Slovak University of Technology in Bratislava
Faculty of Mechanical Engineering
17th Conference on Applied Mathematics **APLIMAT 2018**

Proceedings

ON THE HOSOYA INDEX FOR THE MOLECULAR GRAPHS OF HELICENES

SEIBERT Jaroslav (CZ), ZAHRÁDKA Jaromír (CZ)

Abstract. The matching of an undirected graph $G = (V, E)$ is a subset M of E such that no two edges of M are adjacent in G. The Hosoya index of a graph G is given by the number of all matching of G. This graph invariant is one of the most interesting topological index in chemistry. Helicenes are extremal hexagonal chains with a simple graph representation as an important subclass of benzenoid molecules. We obtain the exact formula for the Hosoya index of the molecular graphs of helicenes as a function of the number of hexagons in it.

Key words. matching of a graph, Hosoya index, molecular graph, decomposition theorem, difference equation, helicene

Mathematics Subject Classification: Primary 12H10, 39A10; Secondary 11B39

1 Introduction

In this contribution, we consider undirected simple graphs (without loops of multiple edges). A subset M of the edge set of a graph G is called the matching if none two edges of M are adjacent in G. Other graph terminology and notation is taken from the book [1]. Denote by $m(G, k)$ the number k-matching, which means the number of k mutually independent edges can be selected in G.

Definition 1. Let $G = (V, E)$ be a simple connected graph. By definition, $m(G,0) = 1$ for all graphs, and $m(G,1) = |E|$. The Hosoya index of G is given by $Z(G) = \sum_{k \geq 0} m(G, k)$ as the number of all matching in G.

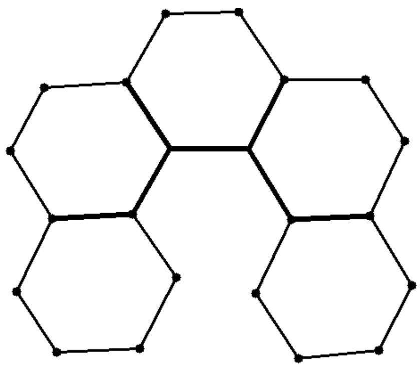
The chemist Haruo Hosoya introduced in 1971 a molecular graph based structure descriptor, which he named the topological index $Z(G)$. He showed that certain physico-chemical properties of saturated hydrocarbons, in particular their boiling points, are well correlated with $Z(G)$. The molecular structural description Z was soon re-named into Hosoya index, whereas the name topological index is used for any of the countless invariants that are found to have some chemical applicability [11].

Many authors found the Hosoya index for various classes of graphs or solved the problem of extremal values of the Hosoya index in special cases. LV and YU [6] investigated the Hosoya index of trees with a given maximum vertex degree. Hamzeh et al. [4] obtained exact formulas of the Hosoya index for the set of bicyclic graphs, caterpillars and dual star. Huang, Kuang and Deng [5] found the average values of the Hosoya index with respect to the set of all polyphenylene chains. Wagner and Gutman [11] collected and classified the results on the extremal values of the Hosoya index and also provided some useful tools and techniques that are frequently used in the study of this type of problem.

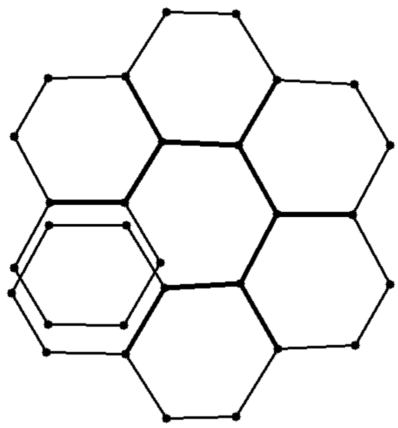
A hexagonal system is a connected plane graph without cut-vertices in which all inner faces are the cycles of the length 6 (hexagons), such that two hexagons are either disjoint or have exactly one common edge, and none three hexagons share a common edge. Two hexagons of a system may have either two common vertices (if they are adjacent) or none (if they are not adjacent). A hexagonal chain with n hexagons, $n \geq 2$, possesses two terminal hexagons and $n-2$ hexagons that have two neighbors. A hexagon adjacent to exactly two other hexagons possesses two vertices of degree 2.

Denote B_n an arbitrary chain with n hexagons and V_3 the set of its vertices of degree 3. The subgraph B_n' of B_n generated by V_3 is an acyclic graph [13]. If the subgraph B_n' is a matching with $n-1$ edges, then B_n is called a linear chain. If the subgraph B_n' is a path, then B_n is called a zigzag chain. Shiu [10] showed that the linear hexagonal spider and zig-zag hexagonal spider attain the extremal values of the Hosoya index. Gutman [3] reported some results on extremal hexagonal chains.

If the subgraph B_n' is a comb, then B_n is called a helicene chain H_n (Fig.1.), where a comb is a graph obtained by joining a single pendant edge to each vertex of a path (see e.g. [9]).



H_5



H_7

Fig. 1. The molecular graphs of $[5]$helicene and $[7]$helicene (the combs are drawn by bold lines).

Carbohelicenes belong to a class of fascinating, chiral, and helicoidal molecules, which have a rich history in chemistry since the very beginning of the 20th century. Helicene chemistry is being considered as an expanding and modern field, leading to several applications in supramolecular chemistry, in nanosciences, in chemical-biology, in polymers and materials science. A comprehensive report on non-stereoselective reactions and methods for producing helicenes can be found in [2].

The main aim of this contribution is to derive the exact formula for the Hosoya index of the molecular graphs of helicenes H_n as a function of n.

2 Preliminary results

In this section, we review some basic and general properties of the Hosoya index (e.g. [4], [11]).

Theorem 1. If $G_1, G_2, \ldots, G_m$ are the connected components of a graph G, then $Z(G) = \prod_{i=1}^{m} Z(G_i)$.

Theorem 2. Let G be a graph with at least two vertices and v be its arbitrary vertex. Then

$$Z(G) = Z(G-v) + \sum_{w} Z(G - \{v, w\}),$$

where the sum is taken over all the vertices w adjacent to v.

Theorem 3. Let vertices u,v be the ends of an edge e in a graph G. Then

$$Z(G) = Z(G-e) + Z(G - \{u, v\}).$$

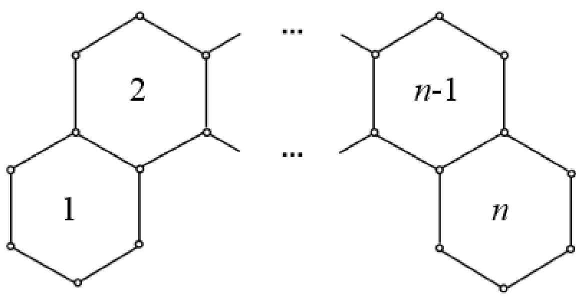
Using Theorem 2 or Theorem 3 it is easy to obtain the following values of the Hosoya index for special graphs. We just recall that the Fibonacci numbers F_n are defined by the second order recurrence $F_{n+2} = F_{n+1} + F_n$, with $F_0 = 0, F_1 = 1$ and the Lucas numbers L_n satisfy the same recurrence but with the initial terms $L_0 = 2, L_1 = 1$.

Theorem 4.

(a) $Z(P_n) = F_{n+1}$, where P_n is a path with n vertices,

(b) $Z(C_n) = L_n$, where C_n is a cycle with n vertices.

Now, we will derive some auxiliary results to finding the Hosoya index for helicenes. Consider therefore the molecular graph H_n of $[n]$helicene and its subgraphs I_n, J_n, K_n and M_n (Fig.2).



H_n

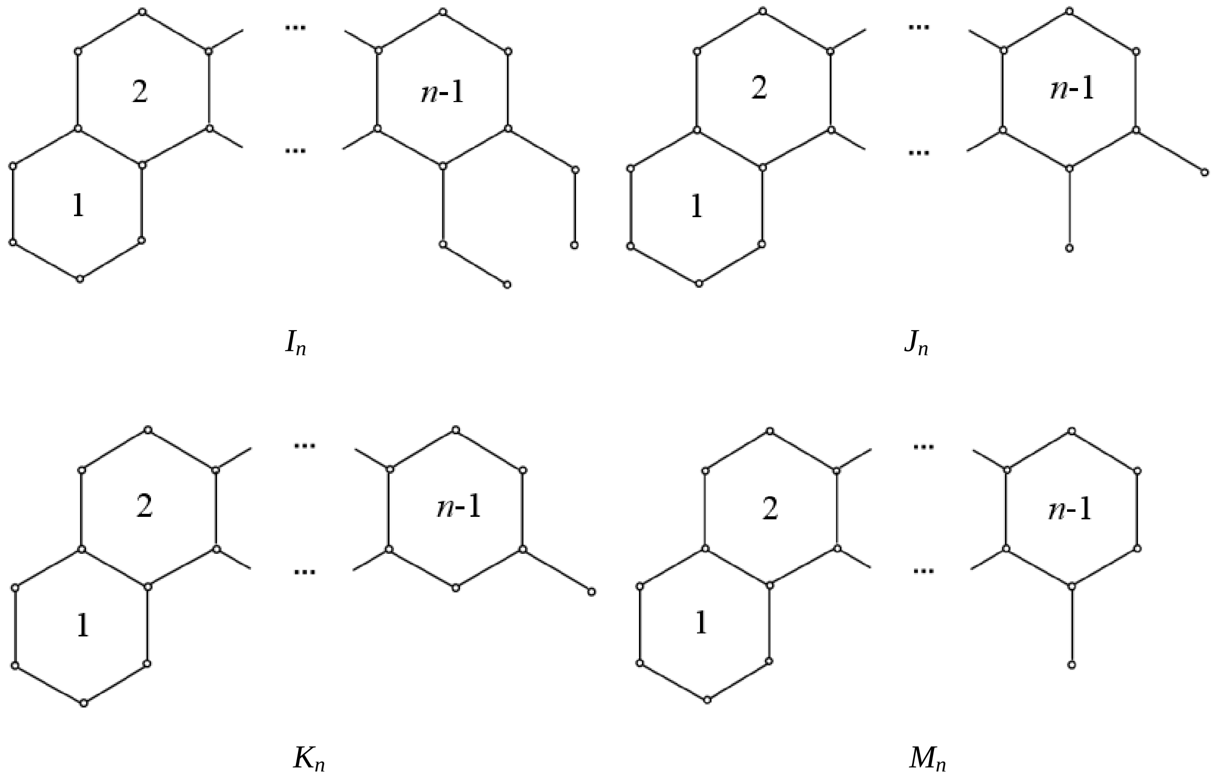


Fig. 2. The molecular graph H_n of $[n]$helicene and its subgraphs I_n, J_n, K_n and M_n.

Let us denote their Hosoya indices $Z(H_n) = h_n$, $Z(I_n) = i_n$, $Z(J_n) = j_n$, $Z(K_n) = k_n$ and $Z(M_n) = m_n$, for short. By direct using of Theorems 1-4 we can obtain the values of these indices for the small numbers of n. These values are collected in Tab. 1.

n	h_n	i_n	j_n	k_n	m_n
1	18	13	5	3	3
2	148	109	39	26	26
3	1233	906	327	213	218
4	10244	7531	2713	1778	1807
5	85169	62605	22564	14764	15033

Tab. 1. The Hosoya index of the graphs H_n, I_n, J_n, K_n, M_n for $1 \le n \le 5$.

Lemma 1. The terms of sequences $\{h_n\}$, $\{i_n\}$, $\{j_n\}$, $\{k_n\}$ and $\{m_n\}$ satisfy the following difference equations for any $n \ge 2$

$$h_n = i_n + j_n \qquad (1)$$

$$i_n = j_n + k_n + m_n + h_{n-1} \qquad (2)$$

$$j_n = k_n + 3k_{n-1} + 2h_{n-2} \qquad (3)$$

$$k_n = h_{n-1} + j_{n-1} + m_{n-1} \qquad (4)$$

$$m_n = h_{n-1} + 2k_{n-1} + h_{n-2} \qquad (5)$$

Proof.

Relation (1) – Using Theorem 3 we choose the edge e of the graph H_n as in Fig. 3. Then we have directly $Z(H_n) = Z(I_n) + Z(J_n)$.

Relation (2) – Using Theorem 2 two times (see Fig. 3) we have successively

$Z(I_n) = Z(I_n - v_1) + Z(I_n - \{v_1, w_1\}) =$

$= Z(I_n') + Z(I_n'') = Z(I_n' - v_2) + Z(I_n' - \{v_2, w_2\}) + Z(I_n'' - v_2) + Z(I_n'' - \{v_2, w_2\})$, but it means that $Z(I_n) = Z(J_n) + Z(K_n) + Z(M_n) + Z(H_{n-1})$ which was to prove.

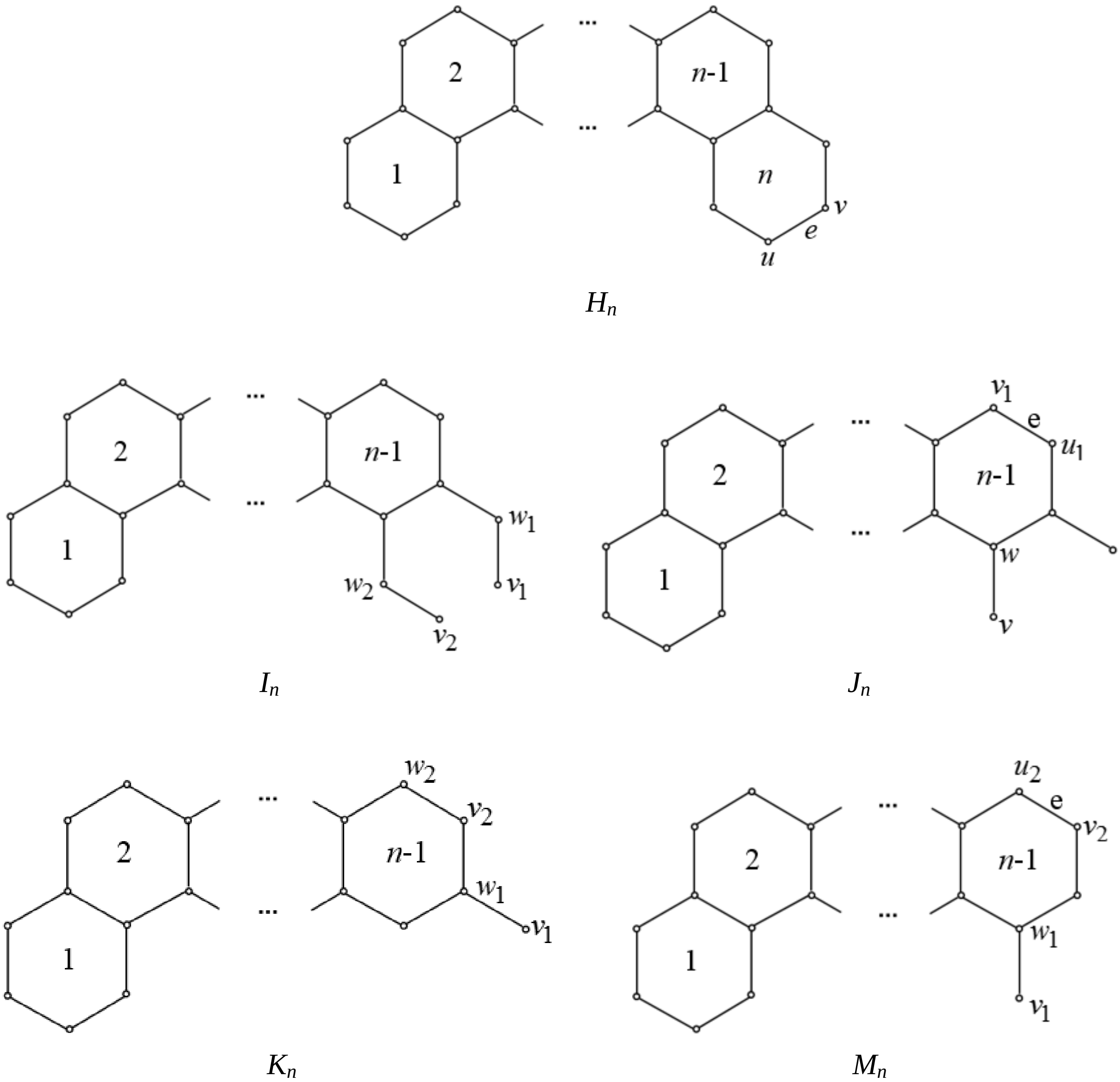


Fig. 3.

Relation (3) – First, we will use Theorem 2 so that we choose the vertex v of the graph J_n as in Fig. 3. Then $Z(J_n) = Z(K_n) + Z(J_n - \{v, w\})$. Now, with the respect of Theorem 3 by using the edge

$e = u_1 v_1$ we obtain $Z(J_n - \{v, w\}) = Z(P_3 \cup K_{n-1}) + Z(P_2 \cup H_{n-2})$. Using Theorem 1 we have $Z(J_n) = Z(K_n) + Z(P_3)\,Z(K_{n-1}) + Z(P_2)\,Z(H_{n-2}) = Z(K_n) + F_4\,Z(K_{n-1}) + F_3\,Z(H_{n-2})$, which gives the proved relation.

Relation (4) – Now, we will again use Theorem 2 two times (see Fig. 3). Thus

$$Z(K_n) = Z(K_n - v_1) + Z(K_n - \{v_1, w_1\}) = Z(H_{n-1}) + Z(K'_n) = Z(H_{n-1}) + Z(K'_n - v_2) + Z(K'_n - \{v_2, w_2\}) =$$
$$= Z(H_{n-1}) + Z(J_{n-1}) + Z(M_{n-1}).$$

Relation (5) – Using Theorem 2 we first obtain

$Z(M_n) = Z(M_n - v_1) + Z(M_n - \{v_1, w_1\}) = Z(H_{n-1}) + Z(M'_n)$. If we want to use Theorem 3 for the graph M'_n it is suitable to choose the edge e as in Fig. 3. Then

$Z(M_n) = Z(H_{n-1}) + Z(M'_n - e) + Z(M'_n - \{u_2, v_2\}) = Z(H_{n-1}) + Z(P_2)\,Z(K_{n-1}) + Z(P_1)\,Z(H_{n-2})$ and since $Z(P_2) = F_3 = 2$, $Z(P_1) = F_2 = 1$ the proof is over.

It is easy to see that $h_0 = Z(P_2) = F_3 = 2$ and therefore relations (1) and (2) are also valid for $n = 1$.

3 Main results

Theorem 5. For the Hosoya index k_n of the graphs K_n the following linear difference equation

$$k_{n+4} - 6k_{n+3} - 19k_{n+2} - 2k_{n+1} + k_n = 0 \tag{7}$$

holds for each positive integer n.

Proof. First, we reduce the number of variables in the system relations from Lemma 1 for any $n \geq 2$

$$j_n = k_n + 3k_{n-1} + 2h_{n-2} \tag{3}$$
$$j_n = k_{n+1} - 2k_{n-1} - h_n - h_{n-1} - h_{n-2} \tag{8}$$
$$2\,j_n = -k_n - 2k_{n-1} + h_n - 2h_{n-1} - h_{n-2} \tag{9}$$

when we obtain relation (8) from relations (4), (5) and relation (9) from relations (1), (2), (5).

We can eliminate j_n by a suitable way and then we have two relations for the members of the sequences $\{h_n\}$ and $\{k_n\}$

$$k_{n+2} - 4k_{n+1} - 13k_n = 3h_n + 8h_{n-1} \tag{10}$$
$$k_{n+3} - 6k_{n+2} - 14k_{n+1} + 2k_n = 8h_n - h_{n-1} \tag{11}$$

and the equality

$$h_n = \frac{1}{67}\left(8k_{n+3} - 47k_{n+2} - 116k_{n+1} + 3k_n\right) \tag{12}$$

is directly obtained from the previous system.

Finally, we substitute (12) into (10) and after simplification we obtain difference equation (7).

Remark. It is known fact that each of the sequences from Lemma 1 satisfy the same linear difference equation as the sequence $\{k_n\}$. Therefore the general solution for these sequences are the same, but the concrete function expression (particular solution) of the members of sequences will be different with respect to the various initial members of sequences.

Theorem 6. The Hosoya index of the molecular graph H_n of $[n]$helicene can be written for any positive integer n in the form $h_n = c_1\alpha^n + c_2\beta^n + c_3\gamma^n + c_4\delta^n$, where the constants c_1, c_2, c_3, c_4 can be uniquely expressed by the numbers $\alpha, \beta, \gamma, \delta$.

Proof. The members of the sequence $\{h_n\}$ satisfy the homogeneous linear difference equation of the fourth order with constant coefficients

$$h_{n+4} - 6h_{n+3} - 19h_{n+2} - 2h_{n+1} + h_n = 0.$$

Its characteristic equation is an algebraic equation of the fourth order

$$x^4 - 6x^3 - 19x^2 - 2x + 1 = 0$$

with the different roots $\alpha, \beta, \gamma, \delta$. It means that $h_n = c_1\alpha^n + c_2\beta^n + c_3\gamma^n + c_4\delta^n$ and with help of the initial members $h_0 = 2, h_1 = 18, h_2 = 148, h_3 = 1233$ we obtain the following system of four linear equations for unknown c_1, c_2, c_3, c_4

$$\begin{aligned} c_1 + c_2 + c_3 + c_4 &= 2 \\ c_1\alpha + c_2\beta + c_3\gamma + c_4\delta &= 18 \\ c_1\alpha^2 + c_2\beta^2 + c_3\gamma^2 + c_4\delta^2 &= 148 \\ c_1\alpha^3 + c_2\beta^3 + c_3\gamma^3 + c_4\delta^3 &= 1233. \end{aligned}$$

It is easy to find the solution of this system in the symbolic form

$$c_1 = \frac{1233 - 148(\beta+\gamma+\delta) + 18(\beta\gamma+\beta\delta+\gamma\delta) - 2\beta\gamma\delta}{(\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)}$$

$$c_2 = \frac{1233 - 148(\alpha+\gamma+\delta) + 18(\alpha\gamma+\alpha\delta+\gamma\delta) - 2\alpha\gamma\delta}{(\beta-\alpha)(\beta-\gamma)(\beta-\delta)}$$

$$c_3 = \frac{1233 - 148(\alpha+\beta+\delta) + 18(\alpha\beta+\alpha\delta+\beta\delta) - 2\alpha\beta\delta}{(\gamma-\alpha)(\gamma-\beta)(\gamma-\delta)}$$

$$c_4 = \frac{1233 - 148(\alpha+\beta+\gamma) + 18(\alpha\beta+\alpha\gamma+\beta\gamma) - 2\alpha\beta\gamma}{(\delta-\alpha)(\delta-\beta)(\delta-\gamma)}.$$

Then the proof is finished.

4 Numerical results

We also had to calculate the previous results using Symbolic Math Toolbox of MATLAB.

The roots of the characteristic equation $x^4 - 6x^3 - 19x^2 - 2x + 1 = 0$ can be found by MATLAB in the numerical form through the command

```
>> syms x, R=solve(x^4-6*x^3-19*x^2-2*x+1)
R =
  8.3128260775610574875139058467921
 0.17898242638989329333757646913108
```

-0.3077323367373107047565344171426
-2.1840761672136400760949478987805

The roots are four irrational numbers, which were denoted by $\alpha, \beta, \gamma, \delta$ in Theorem 6. It is obvious that the above mentioned values are only approximations these numbers.

The system of four linear equations for unknown c_1, c_2, c_3, c_4 was solved by using commands of MATLAB. Instead of the variables $\alpha, \beta, \gamma, \delta$ are used in MATLAB the symbolic variables a, b, c, d and instead c_1, c_2, c_3, c_4 are used the identifiers C1, C2, C3, C4.

```
>> syms a b c d , A = [1,1,1,1; a,b,c,d; a^2, b^2, c^2, d^2 ; a^3, b^3, c^3, d^3];
>> AC1=A; AC1(:,1)=[2; 18; 148; 1233]; C1=factor(det(AC1))/factor(det(A))
C1 =
-(148*b + 148*c + 148*d - 18*b*c - 18*b*d - 18*c*d + 2*b*c*d - 1233)/((a - b)*(a - c)*(a - d))
>> AC2=A; AC2(:,2)=[2; 18; 148; 1233]; C2=factor(det(AC2))/factor(det(A))
C2 =
(148*a + 148*c + 148*d - 18*a*c - 18*a*d - 18*c*d + 2*a*c*d - 1233)/((a - b)*(b - c)*(b - d))
>> AC3=A; AC3(:,3)=[2; 18; 148; 1233]; C3=factor(det(AC3))/factor(det(A))
C3 =
-(148*a + 148*b + 148*d - 18*a*b - 18*a*d - 18*b*d + 2*a*b*d - 1233)/((a - c)*(b - c)*(c - d))
>> AC4=A; AC4(:,4)=[2; 18; 148; 1233]; C4=factor(det(AC4))/factor(det(A))
C4 =
(148*a + 148*b + 148*c - 18*a*b - 18*a*c - 18*b*c + 2*a*b*c - 1233)/((a - d)*(b - d)*(c - d))
```

Calculation of approximations of the irrational constants c_1, c_2, c_3, c_4 can be realized by command

```
>> c1=subs(C1,[a,b,c,d],R), c2=subs(C2,[a,b,c,d],R), c3=subs(C3,[a,b,c,d],R), c4=subs(C4,[a,b,c,d],R)
c1 =
2.1454822181534150647686192500525
c2 =
0.046858209049296998559385477378165
c3 =
-0.14042920424633548952728375521586
c4 =
-0.051911222956376573800720972214777 .
```

The values of Hosoya index calculated by using MATLAB through the recurrence formula $h_{n+4} = 6h_{n+3} + 19h_{n+2} + 2h_{n+1} - h_n$ for the positive integers $1 \le n \le 20$ are

```
>> h=[sym(18) sym(148) sym(1233) sym(10244)];
>> for  n=1:16, h(n+4)= 6*h(n+3)+19*h(n+2)+2*h(n+1)-h(n); end, h
h =
[ 18, 148, 1233, 10244, 85169, 707968, 5885274, 48923130, 406689753, 3380740568,
28103509701, 233619585374, 1942038987946, 16143832328616, 134200870403717,
```

1115588495056524, 9273693133679053, 77090598116627092, 640840734359373890, 5327197568144459670] .

The values of Hosoya index can also be calculated through the previously derived formula $h_n = c_1\alpha^n + c_2\beta^n + c_3\gamma^n + c_4\delta^n$ for the positive integers $1 \le n \le 20$.

```
>> syms n, H=c1*R(1)^n+ c2*R(2)^n+ c3*R(3)^n+ c4*R(4)^n;
>> for  k=1:20, h(k)=subs(H,n,sym(k)); end, h
 h =
 [ 18.0, 148.0, 1233.0, 10244.0, 85169.0, 707968.0, 5885274.0, 48923130.0, 406689753.0,
3380740568.0, 28103509701.0, 233619585374.0, 1942038987946.0, 16143832328616.0,
134200870403717.0, 1115588495056524.0, 9273693133679053.0, 77090598116627092.0,
640840734359373890.0, 5327197568144459670.0]
```

It is easy to see that both used methods of calculating of the Hosoya index of the graphs H_n for the positive integers $1 \le n \le 20$ give identical results with respect to integers. The Hosoya index of $[n]$helicenes for $3 \le n \le 10$ is presented in Tab. 2.

n	3	4	5	6	7	8	9	10
h_n	1 233	10 244	85 169	707 968	5 885 274	48 923 130	406 689 753	3 380 740 568

Tab. 2.

5 Concluding remarks

The Hosoya index of a graph, also known as Z-index, is one of the large family of topological indices of graphs which are closely connected with selected physico-chemical characteristics of the respective compounds. Similar connections are also known for the Merrifield-Simmons index of a graph, which is defined via the number of ways in which mutually independent vertices can be selected in a graph. The two indices do not only have very similar definitions, they are also quite related in another respect. Mostly it is true that the graph which minimizes the Merrifield-Simmons index is also the one which maximizes the Hosoya index, and vice versa.

In the previous years we have calculated the Merrifield-Simmons index for special classes of hydrocarbons. In this contribution we have turned attention to the Hosoya index for one interesting class of molecular graphs. The methods of calculation are possible to use for other classes of the molecular graphs. The main problem is again in a rather complicated solution of system of difference equations which often leads only to numeric results.

Acknowledgment. The authors gratefully acknowledge the anonymous referee for his useful remarks and comments.

References

[1] DIESTEL, R. *Graph Theory*, Springer, 2006.

[2] GINGRAS, M. *One hundreds years of helicene chemistry. Part 1: non-stereoselective syntheses of carbohelicenes.* Chem. Soc. Rev., **42** (2013), 968-1006.

[3] GUTMAN, I. *Extremal hexagonal chains*. Journal of Mathematical Chemistry, **12** (1993), 197–210.

[4] HAMZEH, A., IRANMANESH, A., HOSSEIN-ZADEH, S. et al. *The Hosoya index and the Merrifield-Simmons index of some graphs.* Transactions on Combinatorics, **1** (2012), 51-60.

[5] HUANG, G., KUANG, M., DENG, H. *The expected values of Hosoya index and Merrifield-Simmons index in a random polyphenylene chain.* J. Comb. Optim., **12** (2015), 55-67.

[6] LV, X., YU, A. *The Merrifield-Simmons indices and Hosoya indices of trees with a given maximum degree.* MATCH Commun. Math. Comput. Chem., **56** (2006), 605-616.

[7] MERRIFIELD, R. E., SIMMONS, H. E. *Topological Methods in Chemistry*, Wiley, New York, 1989.

[8] PAN, X.F., XU, J.M., YANG, Ch. et al. *Some graphs with minimum Hosoya index and maximum Merrifield-Simmons index.* MATCH Commun. Math. Comput. Chem., **57** (2007), 235-242.

[9] SANDHYA, S. S., RAY, C. D., JAYASEKARAN, C. *Some new results on harmonic mean graphs.* International Journal of Mathematical Archive, **4** (2013), 240-245.

[10] SHIU, W. C. *Extremal Hosoya index and Merrifield-Simmons index of hexagonal spiders.* Discrete Applied Mathematics, **156** (2008), 2978-2985.

[11] VAGNER, S., GUTMAN, I. *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.

[12] YEH, L. *Fibonacci numbers of product graphs.* The Journal of Combinatorial Mathematics and Combinatorial Computing, **32**(2000), 223-229.

[13] ZHANG, L. *Extremal hexagonal chains concerning largest eigenvalue.* Science of China (Series A), **44** (2001), 1089-1097.

Current address

Seibert Jaroslav, doc. RNDr., CSc.
Institute of Mathematics and Quantitative Methods, Faculty of Economics and Administration
University of Pardubice
Studentská 95, 532 10 Pardubice, Czech Republic
E-mail: jaroslav.seibert@upce.cz

Zahrádka Jaromír, RNDr., PhD.
Institute of Mathematics and Quantitative Methods, Faculty of Economics and Administration
University of Pardubice
Studentská 95, 532 10 Pardubice, Czech Republic
E-mail: jaromir.zahradka@upce.cz