MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **81** (2019) 505-522

ISSN 0340 - 6253

Extremal Polygonal Arrays for the Merrifield–Simmons Index

Guillermo De Ita Luna[a], **J. Raymundo Marcial-Romero**[b], **J. A. Hernández-Servín**[b]

[a]*Faculty of Computer Science, Autonomous University of Puebla, Puebla, México*
deita@cs.buap.mx

[b]*Faculty of Engineering, Autonomous University of Mexico State, Campus Toluca, Mexico*
jrmarcialr@uaemex.mx , xoseahernandez@uaemex.mx

(Received March 29, 2018)

Abstract

For any polygonal array, independently of the number of sides on each polygon, the zig-zag polygonal array has the extremal minimum value for the Merrifield–Simmons index. This result generalizes a well known fact obtained for hexagonal chains.

We analyze the product between two Fibonacci numbers with complementary indices. The results of the analysis will be used in our proposal. Our method does not require the explicit computation of the number of independent sets on the involved array graphs, instead it is based on the application of the edge division rule as a way to decompose polygonal array graphs.

1 Introduction

Polygonal array graphs have been widely investigated, and they represent a relevant area of interest in mathematical chemistry because they have been used to study intrinsic properties of molecular graphs [1]. Several works have been developed to analyze extremal values for the number of independent sets (known in mathematical chemistry area as the Merrifield–Simmons index) on polygonal arrays [2–7].

Merrifield and Simmons showed the correlation between the number of independent sets and the boiling points on polygonal chain graphs that represent chemical molecules. It is well known that the Merrifield–Simmons index is an important invariant of the structural chemistry [3, 5].

Gutman [8] analyzed extremal hexagonal chains according to three topological invariants: Hosoya index, largest eigenvalue and Merrified-Simmons index. Gutman showed that the extremal topology for the maximum Merrifield–Simmons index, in the particular case of hexagonal chains, is the linear hexagonal chain. He conjectured that the chain with the smallest Merrifield–Simmons index is unique and it corresponds to the zig-zag polyphenegraph. The researchers of Gutman [4, 8] greatly inspired the study of extremal polygonal chains.

Zhang et al. [6, 7] solved Gutman's conjecture. They showed that the minimum of the Merrifield–Simmons index on hexagonal chains is achieved by the zig-zag polyphenegraph. Later on, Cao et al. [2] showed extremal polygonal chains for k-matchings (Hosoya index), considering the topology of polygonal arrays that provide maximum as well as minimum values. His proofs are based on the use of the Z-polynomial (Z-counting polynomial). While Zhang et al. [9] determined extremal hexagonal chains concerning the total ϕ-electron energy, which are similar to the extremal chains in [7].

Several works deal with the characterization of the extremal graphs with respect to Hosoya and Merrifield–Simmons indices in several given graph classes, like: trees, unicyclic graphs, and certain structures involving pentagonal and hexagonal cycles [1–3, 5, 10–12]. For example, Ren et al. [13] determined the minimal Merrifield–Simmons index of double hexagonal chains. In [14], Li et al. characterized the tree with the maximal Merrifield–Simmons index among the trees with a given diameter. In [12], a survey about extremal graphs for Hosoya and Merrifield–Simmons indices involving different graph topologies is considered.

In this paper, we determine the extremal graph for the minimum Merrifield–Simmons index regarding any kind of polygonal arrays, generalizing a result previously obtained for hexagonal and pentagonal chains. Our proofs are based on properties that are derived from the product between two Fibonacci numbers that have complementary indices. Furthermore, our proofs do not require the explicit computation of the Merrifield–Simmons index on those polygonal arrays, instead the edge division rule is applied as a way to decompose polygonal arrays. We believe that our method can be adapted to compute other intrinsic properties on molecular graphs.

2 Array of polygons

Let $G = (V, E)$ be an undirected graph with vertices set V and set of edges E. It is assumed that G is a simple graph, it has not loops nor parallel edges. The *neighborhood* of $x \in V$ is the set $N(x) = \{y \in V : xy \in E\}$, and its *closed neighborhood* is $N(x) \cup \{x\}$ which is denoted by $N[x]$. The degree of a vertex x in the graph G, denoted by $\delta_G(x)$, is $|N(x)|$. The degree of the graph G is $\Delta(G) = \max\{\delta_G(x) : x \in V\}$.

A path between two vertices v and w, denoted as P_{vw}, or simply as P_n, is a sequence of edges: $v_0v_1, v_1v_2, \ldots, v_{n-1}v_n$ such that $v = v_0$, $v_n = w$, and $v_kv_{k+1} \in E$, for $0 \leq k < n$; the length of the path is n. A simple path is a path where $v_0, v_1, \ldots, v_{n-1}, v_n$ are all distinct. A cycle is a non-empty path such that the first and last vertices are identical, and a simple cycle is a cycle in which no vertex is repeated, with the exception that the first and last vertices are identical.

A subset $S \subseteq V$ is called independent if for every $u, v \in S$ implies that $uv \notin E$. The corresponding counting problem on independent sets, denoted by $i(G)$, consists of counting the number of independent sets of a graph G. Computing $i(G)$ is a $\sharp$P-complete problem for graphs G where $\Delta(G) \geq 3$. The computation of $i(G)$ remains $\sharp$P-complete even if it is restricted to 3-regular graphs.

Let $G = (V, E)$ be a molecular graph. Denote by $n(G, k)$ the number of ways in which k mutually independent vertices can be selected in G. By definition, $n(G, 0) = 1$ for all graphs, and $n(G, 1) = |V(G)|$. Furthermore, $i(G) = \sum_{k \geq 0} n(G, k)$ is the *Merrifield–Simmons index* of G, that is, exactly the number of independent sets of G.

A polygon (also called a polygonal graph) is a simple cycle graph. Therefore, a cycle graph C_n of length n represents a polygon of n sides, and it forms a n-gon. The way that two k-gons are joined, via a common vertex or via a common edge, defines different classes of polygonal chemical compounds. Two polygons that have an edge in common are called *adjacent*.

A polygonal chain is a 2-connected simple graph G obtained by identifying a finite number of congruent regular polygons (called basic polygons) one by one such that each vertex of G has degree 2 or 3 and each basic polygon, except the first one and the last one, is adjacent to exactly two basic polygons. A polygonal array is a graph $P_{k,t}$ obtained by identifying a finite number of t congruent polygons each of size k, also known as a chain of t k-gons.

A special class of polygonal arrays is the class of hexagonal chains, which are chains formed by n 6-gons. Hexagonal chains are the graph representations of an important subclass of benzenoid molecules, namely of the so called unbranched catacondensed benzenoids. The structure of these graphs is apparently the simplest among all hexagonal systems. Therefore, it is not surprising that a great deal of mathematical and mathematico–chemical results known in the theory of hexagonal systems apply, in fact, only to hexagonal chains [8].

Let $H_n = h_1 h_2 \cdots h_n$ be a polygonal array with n basic polygons, the polygons do not have the same number of sides necessarily, and where each h_i and h_{i+1} have exactly one common edge $e_i, i = 1, 2, \ldots, n-1$. A polygonal array with at least two polygons has two end-polygons, h_1 and h_n, while $h_2, \ldots, h_{n-1}$ are the internal polygons of the array. In a polygonal array, each vertex has degree either 2 or 3. The vertices of degree 3 are exactly the end points of the common edges between adjacent polygons. Let H be the subgraph from H_n induced by the vertices of degree 3. A polygonal array H_n is called a chain of type one if H is a path [2].

The distance $d_G(x, y)$ from a vertex x to another vertex y is the minimum number of edges in an $x - y$ path of G. The distance $d_G(x, S)$ from a vertex x to a set S is the $\min_{y \in S} d_G(x; y)$. Similarly, we define the distance between two edges e_1, e_2 on the graph G: $d_G(e_1, e_2)$, as the minimum number of edges in an $e_1 - e_2$ path of G, without consider the same edges e_1 and e_2.

Let H_n be a hexagonal chain with n basic 6-gons joined by one common edge between two adjacent hexagons. If for each pair of consecutive sharing edges e_i and e_{i+1} of H_n, it holds that $d_{H_n}(e_i, e_{i+1}) = 2$ according to the clockwise direction, then H_n is known as the linear hexagonal chain (L_n), and if $d_{H_n}(e_i, e_{i+1}) = 1$ according to the clockwise direction on each pair of consecutive common edges, then H_n is known as the zig-zag hexagonal chain (Z_n). Notice that in a zig-zag polygonal array the induced subgraph H (formed by the vertices of degree 3) is a path.

In some articles, the number of independent sets ($i(G)$) is also called the Fibonacci number of the graph G. For example, in [14], Li et al., characterized the tree with the maximum Fibonacci number among the trees with a given diameter. In [16], Zhao and Li investigated the orderings of two classes of trees by their Fibonacci numbers. In [1], Pedersen and Vestergaard studied the Fibonacci number for the unicyclic graphs. In [19], Yu and Tian studied the Fibonacci numbers of the graphs with given edge-independence number and cyclomatic number. Yu and Lv [17, 20] studied the Fibonacci numbers of trees with maximum degree and given pendent vertices, respectively. Ye et al., ordered the unicyclic graphs with given girth according to the Fibonacci numbers in [18]. More related to our work, in [15] is analyzed the array of polyphenylene compounds represented by graphs obtained from a hexagonal cactus by expanding each of its cut-vertices to an edge.

In recent years, several works have been done for determining topology graphs corresponding to extremal Hosoya and the Merrifield–Simmons indices [2, 5, 10]. For many graph classes that have been studied so far, graphs that minimizes the Hosoya index coincide with those that maximizes the Merrifield–Simmons index, and vice versa, although its relation is still not totally understood. For example, Deng [3] showed that graphs with n vertices and $n+1$ edges, denoted as $(n, n+1)$-graphs, its smallest Merrifield–Simmons index does not coincide with the maximum for the Hosoya index.

Here, we determine the extremal topology for the minimum Merrifield–Simmons index for any polygonal array, not only for hexagonal chains. In fact, it does not matter the number of sides of each polygon in the array. For this, we develop in the following chapter results about the product between Fibonacci numbers.

3 Product between two Fibonacci numbers

It is well known that for any simple path P_n of length $n-1$, that is P_n has n vertices and $n-1$ edges, P_n fulfills $i(P_n) = F_{n+2}$, where F_n is the nth-Fibonacci number with initial values $F_0 = 0, F_1 = 1$ and $F_n = F_{n-1} + F_{n-2}$. Let us consider an isolated vertex as a linear path of length zero, therefore, $i(P_1) = F_3 = 2$.

Let $k > 0$ be a constant integer and let P_i and P_j be two disjointed simple paths, such that $i + j = k$. It is known that $i(P_i \oplus P_j) = i(P_i) \cdot i(P_j) = F_{i+2} \cdot F_{j+2}$. Let the sequence $\beta_{k,s} = F_s \cdot F_{k-s}$, defined for all $k = 2, 3, ...$ and $1 \leq s < k$. Given a constant $k > 0$, we want to determine for which pair: (i, j), $i, j \geq 1$, and $j + i = k$, the product $i(P_i)i(P_j)$ has extremal values, that is, when it achieves the maximum and the minimum values. We establish which are the extremal values for the sequence $\beta_{k,s} = F_s \cdot F_{k-s}$ in the following proposition.

Proposition 1 *For any integers s with $1 \le s < k$,*

1. *if $k \ge 3$ then $\min_s \{F_s F_{k-s}\} = F_2 F_{k-2} = F_{k-2}$,*
2. *and if $k \ge 2$ then $\max_s \{F_s F_{k-s}\} = F_1 F_{k-1} = F_{k-1}$*

Proof 1 *Let $\beta_{k,s} = F_s F_{k-s}$ be the sequence of the Fibonacci product with $2 \le k$ and $1 \le s < k$. The number of terms in the sequence $\{\beta_{k,s}\}_s$ is $k-1$. The proof is done, first, by assuming that k is an even number and then, assuming that k is odd. Let us recall that Fibonacci numbers can be given using Binet's formula, that is for any k, $F_k = (a^k - b^k)/(a-b)$ where $a = (1+\sqrt{5})/2$ and $b = (1-\sqrt{5})/2$ are the roots of the polynomial $P(r) = r^2 - r - 1$. This also means that a, b satisfies the two equations: $a + b = 1$ and $ab = -1$.*

Let us assume that k is even, in other words $k = 2r$ for some $r \ge 1$. Thus, $\beta_{k,\lfloor\frac{k}{2}\rfloor - s} = F_{r-s}F_{r+s}$. However, $\beta_{k,\lfloor\frac{k}{2}\rfloor+s} = \beta_{2r,r+s} = F_{r+s}F_{r-s}$, therefore, $\beta_{k,\lfloor\frac{k}{2}\rfloor-s} = \beta_{k,\lfloor\frac{k}{2}\rfloor+s}$. On the other hand, if $k = 2r+1$ for some r, it is obvious that $\lfloor\frac{2r+1}{2}\rfloor = r$, then $\beta_{2r+1,r+1+j} = F_{r+1+j}F_{r-j} = \beta_{2r+1,r-j}$. This means that the sequence is symmetric, thus, it can be assumed without losing generality that $1 \le s \le \lfloor\frac{k}{2}\rfloor$. If we now choose the subsequence $\beta_{k,2}$, $\beta_{k,4}$,..., that is the sequence where s takes only even values, then by using Binet's formula and some algebraic manipulations, it could be checked that

$$\beta_{k,2(p+1)} = \beta_{k,2p} + F_{k-2(2p+1)} \tag{1}$$

where the facts $ab = -1$ and $a + b = 1$ are used in the calculations. Equation 1 only makes sense for $k \ge 6$ or $r \ge 3$, since p runs from one onwards. Furthermore, in the case where $p = 1$, we have that $\beta_{6,4} = F_4F_2 = F_2F_4 = \beta_{6,2}$, since $F_0 = 0$ as expected of the symmetry of the sequence. The other cases, when $k \le 5$, are easily handled since when $k = 2$ the sequence $\beta_{2,s}$ has only one term and, therefore, there is no minimum value. As for $k = 4$, there are only three terms in the sequence $\beta_{k,s}$, that is $\{\beta_{k,s}\}_s = \{2, 1, 2\}$ thus $\min_s\{\beta_{k,s}\} = 1 = F_{k-2}$ and $\max_s\{\beta_{k,s}\} = 2$. Therefore, for $k \ge 2$ the sequence satisfies $\beta_{k,2} \le \beta_{k,3} \le \cdots \le \beta_{k,\lfloor\frac{k}{2}\rfloor}$. From the above discussion, and from the fact that $F_{k-2(2p+1)} > 0$ in Equation 1 for $k > 6$, it makes the inequalities strict.

n	F_n	$\beta_{1,k}$ Max	$\beta_{2,k}$ Min	$\beta_{3,k}$	$\beta_{4,k}$	$\beta_{5,k}$	$\beta_{6,k}$	$\beta_{7,k}$	$\beta_{8,k}$	$\beta_{9,k}$	$\beta_{10,k}$	$\beta_{11,k}$	$\beta_{12,k}$	$\beta_{13,k}$
1	1	0												
2	1	1	0											
3	2	1	1	0										
4	3	2	1	2	0									
5	5	3	2	2	3	0								
6	8	5	3	4	3	5	0							
7	13	8	5	6	6	5	8	0						
8	21	13	8	10	9	10	8	13	0					
9	34	21	13	16	15	15	16	13	21	0				
10	55	34	21	26	24	25	24	26	21	34	0			
11	89	55	34	42	39	40	40	39	42	34	55	0		
12	144	89	55	68	63	65	64	65	63	68	55	89	0	
13	233	144	89	110	102	105	104	104	105	102	110	89	144	0
14	377	233	144	178	165	170	168	169	168	170	165	178	144	233

Table 1. The product of two Fibonacci numbers with complementary indices.

Analogously, taken the subsequence $\beta_{k,2p+1}$, that is when s takes odd values, then

$$\begin{aligned}(a-b)^2\beta_{k,2p+3} &= (a^{2p+3}-b^{2p+3})(a^{k-2p-3}-b^{k-2p-3}) \\ &= a^k - a^{2p+3}b^{k-2p-3} - a^{k-2p-3}b^{2p+3} + b^k \\ &= (a-b)^2\beta_{k,2p+1} + a^{2p+3}b^{2p+3}\frac{a^2-b^2}{a^2b^2}(a^{k-4(p+1)}-b^{k-4(p+1)}). \qquad (2)\end{aligned}$$

By using $ab=-1$, $a+b=1$ and Binet's formula again, Equation 2 is equivalent to $\beta_{k,2p+3}=\beta_{k,2p+1}-F_{k-4(p+1)}$. Once again, $F_{k-4(p+1)}$ has only meaning for $k\geq 8$, but the other cases were discussed above. Thus, the sequence satisfies $\beta_{k,1}\geq\beta_{k,3}\geq\cdots\geq\beta_{k,\lfloor\frac{k}{2}\rfloor}$. Therefore, $\min_s\{F_sF_{k-s}\}=\beta_{k,2}=F_{k-2}$ and $\max_s\{F_sF_{k-s}\}=\beta_{k,1}=F_{k-1}$ and the proposition follows. ■

In Table 1, we present some of the values of the sequence $\beta_{k,s}=F_s\cdot F_{k-s}$. Notice that different relations can be inferred when we consider the values of the table arranged like the Pascal's triangle.

Notice that the maximum $F_1\cdot F_{k-1}=F_{k-1}$ for the row (k) of the table results to be the minimum $F_2\cdot F_{k-1}=F_{k-1}$ for the row $(k+1)$. Also, the difference between the maximum and minimum in the row k is $F_{k-1}-F_{k-2}=F_{k-3}$. The fact that the extremal values of $\beta_{k,s}$ are in the first two consecutive columns of the Table 1 will have logical consequences on the topologies that represent the extremal values for the Merrifield–Simmons index on polygonal arrays.

Observe the symmetrical behavior of the sequence $\beta_{k,s}$ at the position $s>\lfloor\frac{k}{2}\rfloor$. In fact, $\beta_{k,\lfloor\frac{k}{2}\rfloor-j}=\beta_{k,\lfloor\frac{k}{2}\rfloor+j}$ if k is even, and $\beta_{k,\lfloor\frac{k}{2}\rfloor-j}=\beta_{k,\lfloor\frac{k}{2}\rfloor+j+1}$ if k is odd, and for all j

such that $1 \leq j \leq \left\lfloor \frac{k}{2} \right\rfloor - 2$.

Also, the sequence $\beta_{k,s}$ is increasing on the even indices of s, and it has a decreasing behavior on the odd indices of s. For example, $\beta_{k,2p} < \beta_{k,2(p+1)}$ for every $p \in \left\{1, 2, ..., \left\lfloor \frac{k}{4} \right\rfloor\right\}$, and all k. While, $\beta_{k,2p+1} > \beta_{t,2p+3}$ for every $p \in \left\{0, 1, ..., \left\lfloor \frac{k}{4} \right\rfloor - 1\right\}$ and all k.

We show in the following chapters, how the properties about the sequence $\beta_{k,s}$, and how the application of the edge division rule, are useful for the computation of the extremal values of the Merrifield–Simmons index on polygonal arrays.

4 Extremal topologies for a polygon joined to two paths

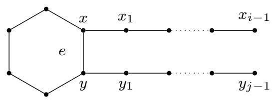
Some reductions rules have been useful to count combinatorial objects on graphs, particularly, the following rules are commonly used to count independent sets on a graph G:

1. Vertex reduction rule: Let $v \in V(G)$,
$$i(G) = i(G - v) + i(G - (N[v]))$$
2. Edge division rule : let $e = \{x, y\} \in E(G)$,
$$i(G) = i(G - e) - i(G - (N[x] \cup N[y]))$$

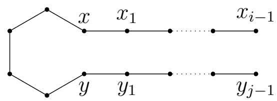
Let h_r be a polygon of r sides. Let P_i and P_j be two different simple paths of lengths $(i-1)$ and $(j-1)$, respectively, and such that $i+j = k$ becomes a constant. $P_i \cup_e h_r \cup_e P_j$ denotes the graph formed by joining P_i and P_j to the end-vertices of an edge $e \in E(h_r)$, as it is illustrated in Figure 1a. Notice that e can be any edge of the polygon since in fact, the initial polygon is a cycle and all of its edges are indistinguishable.

We show that $i(G)$ is maximum under the restriction $|P_i| + |P_j| = k$ when $i = 2$ (P_i has exactly two vertices and only one edge) and $j = k - 2$ (a path of $k - 2$ vertices and $k - 3$ edges).

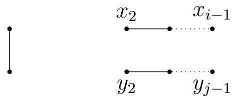
Let $e = \{x, y\} \in E(h_r)$. Let $P_i = \{x, x_1, \ldots, x_{i-1}\}$ and $P_j = \{y, y_1, \ldots, y_{j-1}\}$ be two disjointed paths, where $i + j = k$. $P_i \cup_x h_r \cup_y P_j$ denotes the resulting graph of joining P_i with h_r in the vertex x, and P_j with h_r in the vertex y, this means that $V(P_i) \cap V(h_r) = \{x\}$, $V(P_j) \cap V(h_r) = \{y\}$, and $V(P_j) \cap V(P_i) = \emptyset$.



(a) The graph G with the edge $e = xy$ to be removed



(b) The graph G after the operation $(G - e)$



(c) The graph G where all neighbours of x and y are removed: $G - (N[x] \cup N[y])$

Figure 1. A base graph G

Lemma 1 *If $G = P_i \cup_x h_r \cup_y P_j$ then the maximum for $i(G)$ is achieved when the path P_i has length 1 (two vertices $\{x, x_1\}$) and the path P_j has length $k - 3$ ($j = k - 2$ vertices).*

Proof 2 *Applying the edge division rule on $e = \{x, y\} \in E(h_r)$, it results in $i(G) = i(G - e) - i(G - (N[x] \cup N[y]))$. Notice that $(G - e)$ is a simple path of length $r + (i - 1) + (j - 1) = r + k - 2$; therefore, $i(G - e) = F_{r+k}$. Furthermore, $i(G - e)$ is invariant with respect to the selected position of the edge $e \in E(h_r)$.*

On the other hand, $(G - (N[x] \cup N[y]))$ is formed by three disjointed paths: P_{i-2}, P_{j-2} and the path that results from eliminating e and its two adjacent edges from h_r, e.g. P_{r-4}. Then, $i(G - (N[x] \cup N[y]))) = F_{r-2} \cdot F_i \cdot F_j$. In fact, the result of this product does not depend on the initial position of e in h_r, because the three resulting paths will have same lengths independently of the position of e in h_r.

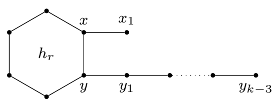
Then, the maximization of $i(G)$ is equivalent to the minimization of $F_i \cdot F_j$ because that term appears as minus in the equation to compute $i(G)$, and they are the unique parameters that can vary under the restriction $i + j = k$. According to part 1 of Proposition 1, $F_i \cdot F_j$ has a minimum value when $F_i = F_2$ and $F_j = F_{k-2}$ which means that the resulting path P_{i-2} after removing $N[x]$ from G must be empty and the resulting path P_{j-2} after removing $N[y]$ from G should have $k - 4$ vertices. This also means that the initial path P_i has two vertices $\{x, x_1\}$ and the original path P_j has $k - 2$ vertices, (see Figure 2). ■

Lemma 2 *If $G = P_i \cup_x h_r \cup_y P_j$ then the minimum for $i(G)$ is achieved when the path P_i has length zero ($i = 1$ vertices) and the path P_j has length $k - 2$ ($j = k - 1$ vertices).*

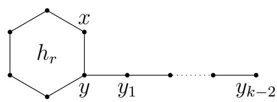
Proof 3 *Similar to Lemma 1, using case 2 from Proposition 1.* ∎

Furthermore, the maximum and minimum values for $i(G)$ are achieved independently of the number of edges in the polygon h_r. Consequently, our results are fulfilled for any polygon joined with two disjointed paths in the end-points of any of its edges. We show in Figure 2, the extremal topologies for $i(G)$ for the class of graphs: $G = P_i \cup_x h_r \cup_y P_j$.

From the previous two proofs, it is inferred that extremal value for $G = P_i \cup_x h_r \cup_y P_j$ is not unique, since due to the commutativity of $F_i \cdot F_j$ and the symmetry of $G' = P_j \cup_x h_r \cup_y P_i$ with G, we obtain that $i(G) = i(G')$, and both subgraphs get the same value for the Merrifield–Simmons index, and in its extremal value.



(a) $\max i(G)$



(b) $\min i(G)$

Figure 2. Extremal arrays for $i(G)$, for $G = P_i \cup_x h_r \cup_y P_y$

The following Lemmas and the Corollary will be useful for our analysis. They show that given an initial graph $G = (V, E)$, if new edges are added to $E(G)$ then $i(G)$ is decreasing, while if new vertices are added to $V(G)$ then $i(G)$ is increasing, even if the new vertices are connected to all original $v \in V(G)$.

Lemma 3 *Let $G = (V, E)$ be an undirected graph, let $x, y \in V(G)$, and $e = \{x, y\} \notin E(G)$, then $i(G) > i(G \cup e)$.*

Proof 4 *Let $S_e = \{S \in i(G) : x, y \in S\}$ be the independent sets in G containing the two vertices $x, y \in V$. $|S_e| > 0$ since at least the set $\{x, y\} \in S_e$ because $e \notin E(G)$. As, $i(G \cup e) = i(G) - |S_e|$ then $i(G) > i(G \cup e)$.*

Lemma 4 *Let $G = (V, E)$ be an undirected graph, and let $x \notin V$. Let $G_1 = G \cup \{\{x, v\} : \forall v \in V\}$, then $i(G_1) = i(G) + 1$.*

Proof 5 *$i(G_1) = i(G) \cup |\{\{x\}\}|$, since there are no more independent set including x and any other vertex from V. Then, $i(G_1) = i(G) + 1$.*

Corollary 1 *Let $G = (V, E)$ be an undirected graph, and let x, v be two vertices such that $x \notin V$, $v \in V$. Let $G_1 = G \cup \{x, v\}$, then $i(G_1) > i(G)$.*

Proof 6 *According to previous lemma,* $i(G_1) = i(G)+1$ *if there are no more edges between* x *and any other vertex* $v \in V$*. If any edge* $\{v, x\}$ *is omitted in* $E(G_1)$ *then* G_1 *is even greater than* $i(G) + 1$*. In whatever case,* $i(G_1) > i(G)$*.*

5 Extremal topologies on polygonal arrays

Let $H_{r_1,r_2...,r_p}$ be a linear array of p polygons, each one of r_i sides $1 \leq i \leq p$, e.g, $H_{r_1,r_2...,r_p} = h_{r_1}h_{r_2} \cdots h_{r_p}$. We denote e_i as the common edge between the polygon h_{r_i} and $h_{r_{i+1}}$. Let us write $H^{l_1,l_2...,l_{p-2}}_{r_1,r_2...,r_p}$ to represent $H_{r_1,r_2...,r_p}$ where l_i denotes the distance between the common edges e_i and e_{i+1} of three consecutive polygons.

Lemma 5 *Let* l *be the distance between the common edges* e_1 *and* e_2 *in* $H^l_{r_1,r_2,r_3}$ *such that* $1 \leq l < \lfloor \frac{r-1}{2} \rfloor$

$$\max\{i(H^l_{r_1,r_2,r_3})\} = i(H^2_{r_1,r_2,r_3})$$

Proof 7 *By applying the edge division rule on the edge* $e_2 = \{x, y\} \in E(h_{r_2})$*, we obtain:*

$$i(H^l_{r_1,r_2,r_3}) = i(H^l_{r_1,r_2,r_3} - e_2) - i(H^l_{r_1,r_2,r_3} - (N[x] \cup N[y])) \tag{3}$$

Notice that $(H^l_{r_1,r_2,r_3} - e_2)$ *is an array of two polygons,* $H : h_{r_1}h_{r_2+r_3-2}$*, with the common edge:* e_1*. Furthermore,* $i(H^l_{r_1,r_2,r_3} - e_2)$ *is invariant with respect to* l *since without regarding the value of* l*, it consists of a linear array of two polygons that are invariants in its lengths independently of the position of* e_2*.*

$(H^l_{r_1,r_2,r_3} - (N[x] \cup N[y]))$ *is the graph formed by two connected components. The first component is a simple path* P_{r_3-4} *with* $r_3 - 4$ *vertices and therefore* $i(P_{r_3-4}) = F_{r_3-2}$*. The second component depends on the value of* l*. When* $l = 1$*, this second component, denoted by* G_1*, is a path* $P_{r_1+r_2-6}$ *with* $r_1 + r_2 - 6$ *vertices and* $r_1 + r_2 - 7$ *edges. But in the case* $l > 1$*, the second connected component is* $G_2 = P_{l-1} \cup_x h_{r_1} \cup_y P_{r_2-l-3}$*, that is a subgraph with* $r_1 + r_2 - 6$ *vertices and* $r_1 + r_2 - 6$ *edges.*

Both connected subgraphs, G_1 *and* G_2 *come from the same original array:* $h_{r_1}h_{r_2}$*. They have the same number of vertices, but* G_2 *has one edge more than* G_1 *because the neighbor of* x *in* h_{r_1} *has a degree bigger (one more) when* $l = 1$ *than when* $l > 1$*. And according to Lemma 3 and corollary 1,* $i(G_2) < i(G_1)$*.*

Hence, in order to maximize $i(H^l_{r_1,r_2,r_3})$*, we must select the minimum between* $i(G_2)$ *and* $i(G_1)$*. As* $i(G_2) < i(G_1)$ *then the minimum is achieved for* $l > 1$*. However,* $l > 1$ *has* *several possibilities for the size of the path* P_{l-1} *for* G_2*. By Lemma 2, the minimum for* $i(G_2)$ *is achieved when the path* P_{l-1} *has just one vertex* ($P_{l-1} = \{x\}$) *and then* P_{r_2-l-3} *has* $r_2 - 4$ *vertices. This topology is only achieved if the original distance between* e_1 *and* e_2 *is two, and therefore,* $l = 2$. ■

Notice that distance $l \geq 2$ between common edges in a polygonal array can only be obtained for polygons with size greater than 5. In fact, distance two between adjacent hexagons gives us an unique topology. This is because the common edge is positioned in the same edge of the last hexagon independently of the direction of how it was counted, clockwise or counterclockwise direction.

Lemma 6 *Let* l *be the distance between the common edges* e_1 *and* e_2 *in* $H^l_{r_1,r_2,r_3}$ *such that* $1 \leq l < \left\lfloor \frac{r-1}{2} \right\rfloor$.

$$\min\{i(H^l_{r_1,r_2,r_3})\} = i(H^1_{r_1,r_2,r_3})$$

Proof 8 *By applying the edge division rule on the edge* $e_2 = \{x, y\} \in E(h_{r_2})$*, we obtain that*

$$i(H^l_{r_1,r_2,r_3}) = i(H^l_{r_1,r_2,r_3} - e_2) - i(H^l_{r_1,r_2,r_3} - (N[x] \cup N[y])) \tag{4}$$

Notice that $(H^l_{r_1,r_2,r_3} - e_2)$ *is an array of two polygons:* $h_{r_1}h_{r_2+r_3-2}$*, with a common edge:* e_1*. Furthermore,* $i(H^l_{r_1,r_2,r_3} - e_2)$ *is invariant with respect to* l *because* $H^l_{r_1,r_2,r_3} - e_2$ *consists of an array of two polygons that maintain the same number of sides independently of the value* l*.*

$(H^l_{r_1,r_2,r_3} - (N[x] \cup N[y]))$ *is a graph formed by two connected components. The first component is a simple path* P_{r_3-4} *with* $r_3 - 4$ *vertices and therefore* $i(P_{r_3-4}) = F_{r_3-2}$*. The second component depends on the value of* l*. Let* G_1 *be such component when* $l = 1$*. In this case,* G_1 *is a path* $P_{r_1+r_2-6}$ *with* $r_1 + r_2 - 6$ *vertices and* $r_1 + r_2 - 7$ *edges. But in the case* $l > 1$*, the second connected component is* $G_2 = P_{l-1} \cup_x h_{r_1} \cup_y P_{r_2-l-3}$*, that is a subgraph with* $r_1 + r_2 - 6$ *vertices and* $r_1 + r_2 - 6$ *edges.*

Both connected subgraphs, G_1 *and* G_2 *come from the same original array:* $h_{r_1}h_{r_2}$*. They have same number of vertices, but* G_2 *has one edge more than* G_1 *because the neighbor of* x *in* h_{r_1} *has a degree bigger (one more) when* $l = 1$ *than when* $l > 1$*. And according to Lemma 3 and corollary 1,* $i(G_2) < i(G_1)$*.*

Hence, in order to minimize $i(H^l_{r_1,r_2,r_3})$ *we need to select the maximum between* $i(G_1)$ *and* $i(G_2)$*. As* $i(G_1)$ *is greater than* $i(G_2)$*, then the case* $l = 1$ *achieves the maximum. Thus, the distance 1 between* e_1 *and* e_2 *minimize* $i(H^l_{r_1,r_2,r_3})$*.* ■

Similar to lemma 5, distance $l \geq 1$ between common edges in a polygonal array can only be obtained for polygons with size greater than 3. In fact, distance one between adjacent quadrangles defines an unique topology. This is because the common edge is positioned in the same edge of the last quadrangle independently of the direction of how it was counted, clockwise or counterclockwise direction.

Let $H_n : h_1 h_2 \cdots h_n$ be an array of $n \geq 2$ polygons and let h_{n+1} be a new polygon of k sides. We denote by $e_i = \{x_i, y_i\}, i = 1, \ldots, n-1$ the common edge between polygons h_i and h_{i+1}. Let us enumerate the edges of h_n as $b_0, b_1, \ldots, b_{k-1}$, where $b_0 = e_{n-1}$ is the common edge between h_n and h_{n-1} and the sequence follows the clockwise direction. We must select $e \in E(h_n)$ such that $i(H_n \cup_e h_{n+1})$ is maximum into the set of possible selections of edges in h_n. For example, e can not be any of the edges b_0, b_1, b_{k-1} because if we join h_{n+1} to H_n in any of those edges then $H_n \cup h_{n+1}$ looses the topology to be a polygonal array.

(a) Distance 2 **(b)** Distance 1

Figure 3. An octagonal array with distance 2 and 1, respectively.

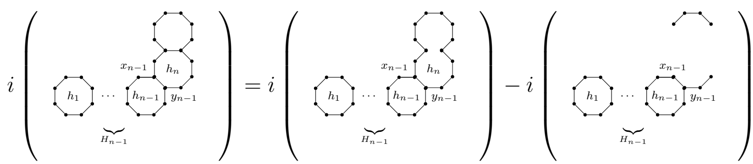


Figure 4. Application of the edge division rule on a $(n+1)$-octagonal array, with distance $l = 2$ between common edges

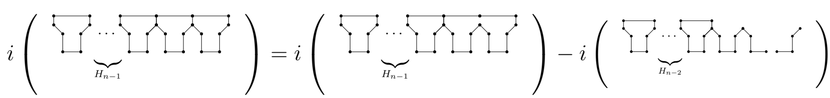


Figure 5. Application of the edge division rule on a $(n+1)$-octagonal array, with distance $l = 1$ between common edges

The extremal minimum topology for the Merrifield–Simmons index for hexagonal arrays has been identified as the zig-zag array (Z_n) - array with distance 1 between adjacent hexagons [6]. We present a generalized result taking into consideration all kind of polygonal arrays. The following theorem shows such result.

Theorem 1 *The minimum for the Merrifield–Simmons index for any polygonal array H_n is achieved when the distance between common edges for any pair of adjacent polygons is one.*

Proof 9 *By induction on n - the number of polygons in the array.*

1. *The base case for H_3 was shown previously (Lemma 6).*
2. *Suppose that the hypothesis holds for any polygonal array H_i with $i \leq n$.*
3. *Let H_n be a polygonal array with n polygons, and let h_{n+1} be an additional polygon of $k \geq 4$ sides, joined to H_n in the position $e = \{x, y\} \in E(h_n)$. By applying the edge division rule on e, we obtain*

$$i(H_{n+1}) = i(H_{n+1} - e) - i(H_{n+1} - (N[x] \cup N[y])) \tag{5}$$

$(H_{n+1} - e)$ is an array of n polygons where the inductive hypothesis is held. Then, $i(H_{n+1} - e)$ is minimum when the list of common edges $e_1, \ldots, e_{n-1}$ is arranged at distance 1 between any pair of consecutive edges in the list. Add more, the term $i(H_{n+1} - e)$ is invariant from e because its value does not depend on the selected position of e to join h_{n+1} to H_n.

On the other hand, $H_{n+1} - (N[x] \cup N[y])$ is a graph formed by two connected components. The first component is a simple path P_{k-4} with $k-5$ edges and $k-4$ vertices, where $i(P_{k-4}) = F_{k-2}$ is an invariant value independent to the position of $e \in E(h_n)$. The other connected component depends on the position of e in h_n.

When the distance between e_n and e is $l = 1$, the second component is formed from H_n for removing the vertices $S = \{x_n, y_n, x_{n-1}, y_b\}$, where $x_{n-1} = N_{H_n}(x_n)$, $y_b = N_{H_n}(y_n)$, $x_{n-1} \neq y_b$. We denote by G_1 such component, where $V(G_1) = V(H_n) - S$. Then, $i(H_{n+1} - (N[x] \cup N[y])) = i(G_1)F_{k-2}$. Notice that $|V(G_1)| = |V(H_n)| - 4$. Furthermore, $\delta_{H_n}(x_n) = \delta_{H_n}(y_n) = \delta_{H_n}(y_b) = 2$, and $\delta_{H_n}(x_{n-1}) = 3$. In this case, *there are 5 edges incident to S from $E(h_n)$ and two edges from $E(h_{n-1})$ incident to S, this is, $|E_{H_n}(S)| = 6$.*

In the case $l > 1$, the second component is formed from H_n due to the same set of vertices S being removed. But in this case, $x_a = N_{H_n}(x_n)$, $y_b = N_{H_n}(y_n)$, $x_a \neq y_b$. Then, $\delta_{H_n}(x_n) = \delta_{H_n}(y_n) = \delta_{H_n}(y_b) = \delta_{H_n}(x_a) = 2$. Let $G_2 = (H_n - S)$, then $G_2 = P_i \cup_{x_{n-1}} H_{n-1} \cup_{y_{n-1}} P_j$ is a polygonal array where the edge $e_{n-1} = \{x_{n-1}, y_{n-1}\}$ from the polygon h_{n-1} is joined to two disjointed paths P_i and P_j. In this case, $i(H_{n+1} - (N[x] \cup N[y])) = i(G_2)F_{k-2}$. Notice that $V(G_2) = V(H_n) - S$, then G_2 and G_1 have the same set of vertices. But in the case $l > 1$, there are only 5 edges from h_n incident to S in H_n. In fact, $G_1 = G_2 - \{x_a, x_{n-1}\}$, and for Lemma 3, $i(G_1) > i(G_2)$. Hence, in order to minimize $i(H_{n+1})$ we must select max$\{i(G_1), i(G_2)\}$ since those are the unique values that can vary in equation 5. As $i(G_1) > i(G_2)$, the maximum corresponds to the case $l = 1$, this is, distance one between common edges in the array minimize $i(H_{n+1})$. ■

When we consider the class of hexagonal arrays theorem 1 solves the Gutman's conjecture [8], showing that the zig-zag hexagonal chain is effectively the topology with a minimum value for $i(H_n)$, as it was already noticed by Zhang [6]. The last theorem generalizes such result, showing that independently of the kind of polygons, considering polygons with more than 5 sides, distance one between adjacent polygons minimize the Merrifield–Simmons index for any polygonal array, even for hexagonal arrays.

Notice that the selection of the vertex v, that holds that $\delta(v) = 3$ and $v \in (N(x) \cup N(y))$, could belong to $N(x)$ or to $N(y)$. Afterwards, two different but symmetrical subgraphs H_n and H'_n can be formed, such that $i(H_n) = i(H'_n)$.

Then, distance one between adjacent polygons can be obtained in clockwise or in counterclockwise direction. Following the same direction for distance one between adjacent polygons (the clockwise or the counterclockwise), then two different but symmetrical array of polygons H_n and H'_n are formed. They hold $i(H_n) = i(H'_n)$, and both topologies achieve the minimum value for the Merrifield–Simmons index.

We have shown that the minimum value for $i(H_n)$ is achieved independently of the number of edges in each polygon in the array, while the distance one between a pair of adjacent polygons is constant following, for example, the clockwise direction.

Since the topology that minimizes $i(H_n)$ is achieved when there is one edge of distance between a pair of adjacent polygons, and if we want to extend the array with an extra polygon and keep a minimum value for $i(H_{n+1})$, then there must be one edge separating the edges: e_{n-1} and e_n. Thus, the direction has to be constant, for example, following the clockwise direction.

There is a question for further analysis that we establish in the following hypothesis.

Conjecture 1 *The polygonal array with a maximum Merrifield–Simmons index is the one that maintains distance two between consecutive common edges.*

6 Conclusions

We have shown how the properties of the product between two Fibonacci numbers can be used for the computation of the Merrifield–Simmons index on polygonal arrays. We have proved that the zig-zag polygonal array (polygonal arrays where each pair of adjacent polygons is joined at distance 1) has the extremal minimum value for the Merrifield–Simmons index. Also, there exists two symmetrical polygonal arrays obtaining the minimum value for the Merrifield–Simmons index.

This result works independently of the length of the polygons in the array, and it generalizes a well known result obtained for hexagonal chains.

Our method does not require the explicit computation of the number of independent sets of the involved graphs, instead it is based on the application of the edge division rule as a way to decompose polygonal array graphs.

Acknowledgments: The authors are grateful to the anonymous referees for their valuable comments, corrections and suggestions, which led to an improved version of this article. The authors also thank SNI-Conacyt México.

References

[1] A. S. Pedersen, P. D. Vestergaard, The number of independent sets in unicyclic graphs, *Discr. Appl. Math.* **152** (2005) 246–256.

[2] Y. Cao, F. Zhang, Extremal polygonal chains on k-matchings, *MATCH Commun. Math. Comput. Chem.* **60** (2008) 217–235.

[3] H. Deng, The smallest Merrifield–Simmons index of $(n, n+1)$-graphs, *Math. Comput. Model.* **49** (2009) 320–326.

[4] I. Gutman, *Advances in the Theory of Benzenoid Hydrocarbons II*, Springer, Berlin, 1992.

[5] S. Wagner, I. Gutman, Maxima and minima of the Hosoya index and the Merrifield–Simmons index, a survey of results and techniques, *Acta Appl. Math.* **112** (2010) 323–346.

[6] Z. Lianzhu, The proof of Gutman's conjectures concerning extremal hexagonal chains, *J. Syst. Sci. Math. Sci.* **04** (1998) 460–465.

[7] L. Z. Zhang, F. Zhang, Extremal hexagonal chains concerning k-matchings and k-independent sets, *J. Math. Chem.* **27** (2000) 319–329.

[8] I. Gutman, Extremal hexagonal chains, *J. Math. Chem.* **27** (1993) 197–210.

[9] F. Zhang, Z. Li, L. Wang, Hexagonal chains with maximal total 3.1416-electron energy, *Chem. Phys. Lett.* **337** (2001) 131–137.

[10] H. Deng, Catacondensed benzenoids and phenylenes with the extremal third–order Randić index, *MATCH Commun. Math. Comput. Chem.* **64** (2010) 471–496.

[11] W. C. Shiu, Extremal Hosoya index and Merrifield–Simmons index of hexagonal spiders, *Discr. Appl. Math.* **156** (2008) 2978–2985.

[12] Z. Zhu, S. Li, L. Tan, Tricyclic graphs with maximum Merrifield–Simmons index, *Discr. Appl. Math.* **158** (2010) 204–212.

[13] H. Ren, F. Zhang, Double hexagonal chains with maximal Hosoya index and minimal Merrifield–Simmons index, *J. Math. Chem.* **42** (2007) 679–690.

[14] I. Gutman, X. Li, H. Zhao, On the Merrifield–Simmons index of trees, *MATCH Commun. Math. Comput. Chem.* **54** (2005) 389–402.

[15] T. Došlić, M. S. Litz, Matchings and independent sets in polyphenylene chains, *MATCH Commun. Math. Comput. Chem.* **67** (2012) 313–330.

[16] X. L. Li, H. X. Zhao, On the Fibonacci numbers of trees, *Fib. Quarter.* **44** (2006) 32–38.

[17] X. Z. Lv, A. M. Yu, The Merrifield–Simmons indices and Hosoya indices of trees with a given maximum degree, *MATCH Commun. Math. Comput. Chem.* **56** (2006) 605–616.

[18] Y. L. Ye, X. F. Pan, H. Q. Liu, Ordering unicyclic graphs with respect to Hoyosa indices and Merrifield–Simmons indices, *MATCH Commun. Math. Comput. Chem.* **59** (2008) 191–202.

[19] A. M. Yu, F. Tian, A kind of graphs with minimal Hosoya indices and maximal Merrifield–Simmons indices, *MATCH Commun. Math. Comput. Chem.* **55** (2006) 103–118.

[20] A. M. Yu, X. Z. Lv, The Merrifield–Simmons indices and Hosoya indices of trees with k pendent vertices, *J. Math. Chem.* **41** (2007) 33–43.