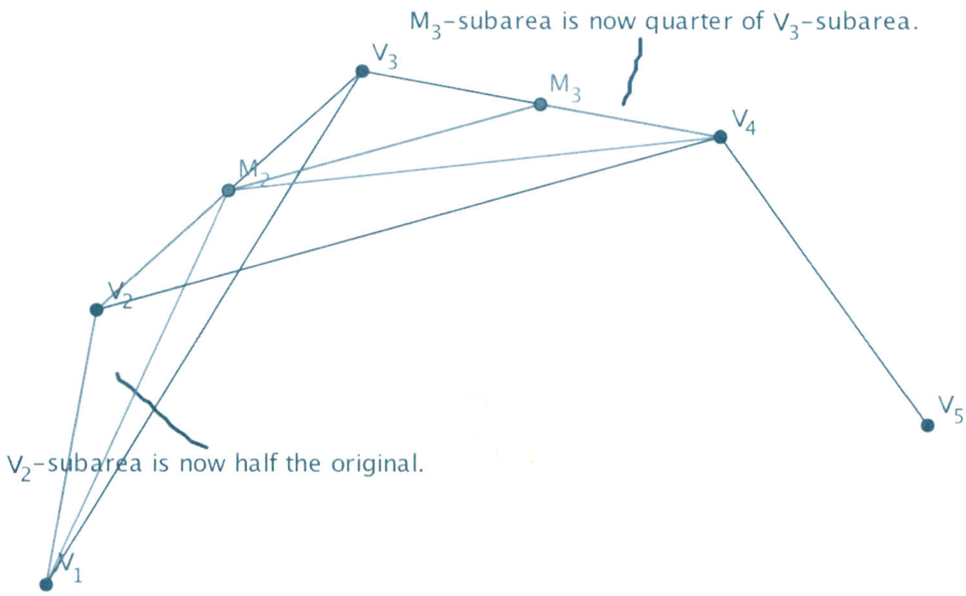


Figure 12: Subareas

By the definition of polygon arrays, we obtain the second to fifth elements of J_{k+1} as $[a+1, b+2, b+2, c+1]$. The conclusion follows. □

As in the above proof, if $[a, b, c]$ is changed to $[a+1, b+2, b+2, c+1]$ in the kth operation, we say the kth operation is $[a, b, c] \to [a+1, b+2, b+2, c+1]$.

Lemma 7.6. *An operation* $[a, b, c] \to [a+1, b+2, b+2, c+1]$ *causes the area of the polygon to decrease by* $\frac{1}{2^b}$.

Proof. The area subtracted is $\frac{1}{4}$ of the b-subarea which is $\frac{1}{2^b}$ by definition. □

7.2 Trees

We construct a tree with $F+2$ layers numbered $-1, 0, ..., F$ as follows where F is the number of operations:

(i) Each element of the polygon array is represented by a vertex. This vertex has the element labelled on it.

(ii) There is a central vertex C (layer -1) connecting to 6 vertices, labelled 0, ($0th$ layer) representing the elements from the initial polygon array.

(iii) All elements from a polygon array J_k are represented by vertices in the kth layer.

(iv) For every operation, as proven, the subarray $[a, b, c]$ of J_k becomes $[a+1, b+2, b+2, c+1]$ in J_{k+1} if an operation is applied to b. We connect the vertex with label b to the vertices with label $b+2$. For the other vertices representing elements of J_k, connect it to the vertex representing corresponding element in J_{k+1}. (For example, a is connected to $a+1$, c is connected to $c+1$.)

(v) If vertex v has label l, we say that the corresponding vertex of l is v.

An example of such a tree is shown below.

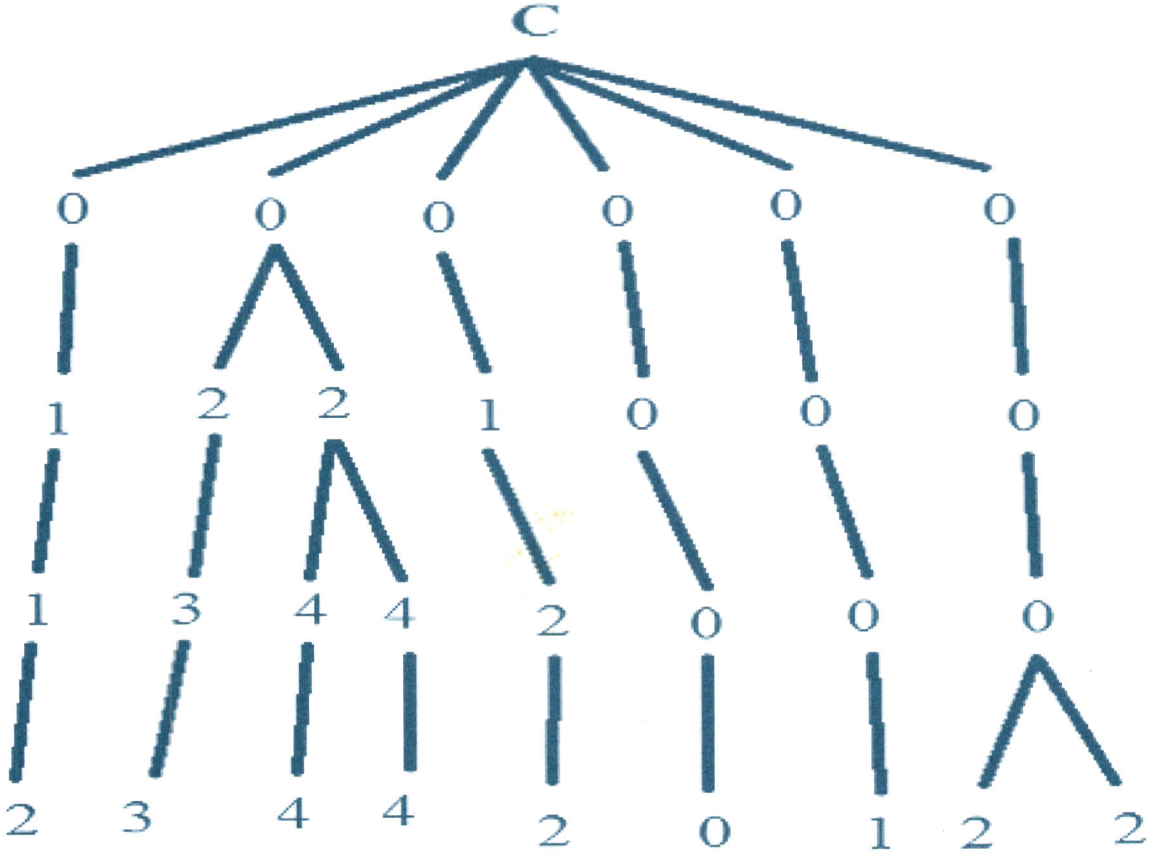


Figure 13: Trees

7.3 Important Definitions

Definition 7.7. *We call a vertex V_1 V_2-originating if we can find a path from V_1 to V_2, in the tree, with the V_1 having a layer number not less than that of V_2. (For example, if V_2 is in the 6th layer, V_1 could possibly be in the 6th or 7th layer but not the 5th.)*

Denote $J[k][j]$ to be the $j(\mod |J_k|)$th **vertex** from the left in the tree on the kth layer for $k \in \{1, 2, ..., F\}, j \in \mathbb{Z}$; and $l(V)$ to be the **label** of a vertex V.

Definition 7.8. *Two vertices $J[k_1][j_1]$ and $J[k_2][j_2]$ are consecutive iff $k_1 = k_2 = k'$ for some k' and $|j_1 - j_2| \equiv \pm 1 \mod k'$.*

Note that the label, not the vertex, is a numerical value. The vertex numbering only determines the position in the tree.

In the kth operation, only one vertex V in the kth layer has two edges connecting to two vertices in the next layer. We say the kth operation is done on a vertex V.

Definition 7.9. *The operation sequence of a sequence of k operations $\{A_1, A_2, ..., A_k\}$ is the sequence satisfying that the ith operation was on vertex $A_i, \forall i \in \{1, 2, ..., k\}$*

Definition 7.10. *The optimal operation sequence is the operation sequence $\{A_1, A_2, ..., A_k\}$ such that $\sum_{i=1}^{k} \frac{1}{2^{l(A_i)}}$ is maximised. (i.e. The maximum possible area is subtracted.) This sequence might not be unique, but it clearly exists.*

Definition 7.11. *The shape of a tree is a tree with the labels removed.*

For example, the following trees have the same shape. They are not identical as their labels are different. However, if we remove the numerical labels, they would be identical.

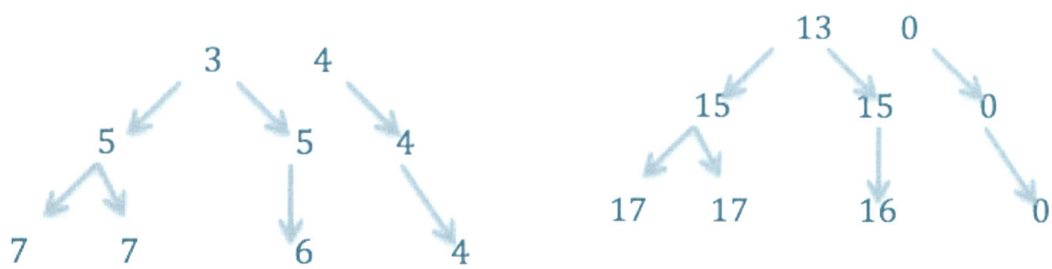


Figure 14: 'Shapes' of trees

8 More Important Lemmas

We now consider the tree formed by performing F operations.

Lemma 8.1. *Consider a vertex a in the tree. Let the operations on a-originating vertices be on vertices $\{q_1, q_2, ..., q_n\}$. Then $S = \sum_{i=1}^{n} \frac{1}{2^{l(q_n)}} < \frac{1}{2^{l(a)-1}}$.*

Proof. If, instead of the operation $[a, b, c] \to [a+1, b+2, b+2, c+1]$ we perform $[a, b, c] \to [a, b+2, b+2, c]$, while still removing $\frac{1}{2^b}$ from the area, S will clearly increase. Thus we consider the latter. We call it a *pseudo-operation*.

Note that a pseudo-operation on a vertex is independent of the pseudo-operations on neighbouring vertices. In other words,

(i) If a, b are distinct vertices in the kth layer of the tree, and we apply a pseudo-operation on an a-originating vertex then on a b-originating vertex, we can do the reverse and keep both the sum and the resultant polygon array the same.

(ii) If we apply a pseudo-operation on an a-originating vertex, it will not change the value of b.

Thus, S is less than the sum attained when we continuously apply the pseudo-operations whenever possible. This can be represented as in the following diagram with every two arrows pointing from a vertex representing one pseudo-operation:

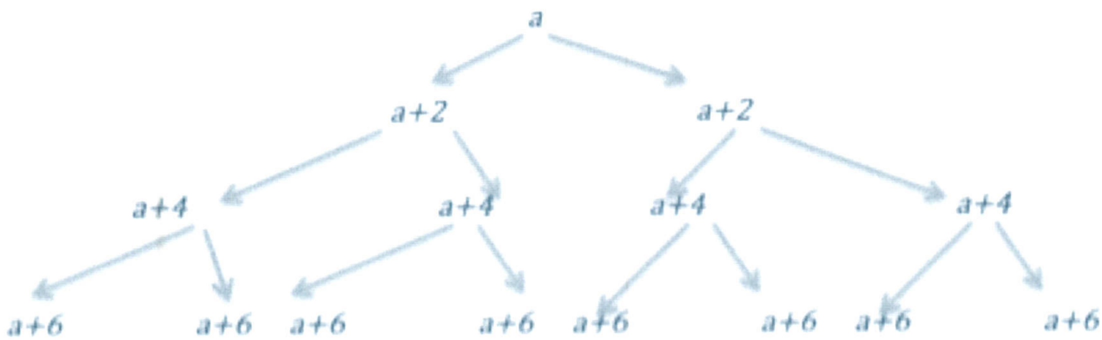


Figure 15: 'Pseudo-operation's

S thus will be strictly less than $\sum_{i=0}^{\infty} 2^i \frac{1}{2^{a+2i}} = \sum_{i=0}^{\infty} \frac{1}{2^{a+i}} = \frac{1}{2^{a-1}}$.

□

Lemma 8.2. *Consider k, c such that $l(J[k][c]) = a+1, l(J[k][c+1]) = a+2$ for some non-negative integer a. The operations on $J[k][c]$ and $J[k][c+1]$-originating vertices removes at most $\frac{1}{2^a}$ from the area.*

Proof. Without loss of generality, suppose $c = 1$.

It suffices to consider only operations on $J[k][1]$ and $J[k][2]$-originating vertices.

Suppose we could subtract more than $\frac{1}{2^a}$ from the area.

Now, suppose the first operation was applied on $J[k][2]$. Then $[a+1, a+2] \to [a+2, a+4, a+4]$. By **Lemma 8.1**, the maximum area deducted is less than $\frac{1}{2^{a+2}} + \frac{1}{2^{a+1}} + 2\frac{1}{2^{a+3}} = \frac{1}{2^a}$, a contradiction.

Thus $[a+1, a+2] \to [a+3, a+3, a+3]$. Now, the subarray consists of the first, second and third elements of J_{k+1}.

Suppose the second operation was applied on $J[k+1][2]$. Then $[a+3, a+3, a+3] \to [a+4, a+5, a+5, a+4]$. By **Lemma 8.1**, the maximum area deducted is less than $\frac{1}{2^{a+1}} + \frac{1}{2^{a+3}} + 2\frac{1}{2^{a+4}} + 2\frac{1}{2^{a+3}} = \frac{1}{2^a}$, another contradiction.

Thus suppose otherwise. We must perform an operation on $J[k+1][1]$ or $J[k+1][3]$. By symmetry, we may just consider $J[k+1][1]$, and the sequence of operations is $[a+1, a+2] \to [a+3, a+3, a+3] \to [a+5, a+5, a+4, a+3]$. This final subarray consists of the first four elements of J_{k+2}.

Let X be the maximum possible area deducted from $J[k][1]$ and $J[k][2]$-originating vertices.

Then the maximum possible area deducted from operations on $J[k+2][3]$ and $J[k+2][4]$-originating vertices is less than $\frac{X}{4}$. Also, by **Lemma 8.1**, the maximum possible area deducted from both $a+5$-originating vertices is less than $2\frac{1}{2^{a+4}} = \frac{1}{2^{a+3}}$.

Hence,

$$X < \frac{X}{4} + \frac{1}{2^{a+3}} + \frac{1}{2^{a+3}} + \frac{1}{2^{a+1}} = \frac{X}{4} + \frac{3}{2^{a+2}} \tag{18}$$

Solving, we obtain $X \leq \frac{1}{2^a}$. This is a contradiction. □

Remark 8.3. *Using these lemmas, we can find a tighter lower bound. Consider $[0,0,0,0,0,0] \to [1,2,2,1,0,0,0]$ without loss of generality. Dividing this into subarrays $[1,2], [2,1], [0], [0], [0]$, we obtain an upper bound of area deducted as the sum of*

(i) $\frac{1}{2^0}$ (from first operation);

(ii) $\frac{1}{2^0}$, (from $[1,2]$ by the above lemma);

(iii) $\frac{1}{2^0}$ (from $[2,1]$ similarly)

(iv) $3 \times \frac{1}{2^{-1}}$ (from $[0], [0], [0]$).

The sum is 9, hence we have a bound of $\frac{5}{8}$ of the initial area of the hexagon left after any finite number of operations.

The Greedy Algorithm involves removing the triangle with the largest possible triangle at each operation. We claim that this is the optimal way to remove as much area as possible for all F operations, and we provide the proof in the following sections.

9 Optimal Operation Sequences

Lemma 9.1. *There exists an optimal operation sequence* $\{A_1, A_2, ..., A_F\}$ *such that the labels on vertices* $A_1, A_2, ..., A_F$ *form a non-decreasing sequence. We call such a sequence a* **good** *sequence.*

Proof. Suppose otherwise. For any optimal operation sequence $\{A_1, A_2, ..., A_F\}$ that is not good, consider the value p, with $1 \leq p \leq F-1$ such that $l(A_p) > l(A_{p+1})$.

Let $J_{p-1} = [b_1, b_2, ..., b_N]$

Suppose the pth operation was operated on vertex $J[p-1][3]$ without loss of generality, then

$$J_p = [b_1, b_2+1, b_3+2, b_3+2, b_4+1, b_5, ..., b_N] \tag{19}$$

The $(p+1)th$ operation was operated on a vertex with label less than b_3, hence A_{p+1} cannot have label b_3+2. Hence the label can be represented as $b_t + t_1$ where $t_1 \in \{0, 1\}, t \in \{1, 2, ..., N\}$.

If the operation was on $J[p][2]$ or $J[p][5]$, we can assume, without loss of generality, that it was on $J[p][2]$. Then the pth and $(p+1)th$ operations were as follows:

$$[b_1, b_2, ..., b_N] \to [b_1, b_2+1, b_3+2, b_3+2, b_4+1, b_5, ..., b_N] \tag{20}$$
$$\to [b_1+1, b_2+3, b_2+3, b_3+3, b_3+2, b_4+1, b_5, ..., b_N] \tag{21}$$

with $\frac{1}{2^{b_3}} + \frac{1}{2^{b_2+1}}$ subtracted from the area. We call this *Situation 1.*

On the other hand, if we had applied the following operations:

$$[b_1, b_2, ..., b_N] \to [b_1+1, b_2+2, b_2+2, b_3+1, b_4, ..., b_N] \tag{22}$$
$$\to [b_1+1, b_2+2, b_2+3, b_3+3, b_3+3, b_4+1, b_5, ..., b_N] \tag{23}$$

with $\frac{1}{2^{b_2}} + \frac{1}{2^{b_3+1}}$ subtracted from the area. We call this *Situation 2.* From these first two operations in both situations, we find that Situation 2 removes $\frac{1}{2^{b_2+1}} - \frac{1}{2^{b_3+1}}$ more area than in Situation 1. We also note that the only two elements that are different, comparing both arrays, are the second and fifth elements.

From here, we consider the trees for both polygon arrays. Call them H_1 and H_2 for the first and second situations respectively. In H_2, denote $J'[k][c]$ to be the cth vertex on the kth layer.

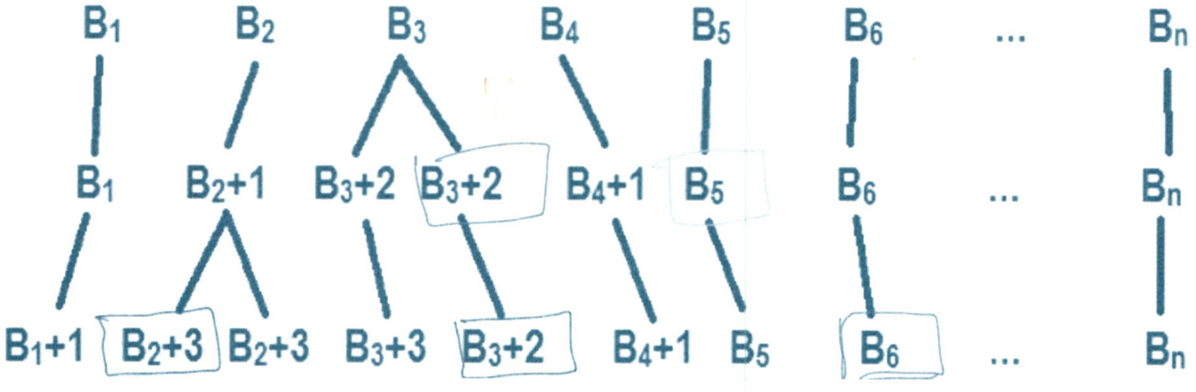


Figure 16: First Situation

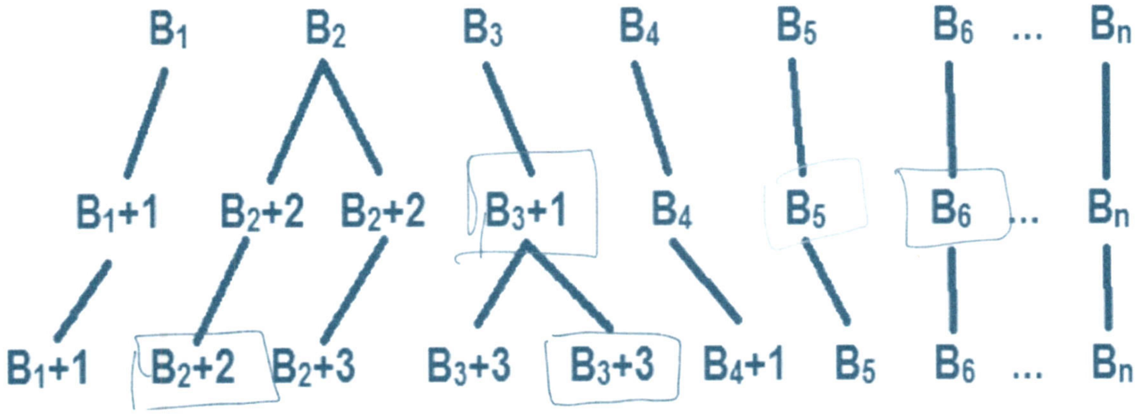


Figure 17: Second Situation

Definition 9.2. *Consider two vertices $A = J[k_1][c_1]$ (in H_1) and $B = J'[k_2][c_2]$ (in H_2).*
*(i) A and B are **similar** iff $k_1 = k_2 = k'$ for some k' and $c_1 \equiv c_2(\mod k')$.*
*(ii) A and B are **constant** iff they are similar and $l(A) = l(B)$.*

For example, the two vertices in orange are constant, those in green are similar, while those in brown are neither.

We then consider the other $F - p - 1$ operations in Situation 1. Consider these operations applied in Situation 2 instead. In other words, after we apply the operations on both trees, H_1 and H_2 have the same shape after the $(p+1)th$ layer.

Let $x = b_2 + 2, y = b_3 + 2$. Then $x < y$.

Now, if on both trees we conduct operations on two constant vertices, there will be no difference in the area subtracted in both situations due to these operations.

Thus we only consider the operations originating from the pairs of vertices that are not constant: in this case, the two pairs

(i) $(J[p+1][2], J'[p+1][2])$, and (ii) $(J[p+1][5], J'[p+1][5])$.

(Their corresponding labels are $x+1, x$ and $y, y+1$ for these pairs respectively.)

Consider those vertices originating from the vertices in the second pair. Any vertices originating from $J'[p+1][5]$ in H_2 will clearly be 1 more than the vertex similar to it and originating from $J[p+1][5]$ in H_1. Hence, suppose we did operations on vertices with labels $z_1, z_2, ...z_n$ in H_1, that originate from $J[p+1][5]$. Then in H_2, we are doing operations on vertices with labels $z_1+1, z_2+1, ..., z_n+1$, originating from $J'[p+1][5]$. The difference in area caused by these operations is thus $\frac{1}{2}\sum_{i=1}^{n}\frac{1}{2^{z_i}}$.

By **Lemma 8.1**, this is less than $\frac{1}{2^y} \le \frac{1}{2^{b_2+2}} \le \frac{1}{2^{b_2+1}} - \frac{1}{2^{b_3+1}}$. Thus the second situation results in at least $\frac{1}{2^{b_2+1}} - \frac{1}{2^{b_3+1}} - \frac{1}{2^y} \ge 0$ more area subtracted than the first.

Suppose we did an operation on a vertex a in H_1 originating from $J[p+1][2]$, and also on the vertex in H_2 originating from $J'[p+1][2]$, similar to a. The area subtracted in the second situation will be more than in the first situation due to these operations.

However, this means that considering all operations, the second situation results in more area subtracted than the first, contradicting the optimality of the optimal sequence.

Therefore, the $(p+1)$th operation is on a vertex not consecutive to $J[p][3]$ or $J[p][4]$ which have label $b_3 + 2$. This means the order of these operations does not matter, and we can swap the pth and $(p+1)th$ operations. Now, $l(A_p) < l(A_{p+1})$.

It suffices to show that these swappings will not go on indefinitely. This will result in a good optimal operation sequence.

Define $S = \sum_{i=1}^{F} i \times l(A_i)$, and H to be the largest label of the vertices in the operation sequence. If we swap 2 vertices $A_y > A_z$ with $y < z$ in the sequence, then S increases by $(z-y)(l(A_y) - l(A_z)) > 0$ while H remains constant.

However, $S \le \sum_{i=1}^{F} iH = \frac{F(F+1)}{2}H$. Since each valid swapping of two non-equal elements in the operation sequence cause S to increase by at least 1, and S is bounded, S will remain constant after some point, and these swappings will end. We then get a non-decreasing optimal sequence.

□

10 Using Greedy Algorithm

We now consider the sequence of operations that results in a good optimal operation sequence.

We prove by induction that we should apply the kth operation on the vertex in the kth layer with minimal label, $\forall k \in \{1, 2, ..., F\}$.

When $k = F$, since it is the final operation, we can take the minimal label without affecting any other operations.

Suppose it is true for $k = m, 1 < m \leq F$. Consider $k = m - 1$. Let $M = J[m-1][c]$ be the such that its label is the minimum element in the polygon array J_{m-1}. Let the 2 vertices consecutive to it be a, c with $a = J[m-1][(c-1) \mod |J_{m-1}|], c = J[m-1][(c+1) \mod |J_{m-1}|]$.

Lemma 10.1. *The $(m-1)th$ operation is on either a or c.*

Proof. Suppose otherwise. Then the $(m-1)th$ operation is not conducted on a, M or c. It is also not operated on any vertex with label equal to that of M. Thus there will still be an element equal to $l(M)$ in J_{m-1}.

By inductive hypothesis, the mth operation is conducted on the vertex with minimal label. Thus it must be on M. However, the $(m-1)th$ operation was conducted on an vertex with label greater than that of M. This is a contradiction as the operation sequence is good. □

Without loss of generality, suppose the operation was conducted on a.

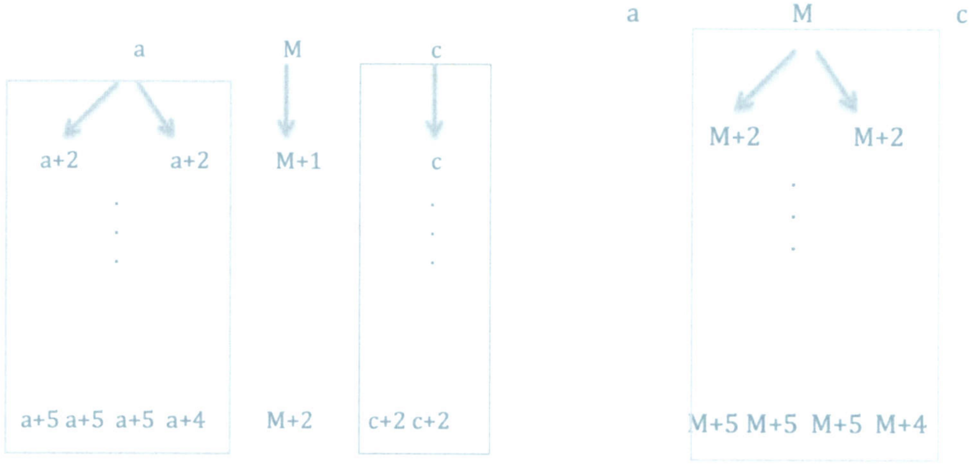
Let k_0 (if exists) to be the minimum positive integer such that a vertex with label $l(M) + k_0$ is M-originating and operated on in some operation.

Lemma 10.2. *k_0 exists.*

Proof. Suppose it does not exist. Let the area cut off due to operations on a, M and c-originating vertices be K_a, K_M and K_c respectively. Denote $D = max(K_a, K_c)$. Without loss of generality, let $D = K_a$.

Consider H_3, the part of the tree formed by a-originating vertices only. Then we apply operations on M such that now, the part of the tree H_4 formed by M-originating vertices has the same shape as in H_3.

An example of how we do this is shown in the two graphs below.



(a) The original sequence of operations

(b) The new sequence of operations

Figure 18

We perform the new sequence of operations such that the two trees in the orange and brown boxes have the same shape *(Refer to Definition 6.11)*.

The area that is subtracted due to operations on M-originating vertices is now $2D \geq K_a + K_c$. By the optimality of the operating sequence, $2D = K_a + K_c \implies K_a = K_c$.

(i) If no operation was applied on c-originating vertices, then $K_a = K_c \implies$ no operation has been applied on a-originating vertices. This is a contradiction.

(ii) If operations were applied on c-originating vertices, then we have found a shorter operation sequence that removes an area not less than the original operation sequence. Hence the original operation sequence is not optimal. This is a contradiction. □

Let r be the vertex $J[m-1][(c-2) \mod |J_{m-1}|]$.

Clearly, $l(a) > l(M)$ otherwise an operation is applied on a vertex with same label as M. If $l(c) = l(M)$, then by inductive hypothesis c should be operated on the mth operation, contradicting that the optimal operating sequence is non-decreasing. Thus both $l(a), l(c) \geq l(M) + 1$.

Let R, A, M', C be the labels of r, a, M, c respectively. Consider $[R, A, M', C] \to [R+1, A+2, A+2, M'+1, C]$. Denote this five-element subarray by J''. If $C > M' + 1$, then the vertex in with label $M' + 1$ (represented by the fourth element in J'') must be operated on. (If not, then operations must have been conducted on the vertex with label $A+2$ in the subarray J'', or on c, both with labels greater than $l(M)+1$, a contradiction.) Thus $k_0 = 1$.

Let $q = J[m-1][(c+2) \mod |J_{m-1}|]$, and $Q = l(q)$.

Otherwise, $C = M' + 1$. Consider the next operation. It is either on a vertex with label $M' + 1$ or c. If it is on a vertex with label $M' + 1$, then $k_0 = 1$. Otherwise, consider the next operation $[R+1, A+2, A+2, M'+1, C, Q] \to [R+1, A+2, A+2, M'+2, C+2, C+2, Q+1]$. If the vertex with label $M'+2$ (the fourth element in the current subarray) is not operated on, then the vertices consecutive to it must have operations applied on it. This will be needed to cause this label $M' + 2$ to increase over the next polygon arrays, so that an operation can be applied on an M-originating vertex. However, both $A+2, C+2 > M'+2$, hence operations cannot be applied on them before an operation is applied on this vertex, contradiction. Thus $k_0 = 2$.

Hence $k_0 = 1$ or 2.

Definition 10.3. *A left or right corner operation on an element of a subarray is the operation applied on the vertices corresponding to the leftmost or rightmost element of the subarray respectively. For example, if the subarray is* $[3, 4, 1, 2]$, *operations on the corresponding vertices to* 3 *and* 2 *are left and right corner operations respectively, while those on the corresponding vertices to* $4, 1$ *in the tree are not.*

In the following sections, we will show that for all sequences of operations starting from a, there exists a more optimal sequence of operations if we start operating on M instead.

10.1 $k_0 = 1$

We start off with $[R, A, M', C]$.

10.1.1 No operations were applied on r-originating vertices before the M-originating vertex labelled $M' + 1$

Then $[R, A, M', C] \to [R+1, A+2, A+2, M'+1, C]$.

If any operation was applied on a vertex with label C, we have a contradiction. Thus among all the vertices corresponding 5 elements in the subarray in the tree, the first to be operated on is the vertex with label $M' + 1$.

From here, $O_i, i \in \mathbb{N}$ represent a sequence of operations on vertices not originating from r, a, M, c.

We thus deduce that the sequence of operations on vertices are as follows: $[R+1, A+2, A+2, M'+1, C] \to O_1 \to [R+1+x, A+2, A+2, M'+1, C+y] \to [R+1+x, A+2, A+3, M'+3, M'+3, C+1+y]$ where x, y are some non-negative integers. The area subtracted due to operations only on r, a, M, c-originating vertices is $\frac{1}{2^A} + \frac{1}{2^{M'+1}}$.

However, consider the following set of operations: $[R, A, M', C] \to [R, A+1, M'+2, M'+2, C+1] \to O_1 \to [R+x, A+1, M'+2, M'+2, C+1+y] \to [R+x+1, A+3, A+3, M'+3, M'+2, C+1+y]$. The area subtracted is $\frac{1}{2^{M'}} + \frac{1}{2^{A+1}}$ which is at least $\frac{1}{2^{M'+1}} - \frac{1}{2^{A+1}} \geq \frac{1}{2^{A+1}}$ more than the original set of operations. The only element in the final subarray that is more than the original is $A+3$, but we have proven that operations on vertices originating from a vertex with label $A+3$ subtract at most $\frac{1}{2^{A+2}}$ from the area by **Lemma 8.1**. This is still less than $\frac{1}{2^{A+1}}$. Also, all other non-r, a, M, c-originating vertices are not affected by this reordering of operations. This can be verified by checking that the number of left corner operations and right corner operations remain constant in both sets of operations after every operation on r, a, M, c-originating elements. Therefore, the second set of operations is more optimal, contradiction.

10.1.2 An operation was applied on r-originating vertices before the M-originating vertex labelled $M'+1$

Since the operation is applied on a vertex with label at most $R+1$, we have $R+1 \leq M'+1$. Thus $R = M'$. Considering only the operations on vertices originating from r, a, M, c, we have the sequence of operations

Array	Area removed	Cumulative area removed
$[R, A, M', C]$	$\frac{1}{2^A}$	$\frac{1}{2^A}$
$[R+1, A+2, A+2, M'+1, C]$	N/A	$\frac{1}{2^A}$
$[R+1, A+2, A+2, M'+1, C+x]$	$\frac{1}{2^{R+1}}$	$\frac{1}{2^A} + \frac{1}{2^{R+1}}$
$[R+3, R+3, A+3, A+2, M'+1, C+x]$	N/A	$\frac{1}{2^A} + \frac{1}{2^{R+1}}$
$[R+3+y, R+3, A+3, A+2, M'+1, C+x+z]$	$\frac{1}{2^{M'+1}}$	$\frac{1}{2^A} + \frac{1}{2^{R+1}} + \frac{1}{2^{M'+1}}$
$[R+3+y, R+3, A+3, A+3, M'+3, M'+3, C+x+z+1]$	N/A	$\frac{1}{2^A} + \frac{1}{2^{R+1}} + \frac{1}{2^{M'+1}}$

where x, y, z are non-negative integers. The area subtracted is $\frac{1}{2^A} + \frac{1}{2^{M'+1}} + \frac{1}{2^{M'+1}} = \frac{1}{2^{M'}} + \frac{1}{2^A}$.

On the other hand, suppose we apply the following set of operations

Array	Area removed	Cumulative area removed
$[R, A, M', C]$	$\frac{1}{2^{M'}}$	$\frac{1}{2^{M'}}$
$[R, A+1, M'+2, M'+2, C+1]$	N/A	$\frac{1}{2^{M'}}$
$[R, A+1, M'+2, M'+2, C+1+x]$	$\frac{1}{2^R}$	$\frac{1}{2^{M'}} + \frac{1}{2^R}$
$[R+2, R+2, A+2, M'+2, M'+2, C+1+x]$	N/A	$\frac{1}{2^{M'}} + \frac{1}{2^R}$
$[R+2+y, R+2, A+2, M'+2, M'+2, C+1+x+z]$	$\frac{1}{2^{A+2}}$	$\frac{1}{2^{M'}} + \frac{1}{2^R} + \frac{1}{2^{A+2}}$
$[R+2+y, R+3, A+4, A+4, M'+3, M'+2, C+x+z]$	N/A	$\frac{1}{2^{M'}} + \frac{1}{2^R} + \frac{1}{2^{A+2}}$

The area subtracted from these operations is $\frac{1}{2^{M'}} + \frac{1}{2^{M'}} + \frac{1}{2^{A+2}}$ which is at least $\frac{1}{2^{M'}} - \frac{3}{2^{A+2}}$ more than the previous scenario. However, the only two elements here that is more than the previous scenario have label $(A+4)$. The area subtracted due to operations on two vertices labelled $A+4$ is less than

$$\frac{2}{2^{A+3}} \leq \frac{1}{2^{A+2}} \leq \frac{1}{2^{M'}} - \frac{3}{2^{A+2}} \tag{24}$$

And as before, all non-r, a, M, c-originating vertices are not affected by this reordering. Applying all other operations as usual, we have a better operation sequence.

10.2 $k_0 = 2$

From here, $E_i, i \in \mathbb{N}$ represent a sequence of operations not on r, a, M, c, q-originating vertices.

After one operation, we have

$$[R, A, M', C, Q] \rightarrow [R+1, A+2, A+2, M'+1, C, Q] \tag{25}$$

Clearly no operation can be applied on any of the vertices labelled $A+2$ before those labelled $M'+2$ as $A+2 > M'+2$.

Thus, to increase $M'+1$ to $M'+2$ for an operation on a vertex with label $M'+2$ originating from M, we need to operate on the c-originating vertex with label C. Now, the label of c must be have the minimum label among those in the same layer as c. If $C > M'+1$, we would be forced to operate on the vertex with label $M'+1$. Thus $C = M'+1$. Also, if any operation was conducted on a q-originating vertex before a vertex with label $M'+2$, then C would be larger than $M'+1$. Again, this contradicts the inductive hypothesis. Therefore, when a c-originating element is operated on, it has label $C+1$.

Denote R' to be the vertex with label $R+1$ in the current subarray.

10.2.1 No operations were applied to R'-originating vertices before $M+1$

Array	Area removed	Cumulative area removed
$[R, A, M', M'+1, Q]$	$\frac{1}{2^A}$	$\frac{1}{2^A}$
$[R+1, A+2, A+2, M'+1, M'+1, Q]$	N/A	$\frac{1}{2^A}$
$[R+1+x, A+2, A+2, M'+1, M'+1, Q+y]$	$\frac{1}{2^{M'+1}}$	$\frac{1}{2^A}+\frac{1}{2^{M'+1}}$
$[R+1+x, A+2, A+2, M'+2, M'+3, M'+3, Q+1+y]$	N/A	$\frac{1}{2^A}+\frac{1}{2^{M'+1}}$
$[R+1+x+w, A+2, A+2, M'+2, M'+3, M'+3, Q+1+y+z]$	$\frac{1}{2^{M'+2}}$	$\frac{1}{2^A}+\frac{3}{2^{M'+2}}$
$[R+1+x+w, A+2, A+3, M'+4, M'+4, M'+4, M'+3, Q+1+y+z]$	N/A	$\frac{1}{2^A}+\frac{3}{2^{M'+2}}$

where w, x, y, z are non-negative integers. The area subtracted is $\frac{1}{2^A}+\frac{1}{2^{M'+1}}+\frac{1}{2^{M'+2}}$. The follow diagram demonstrates this, without E_1 and E_2.

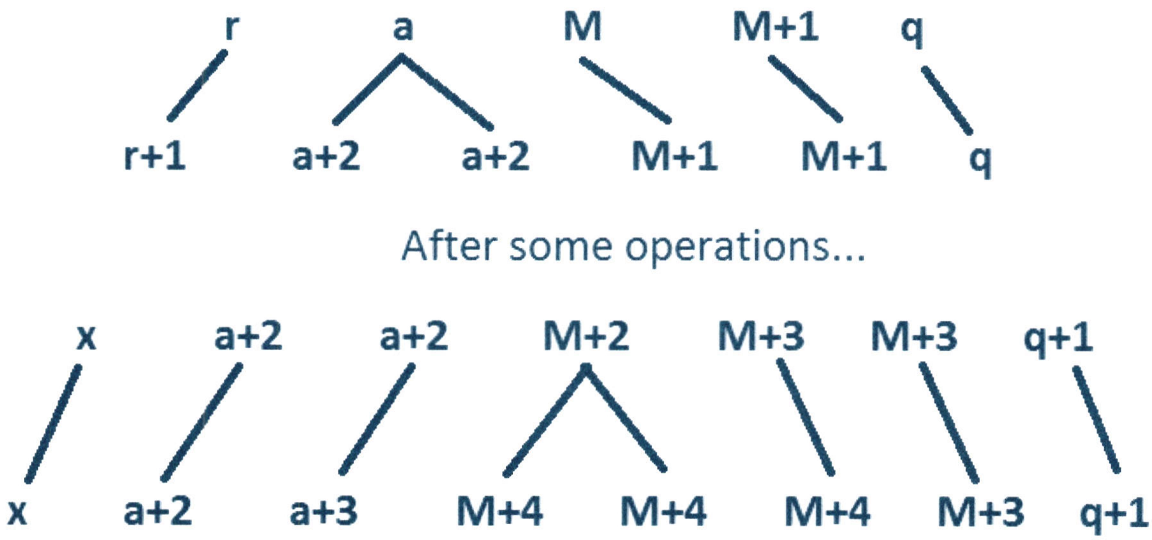


Figure 19: Operating on a

Consider a second scenario with a different set of operations:

Array	Area removed	Cumulative area removed
$[R, A, M', M'+1, Q]$	$\frac{1}{2^{M'}}$	$\frac{1}{2^{M'}}$
$[R, A+1, M'+2, M'+2, M'+2, Q]$	N/A	$\frac{1}{2^{M'}}$
E_1	N/A	$\frac{1}{2^{M'}}$
$[R+x, A+1, M'+2, M'+2, M'+2, Q+y]$	$\frac{1}{2^{M'+2}}$	$\frac{5}{2^{M'+2}}$
$[R+x, A+1, M'+3, M'+4, M'+4, M'+3, Q+y]$	N/A	$\frac{5}{2^{M'+2}}$
E_2	N/A	$\frac{5}{2^{M'+2}}$
$[R+x+w, A+1, M'+3, M'+4, M'+4, M'+3, Q+1+y+z]$	$\frac{1}{2^{A+1}}$	$\frac{5}{2^{M'+2}}+\frac{1}{2^{A+1}}$
$[R+x+w+1, A+3, A+3, M'+4, M'+4, M'+4, M'+3, Q+y+z]$	N/A	$\frac{5}{2^{M'+2}}+\frac{1}{2^{A+1}}$

The following diagram shows this without E_1 and E_2.

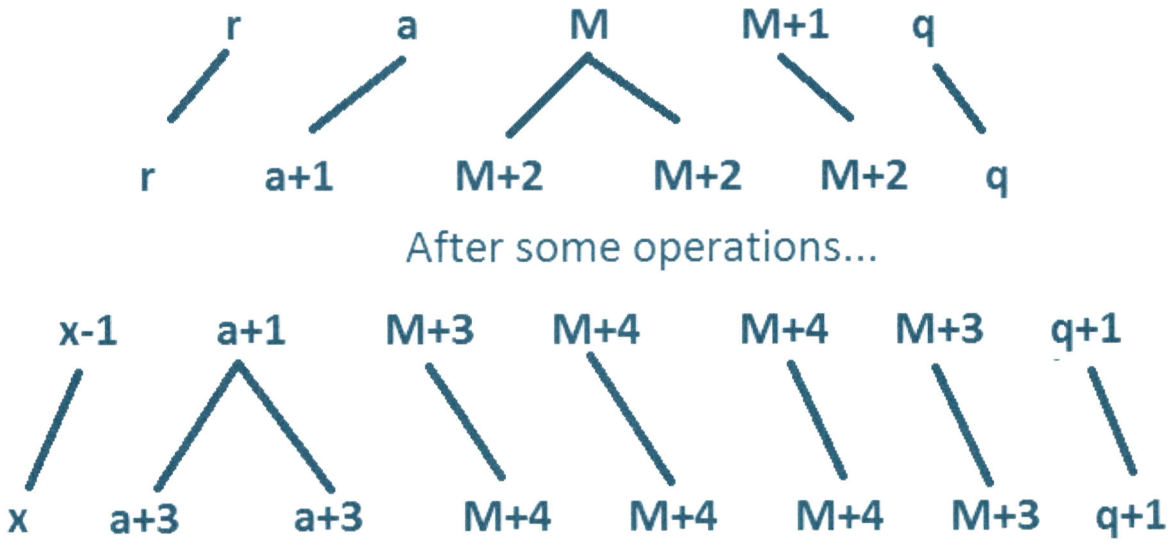


Figure 20: Operating on M

The area subtracted is $\frac{5}{2^{M'+2}} + \frac{1}{2^{A+1}}$. This is $\frac{1}{2^{M'+1}} - \frac{1}{2^{A+1}} \geq \frac{1}{2^{A+1}}$ more than in the first scenario.

The only element in the second subarray more than the corresponding element in the first subarray is $A+3$. However, operations on vertices originating from a vertex labelled $A+3$ subtracts at most $\frac{1}{2^{A+2}} < \frac{1}{2^{A+1}}$ from the area. Again, non-r, a, M, c, q-originating elements are not affected by this switch in operations. Applying all other operations as usual, we have a better operation sequence.

10.2.2 An operation was applied to an R'-originating vertex before a vertex with label $M'+1$

Similarly as with the case $k_0 = 1$, this implies $R = M'$. We deduce that the operation sequence is as follows:

$$[R, A, M', M' + 1, Q] \rightarrow [R+1, A+2, A+2, M'+1, M'+1, Q] \quad (26)$$

Then, we will operate on two of the vertices represented in the current subarray with label $M'+1 = R+1$: the second and sixth elements. (Note that we cannot operate on the fifth element as it is M-originating.)

Since they are not consecutive, the order does not matter. We apply both operations and the operation on the M-originating vertex with label $M+2$ to get

Array	Area removed	Cumulative area removed
$[R+1, A+2, A+2, M'+1, M'+1, Q]$	N/A	0
E_2	N/A	0
$[R+1, A+2, A+2, M'+1, M'+1, Q+y]$	$\frac{1}{2^{R+1}}$	$\frac{1}{2^{R+1}}$
$[R+3, R+3, A+3, A+2, M'+1, M'+1, Q+y]$	N/A	$\frac{1}{2^{R+1}}$
E_3	N/A	$\frac{1}{2^{R+1}}$
$[R+3+x, R+3, A+3, A+2, M'+1, M'+1, Q+y+w]$	$\frac{1}{2^{M'+1}}$	$\frac{1}{2^{R+1}} + \frac{1}{2^{M'+1}}$
$[R+3+x, R+3, A+3, A+2, M'+2, M'+3, M'+3, Q+1+y+w]$	N/A	$\frac{1}{2^{R+1}} + \frac{1}{2^{M'+1}}$
E_4	N/A	$\frac{1}{2^{R+1}} + \frac{1}{2^{M'+1}}$
$[R+3+m+x, R+3, A+3, A+2, M'+2, M'+3, M'+3, Q+1+y+w+n]$	$\frac{1}{2^{M'+2}}$	$\frac{1}{2^{R+1}} + \frac{3}{2^{M'+2}}$
$[R+3+m+x, R+3, A+3, A+3, M'+4, M'+4, M'+4, M'+3, Q+1+y+w+n]$	N/A	$\frac{1}{2^{R+1}} + \frac{3}{2^{M'+2}}$

where x, y, z, w are some non-negative integers. The area subtracted is $\frac{1}{2^A} + \frac{5}{2^{M'+2}}$.

On the other hand, consider

Array	Area removed	Cumulative area removed
$[R, A, M', M'+1, Q]$	$\frac{1}{2^{M'}}$	$\frac{1}{2^{M'}}$
$[R, A+1, M'+2, M'+2, M'+2, Q]$	N/A	$\frac{1}{2^{M'}}$
Apply E_2	N/A	$\frac{1}{2^{M'}}$
$[R, A+1, M'+2, M'+2, M'+2, Q+y]$	$\frac{1}{2^R}$	$\frac{1}{2^{M'}} + \frac{1}{2^R}$
$[R+2, R+2, A+2, M'+2, M'+2, M'+2, Q+y]$	N/A	$\frac{1}{2^{M'}} + \frac{1}{2^R}$
Apply E_3	N/A	$\frac{1}{2^{M'}} + \frac{1}{2^R}$
$[R+2+x, R+2, A+2, M'+2, M'+2, M'+2, Q+y+w]$	$\frac{1}{2^{M'+2}}$	$\frac{1}{2^R} + \frac{5}{2^{M'+2}}$
$[R+2+x, R+2, A+2, M'+2, M'+3, M'+4, M'+4, Q+y+1+w]$	N/A	$\frac{1}{2^R} + \frac{5}{2^{M'+2}}$
Apply E_4	N/A	$\frac{1}{2^R} + \frac{5}{2^{M'+2}}$
$[R+2+x+m, R+2, A+2, M'+2, M'+3, M'+4, M'+4, Q+y+1+w+n]$	$\frac{1}{2^{M'+2}}$	$\frac{1}{2^R} + \frac{3}{2^{M'+1}}$
$[R+2+x+m, R+2, A+3, M'+4, M'+4, M'+4, M'+4, M'+4, Q+y+1+w+n]$	N/A	$\frac{1}{2^R} + \frac{3}{2^{M'+1}}$

The area subtracted is $\frac{5}{2^{M'+1}}$ which is at least $\frac{5}{2^{M'+2}} - \frac{1}{2^A}$ more than the previous scenario. The only element in the final polygon array that is greater than in the previous scenario is the $(M'+4)$, and the maximum area subtracted by operations on vertices originating from a vertex with label $M'+4$ is less than $\frac{1}{2^{M'+4}} \leq \frac{5}{2^{M'+2}} - \frac{1}{2^A}$. Again, we can check left and right corner operations to verify that other non-r, a, M, c, q-originating elements are not affected. If we apply all other operations as usual, we thus get a better sequence.

The two cases $k_0 = 1$ and $k_0 = 2$ both contradict the optimality of the operation sequence. Thus, by mathematical induction, we should operate on M first.

Therefore, we can use Greedy algorithm to find the minimum area left. In other words, we can do the kth operation on the vertex with smallest label in the kth layer to give the maximum area subtracted for all $k \in \{1, 2, ..., F\}$.

Remark 10.4. *There might be multiple ways to obtain the largest area subtracted, but using Greedy is one of the ways. This still does not tell which we should pick. We will consider the possibilities.*

Remark 10.5. *This proof holds only when a finite number of operations is performed on the hexagon. It does not hold when an infinite number of operations are performed.*

11 Algebraic Bounds

11.1 A Bound of $\frac{3}{4}K_0$

Suppose we could achieve a remaining area of less than 18. Let the maximum possible area deducted be X after F operations. Then $X > 6$.

Consider the optimal operation sequence which uses Greedy Algorithm.

Suppose the first operation was $[0, 0, 0, 0, 0, 0] \to [1, 2, 2, 1, 0, 0, 0]$ without loss of generality.

The next operation must be done on a vertex labelled 0, hence there are three possible second operations:

(1) $[1, 2, 2, 1, 0, 0, 0] \to [1, 2, 2, 2, 2, 2, 1, 0]$

(2) $[1, 2, 2, 1, 0, 0, 0] \to [2, 2, 2, 1, 0, 1, 2, 2]$

(3) $[1, 2, 2, 1, 0, 0, 0] \to [1, 2, 2, 1, 1, 2, 2, 1]$

In either of these operations, we have only a total area of 2 subtracted from the area. Thus we have to perform more operations.

Lemma 11.1. *Consider a subarray $[a, a, a]$, the $c, c+1$ and $c+2$th elements in J_k with indices taken modulo $|J_k|$. The area subtracted from operations on vertices originating from $J[k][c], J[k][c+1], J[k][c+2]$ is at most $\frac{1}{2^{a-2}}$.*

Proof. We may ignore all other operations except for those on $J[k][c], J[k][c+1], J[k][c+2]$. Let the area subtracted be Y.

If the first operation was $[a, a, a] \to [a+1, a+2, a+2, a+1]$, by **Lemma 8.2**, the area subtracted is at most $\frac{1}{2^a} + 2\frac{1}{2^a}$ which is less than $\frac{1}{2^{a-2}}$.

Otherwise, without loss of generality, suppose it was $[a, a, a] \to [a+2, a+2, a+1, a]$. Consider the first operation, if exists, done on a vertex originating from the corresponding vertex to the rightmost a. Suppose it is on $a + k_1$.

By Lemmas 8.1, and 8.2, and considering the subarrays $[a+2]$, $[a+2, a+1]$ and $[a]$, we obtain

$$Y \leq \frac{1}{2^a} + \frac{1}{2^{a+1}} + \frac{1}{2^a} + \frac{1}{2^{a+k_1-1}} \leq \frac{1}{2^{a-2}} \tag{27}$$

unless $k_1 = 0$. If $k_1 > 0$, or doesn't exist, we are done with the proof.

If $k_1 = 0$, then $[a, a, a] \to [a+2, a+2, a+1, a] \to [a+2, a+2, a+2, a+2, a+2]$. The area subtracted from $(a+2)$-originating vertices in $[a+2, a+2, a+2]$ is at most $\frac{Y}{4}$, while those from $[a+2, a+2]$ is at most $2\frac{1}{2^{a+1}} = \frac{1}{2^a}$ by Lemma 8.1. We thus obtain

$$Y \leq \frac{1}{2^a} + \frac{1}{2^a} + \frac{Y}{4} + \frac{1}{2^a} \tag{28}$$

Solving, we obtain $Y \leq \frac{1}{2^{a-2}}$. □

For Cases (1) and (2), an operation must be done on a 0, resulting in $J_3 = [2, 2, 2, 2, 2, 2, 2, 2, 2]$. The area subtracted from T_0 is now 3. Dividing J_3 into 3 subarrays $[2, 2, 2]$, we realise that for more than an area of 6 to be subtracted in total, we have to subtract at least 1 from operations on vertices originating from vertices in one of these subarrays. This is a contradiction by **Lemma 9.1**.

For Case (3), the area subtracted from vertices that are $1, 2$-originating in the subarrays $[1, 2]$ or $[2, 1]$ is at most $\frac{1}{2^0} = 1$ from **Lemma 8.2**. Hence the total area subtracted X is at most $2+1+1+1+1 = 6$.

In either case, $X \leq 6$, a contradiction.

Thus at least $\frac{3}{4}$ of the area of T_0 remains after any finite number of operations.

12 Results

By establishing the *polygon array*, we have thus proven that starting from a regular hexagon, at least 0.75 of the original area remains after any finite number of operations. Using geometric and algebraic methods respectively, we obtain $\frac{49}{85}$ (0.576) and $\frac{3}{4}$ (0.75) the original area. (If the original area is K_0, then the corresponding results are 1.496 and 1.949 respectively.)

We have also run a C++ program to estimate the bound for n-gons, $n \geq 6$. Given the number of cuts as input, it can either

(i) randomly perform an operation at each step;

or (ii) randomly choose one of the operations that will result in the most area removed (Greedy Algorithm) and perform it at each step.

We have found that there exists a sequence of cuts that results in area less than 0.80216 of the original by running the program for 1000 cuts a few times. Hence, our bound of 0.75 is actually quite high.

The following shows the results for polygons with more sides, where the polygon is inscribed in a unit circle:

Number of Sides	Random	Greedy	Area of Initial Polygon	Minimum Proportion
7	2.34039	2.33738	2.73641	0.85417
8	2.50551	2.501	2.82847	0.88422
9	2.63194	2.62626	2.89254	0.90794
10	2.7192	2.72527	2.93892	0.92523

Figure 21: Table of results

The following shows the trend for even more sides:

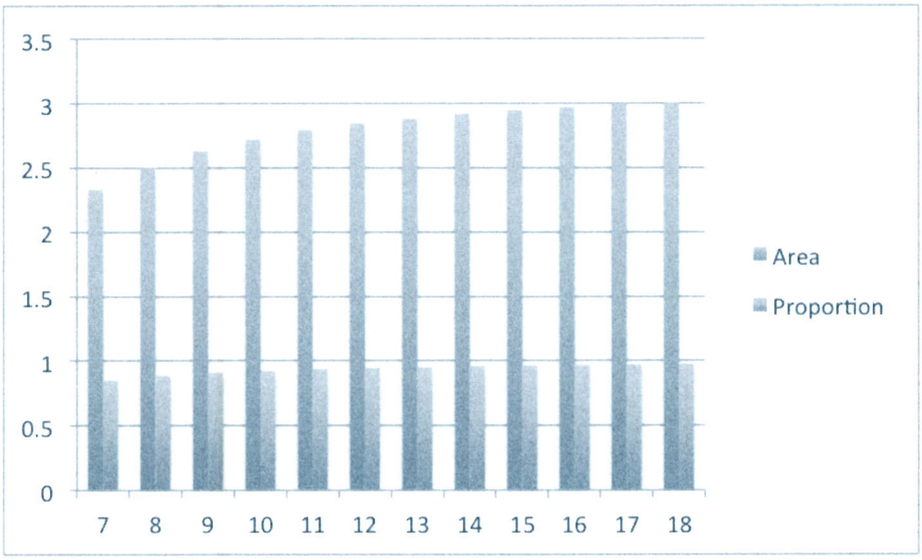
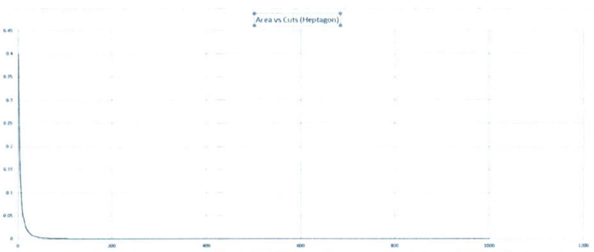
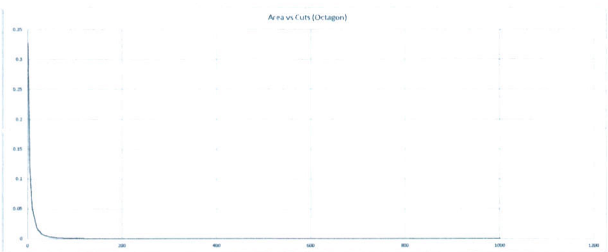


Figure 22: Graph

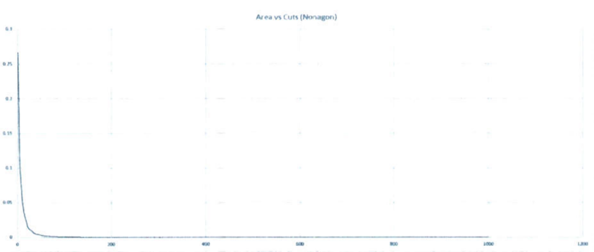
The following graphs show the decrease in area over cuts (Done using Excel):



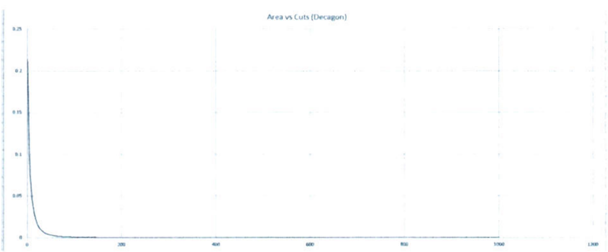
(a) Heptagon



(b) Octagon



(a) Nonagon



(b) Decagon

Figure 24: Various polygons

Note that for the decagon case, the area left after cutting randomly is less than that for Greedy. This is not a surprise, as the program for Greedy randomly chooses one of the operations that will remove the most area at each step, and does not consider all possible outcomes.

13 Further Research, Extensions and Generalisations

The bound of $\frac{3}{4}$ is still short of 0.8 by a fair amount, which is the approximate bound from the program. Further work must be done to tighten this bound further. Also, polygons with more than 6 sides were only touched on in this report. Research can also be done to extend this problem to n-gons with $n \geq 3$. Following that, the problem can be extended to regular 3-dimensional polyhedra, where a vertex is chosen and the cut is made along the plane passing through all the midpoints of the edges from the vertex.

14 Acknowledgments

We would like to express our gratitude to Mr Chia Vui Leong, Dr Lee Chan Lye and Mr Chai Ming Huang for guiding us in this project.

We would also like to acknowledge the Mathematical Association of America for the USAMO problem from which this project is based upon.

15 References

1. Art of Problem Solving. (n.d.). Retrieved September 10, 2015, from http://www.artofproblemsolving.com/wiki/index.php/1997˙USAMO˙Problems

2. Menelaus' Theorem – from Wolfram MathWorld. (n.d.). Retrieved from http://mathworld.wolfram.com/MenelausTheorem.html