STUDIA UNIVERSITATIS BABEŞ-BOLYAI, CHEMIA, XLVII, 1-2, 2002

Dedicated to Professor Ionel Haiduc
on the occasion of his 65th birthday

CLUJ POLYNOMIALS

MIRCEA V. DIUDEA

"Babes-Bolyai" University, Faculty of Chemistry and Chemical Engineering, 11 Arany Janos str., RO-3400, Cluj-Napoca, Romania

ABSTRACT. A novel class of distance property polynomials *P*(*G*,*x*) is proposed, as an extension of the well-known Hosoya polynomial. The polynomial coefficients are calculated by means of layer/shell matrices, built up according to the vertex distance partitions of a graph. Basic definitions and properties for the Cluj matrices and corresponding polynomials, as particular cases of the distance property polynomial, are given.

INTRODUCTION

A graph can be described by: a connection table, a sequence of numbers, a derived number (called sometimes a topological index), a matrix, or a polynomial. Quantum chemistry was the first field in Chemistry that used the polynomial description of a molecular graph. In the early Hűckel theory, the roots of the most studied *characteristic polynomial*:

$$Ch(G,x) = \det[x\mathbf{I} - \mathbf{A}(G)] \tag{1}$$

with **I** being the unit matrix of a pertinent order and **A** the adjacency matrix, are assimilated to the π-electron energy levels of the molecular orbitals in conjugated hydrocarbons. Other related topics are: Topological Resonance Energy TRE, Topological Effect on Molecular Orbitals, TEMO, the Aromatic Sextet Theory, AST, the Kekulé Structure Count, KSC, *etc*.[1-3]

The coefficients a_k of the characteristic polynomial of order N are calculable from the graph G on N vertices:

$$Ch(G,x) = \sum_{k=0}^{N} a_k(G) \cdot x^{N-k} \tag{2}$$

Relation (2), found independently by Sachs, Harary, Milić, Spialter, *etc*.,[2] makes use of the *Sachs graphs*, contained as subgraphs in *G*. More efficient are the numeric methods of linear algebra, such as the recursive algorithms of Le Verier, Frame, or Fadeev.[4,5]

An extension of relation (1) was made by Hosoya[6] and others[7-10] by changing the adjacency matrix with the distance matrix and next by any square topological matrix.

A different field using the polynomial description is that of finite sequences[2] of some graph invariants, such as the distance degree sequence or the sequence of the number of k-independent edge sets. The polynomial corresponding to the last sequence was introduced by Hosoya as the Z-counting polynomial.[11] The polynomial roots and coefficients are used for characterization of the topological nature of hydrocarbons.[2,3,11]

The present paper introduces novel distance-based sequence polynomials whose coefficients are calculable from two kinds of layer matrices.

BASIC DEFINITIONS

Define a *distance property polynomial* as:

$$P(G,x) = \sum_{k=0}^{d(G)} p(G,k) \cdot x^k \tag{3}$$

with $p(G,0) = P = \sum_i p_i$. In relation (1), $p(G,k)$ is twice the contribution to the global (molecular) property $P=P(G)$ of the vertex pairs located at distance k from each other, in the graph G. The summation runs from zero to $d(G)$, which is the *diameter* of G or the longest distance in the graph.[2,12]

When the local property p_i = 1 (*i.e.*, the vertex cardinality), $p(G,k)$ denotes the number of pair vertices separated by distance k in G, and the classical Hosoya polynomial[13] (more exactly twice this polynomial) is recovered. In this case, $p(G,0) = N$, where N stands for the number of vertices in the hydrogen depleted molecular graph.

The polynomial coefficients $p(G,k)$ are calculable as the column sums in the layer matrices **LM** and **SM**. They are non-square arrays collecting *shells of property*, located at distance k around each vertex.

Let us define the k^{th} layer/shell of vertices v with respect to the vertex i as:

$$G(i)_k = \left\{ v \middle| \; v \in V(G); \quad d_{iv} = k \right\} \tag{4}$$

The collection of all its layers defines the partition of G with respect to i:

$$G(i) = \left\{ G(i)_k \; ; \, k \in [0,1,..,ecc_i] \right\} \tag{5}$$

with ecc_i being the *eccentricity* of i (*i.e.*, the largest distance from i to the other vertices in G). Since **LM** was defined elsewhere,[14] (see also ref. 15) we give here only the *shell matrix* **SM**, the entries of which are defined as:

$$[\mathbf{SM}]_{i,k} = \sum_{v | d_{i,v} = k} [\mathbf{M}]_{i,v} \tag{6}$$

where **M** is any square topological matrix. Any other operation over the square matrix entries $[\mathbf{M}]_{i,v}$ can be used. The shell matrix is a collection of the above defined entries:

$$\mathbf{SM} = \left\{ [\mathbf{SM}]_{i,k}\, ; \, i \in V(G);\, k \in [0,1,..,d(G)] \right\} \tag{7}$$

The zero column vector $[\mathbf{SM}]_{i,0}$ collects the diagonal entries in the parent square matrix. In the case that they are zero, by definition, $[\mathbf{SM}]_{i,0}$ =1. The property p_i is usually introduced by means of the zero column. An example is given for the shell Cluj matrix of the graph 1 (see below).

The name of a property polynomial is built up by exchanging the letter *P* for a string including: L (or S) for the type of layer matrix, symbol of the *info matrix* **M**, and the local property p_i. For example, *SUCJ*(*G,x*) reads: the polynomial of the Shell, Unsymmetric Cluj matrix (see below). In the case of a graph theoretical property, p_i is implicitly 1, and therefore omitted.

Vertex contributions to the global polynomial can be written as:

$$P(i,x) = \sum_{k=0}^{d(G)} p(i,k) \cdot x^k \tag{8}$$

where *p*(*i,k*) is the contribution of vertex *i* to the partition *p*(*G,k*) of the global molecular property *P*. Note that *p*(*i,k*)'s are just the entries in **LM** or **SM**.

Usually, the contribution of vertices (*i.e.*, atoms) to the molecular property vary in a molecular graph, so that the polynomial for the whole molecule is obtained by summing all atomic contributions:

$$P(G,x) = \sum\nolimits_i P(i,x) \tag{9}$$

In a vertex transitive graph, the vertex contribution is simply multiplied by *N*:

$$P(G,x) = N \cdot P(i,x) \tag{10}$$

A *distance-extended property* can be calculated by evaluating the *first derivative* of the polynomial, for *x* = 1:

$$P'(G,1) = \sum_{k=1}^{d(G)} k \cdot p(G,k) = D\_P(G) \tag{11}$$

Any square matrix can be used as an info matrix for the layer matrices, thus resulting in an unlimited number of property polynomials. The property *P* can be taken either as a crude property (*i.e.*, column zero in **LM**) or within some weighting scheme (*i.e.*, transformed by the sequence: **W**-operator **W**(**M**1,**M**2,**M**3), **W**(**M**) matrix, **LM**/**SM**).[2] In the present paper we limit discussion to some graph theoretical properties. Various property polynomials, using physico-chemical properties are exposed in a following paper.[16]

HOSOYA-LIKE POLYNOMIAL

In the case: $p_i = 1$, **LM** = **LC**, (*i.e.*, layer matrix of cardinalities) and the property polynomial, called the cardinality polynomial $C(G,x)$, is twice the Hosoya polynomial. The formulas given in the following represent well-known results.[13,17,18]

The index calculated as the polynomial first derivative[13,19] is (twice) the well-known Wiener index,[20] W. It is just a distance-extended property.

$$C'(G,1) = \sum_{k=1}^{d(G)} k \cdot p(G,k) = 2 \cdot W \qquad (12)$$

The hyper-Wiener index WW, originated by Randić,[21] is calculated as:

$$WW(G) = W(G) + \Delta(G) \qquad (13)$$

where $\Delta(G)$ is the non-Wiener part[22,23] of the hyper-Wiener number, calculable from the second derivative[17] of the Hosoya polynomial.

$$WW(G) = H'(G,1) + (1/2)H''(G,1) \qquad (14)$$

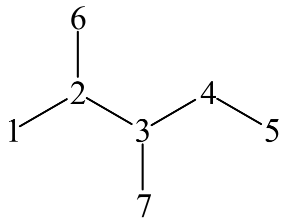
In terms of $C(G,x)$, the relation is trivially deduced from (14):

$$WW(G) = (1/2) \cdot C'(G,1) + (1/4)C''(G,1) \qquad (15)$$

For the graph G_1, the Cardinality polynomial is:

$$C(G,x) = LC(G,x) = 7 + 12x + 14x^2 + 12x^3 + 4x^4 \qquad (16)$$

Real Roots = -0.723, -0.723, -0.134, -0.134.



G_1

CLUJ POLYNOMIALS

Cluj polynomials, as particular cases of the distance property polynomial, make use of the shell matrix **SM** for re-arrange the basic Cluj matrices.

Define the *Cluj fragments*[2,23] $CJ_{i,j,p}$ as the set of vertices v lying closer to i than to j and at least one path $p(i,v)$ exists so that it intersects the path $p(i,j)$ at most in i.

$$CJ_{i,j,p} = \left\{v \middle| v \in V(G);\, d(G)_{i,v} < d(G)_{j,v};\, \text{and } \exists\ p(i,v) \cap p(i,j) = \{i\}\right\} \quad (17)$$

The above definition holds in any undirected graph. The intersecting condition means at least one path $p(i,v)$ is external to the "prohibited" path $p(i,j)$. In trees, due to the unicity of paths joining any two vertices, $CJ_{i,j,p}$ means the number of paths going to j through i. In this way, we characterize the path $p(i,j)$, designed as p hereafter, by a single endpoint, that suffices for the unsymmetric Cluj matrix **UCJ**.[23]

In cycle-containing graphs, more than one path could join the pair (i,j), thus resulting more than one fragment related to i (with respect to j and a given path p). By definition, the entries in the Cluj matrix are taken as the maximum cardinality among all such fragments/sets:

$$[\mathbf{UCJ}]_{\boldsymbol{i,j}} = \max_{p} \left| CJ_{i,j,p} \right| \quad (18)$$

When the path p belongs to the set of *distances* (*i.e.*, geodesics) $D(G)$, the corresponding Cluj matrix is designed as **UCJDi**. When $p \in \Delta(G)$ (*i.e.*, detours), the symbol will be **UCJDe**.

A variant of Cluj fragmentation,[2,24] called $CF_{i,j,p}$ considers all the paths $p(i,v)$ external to $p(i,j)$. It is possible by cutting the "prohibited" path $p(i,j)$, excepting its endpoints. Thus, relation (17) becomes:

$$CF_{i,j,p} = \left\{v \middle| v \in V(G);\, d(G_p)_{i,v} < d(G_p)_{j,v};\ G_p = G - p(i,j)\right\} \quad (19)$$

Relation (18) also holds, in terms of *CF* fragments. The corresponding matrices, denoted by **UCFDi** and **UCFDe**, are in general unsymmetric, apart from some symmetric graphs. This is also true for the **CJ** matrices. They can be made symmetric by the Hadamard multiplication with their transposes:

$$\mathbf{M}_p = \mathbf{UM} \bullet (\mathbf{UM})^{\mathrm{T}} \quad (20)$$

$$\mathbf{M}_e = \mathbf{M}_p \bullet \mathbf{A} \quad (21)$$

The subscript p means that the matrix is defined on paths (*i.e.*, on all pair vertices) while e designates an edge defined matrix. Note that, in trees, the four variants of Cluj matrices are one and the same, so that the symbol **CJ** (unless otherwise specified) will hereafter be used. Only tree graphs are considered here. The shell matrix and its parent **UCJ** for the graph 1 are illustrated in Table 1. Relations for the column sums are given in the bottom of the table.

Table 1

SUCJ(G_1)								**UCJ**(G_1)							
i \ *k*	0	1	2	3	4	$\sum_{k=1}^{d(G)}$		1	2	3	4	5	6	7	*RS*
1	1	1	2	2	1	6		0	1	1	1	1	1	1	6
2	1	15	6	3	0	24		6	0	3	3	3	6	3	24
3	1	15	13	0	0	28		4	4	0	5	5	4	6	28
4	1	8	4	4	0	16		2	2	2	0	6	2	2	16
5	1	1	1	2	2	6		1	1	1	1	0	1	1	6
6	1	1	2	2	1	6		1	1	1	1	1	0	1	6
7	1	1	2	3	0	6		1	1	1	1	1	1	0	6
CS	7	42	30	16	4	92[a]	*CS*	15	10	9	12	17	15	14	92
CS·k		42	60	48	16	166[b]									

(a) 2x*W*; (b) 2x*WW*

For the graph G_1, several Cluj polynomials are exemplified:

$$SUCJ(G,x) = 7 + 42x + 30x^2 + 16x^3 + 4x^4 \tag{22}$$

Real roots = -5.264 -0.389 -0.173 -0.173

$$SURCJ(G,x) = 7 + 8.656x + 5.645x^2 + 3.642x^3 + 0.916x^4 \tag{23}$$

Real roots = -0.783 -0.391 -0.031 -0.031

$$SCJ(G,x) = 7 + 92x + 50x^2 + 20x^3 + 4x^4 \tag{24}$$

Real roots = -12.593 -0.305 -0.122 -0.122

$$SIUCJ(G,x) = -3.975 - 1.804x + 2.464x^2 + 3.225x^3 + x^4 \tag{25}$$

Real roots = -0.515 -0.515 -0.486 1.061

$$SWUCJ(G,x) = 92 + 42x + 30x^2 + 16x^3 + 4x^4 \tag{26}$$

Real roots = -0.324 -0.324 0.096 0.096

In the above polynomials, the matrices are as follows: unsymmetric Cluj (22); unsymmetric reciprocal Cluj (23); symmetric Cluj (24); inverse unsymmetric Cluj (25) and walk (of rank 1) of unsymmetric Cluj (26). The free term in (22) to (24) equals the number of vertices in *G*, which is taken by definition for the zero diagonal matrices. The different free term in the last two examples indicates matrices having non-zero diagonals. The diagonal entries in the walk matrix of rank *e* (relation 26) equals the row sum in the matrix **M** raised to the power *e*, $[{}^{e}\mathbf{WM}]_{ii} = R(\mathbf{M}^{e})_i$. Note that in (26) the free term is twice the Wiener number.

The walk matrix **W**(**M**) is constructed by following the $^{e}\mathbf{WM}$ algorithm,[14] extended for walk count in general graphs and any square matrix. The walk matrix together with the triple matrix walk operator[25] **W**(**M**1,**M**2,**M**3) is used for mixing the info matrices and weighting the topological descriptors (property polynomials included) by various physico-chemical attributes. Details are given in a following paper of this topic.[16] The calculations were performed by the TOPOCLUJ software package.

* * *

Reverting to the polynomials, the distance-extended property, calculated on the Cluj polynomial (*cf*. (11)), is, in trees, twice the hyper-Wiener index:

$$SUCJ'(G,1) = 2 \cdot WW \tag{27}$$

The sum of polynomial coefficients (for $k > 0$) gives twice the Wiener index:

$$SUCJ(G,1,)_{k>0} = 2 \cdot W \tag{28}$$

In case of *SUCJ*(*G*,*x*), the non-Wiener part of the hyper-Wiener index is:

$$2 \cdot \Delta(G) = SUCJ'(G,1) - SUCJ(G,1,)_{k>0} \tag{29}$$

In case of a symmetric Cluj matrix, the sum of polynomial coefficients gives twice the hyper-Wiener index:

$$SCJ(G,1)_{k>0} = 2 \cdot WW \tag{30}$$

The distance-extended property leads, in this case, to twice the Tratch index.[26]

The non-Wiener part of the hyper-Wiener index $\Delta(G)$ (see eq 13) is now obtained as the difference of the Cluj polynomial coefficients for the symmetric **CJ** and unsymmetric **UCJ** matrices.

$$2 \cdot \Delta(G) = SCJ(G,1) - SUCJ(G,1) \tag{31}$$

CONCLUSIONS

Extension of the well-known Hosoya polynomial, grounded on vertex distance partitions of a graph, resulted in a novel class of distance property polynomials *P*(*G*,*x*). The polynomial coefficients are obtained as the column sums in the layer/shell matrices. The polynomial roots and coefficients can be used for the topological (and chemical) characterization of chemical structures. Examples were given for the Cluj matrices and corresponding polynomials.

Acknowledgements. The paper is supported by the CNCSIS Romanian GRANT, 2002. The author thank to the referees for useful comments.

REFERENCES

1. M. V. Diudea and O. Ivanciuc, *Molecular Topology,* Comprex, Cluj, 1995 (in Romanian).
2. M.V. Diudea; I. Gutman; L. Jäntschi, *Molecular Topology*. Nova Science, Huntington, New York, 2001.
3. N. Trinajstić, *Chemical Graph Theory*, IInd Ed. CRC Press, 1992.
4. P. S. Dwyes, *Linear Computations*, Wiley, N. Y. 1951.
5. D. K. Fadeev and I. S. Sominskii, *Problems in Higher Algebra*, Freeman, San Francisco, 1965.
6. H. Hosoya, M. Murakami and M. Gotoh, Distance polynomial and characterization of a graph. *Natl. Sci. Rept. Ochanomizu Univ.,* **1973**, *24*, 27-34.
7. R. L. Graham and L. Lovasz, Distance matrix polynomials of trees. *Adv. Math*. **1978**, *29*, 60-88.
8. M. V. Diudea**,** O. Ivanciuc, S. Nikolić, and N. Trinajstić**,** Matrices of Reciprocal Distance. Polynomials and Derived Numbers. *Commun. Math. Comput. Chem.* (*MATCH*), **1997**, *35*, 41-64.
9. O. Ivanciuc, M. V. Diudea, and P. V. Khadikar, New Topological Matrices and Their Polynomials. *Indian J. Chem.* **1998**, *37A*, 574-585.
10. O. Ivanciuc, T. Ivanciuc, and M. V. Diudea, Polynomials and Spectra of Molecular Graphs. *Roum. Chem. Quart. Rev*. **1999,** *7***,** 41-67.
11. H. Hosoya, Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. *Bull. Chem. Soc. Japan*, **1971**, *44*, 2332-2339.
12. F. Harary, *Graph Theory*, Addison - Wesley, Reading, M.A., 1969.
13. H. Hosoya, On some counting polynomials in chemistry. *Discrete Appl*. *Math*., 1988, *19*, 239-257.
14. M. V. Diudea, Layer Matrices in Molecular Graphs. *J. Chem. Inf. Comput. Sci.,* **1994**, *34*, 1064-1071.
15. M. Randić, Graph valence shells as molecular descriptors. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 627-630.
16. O. Ursu and M. V. Diudea, Layer Matrices and Distance Property Polynomials. *Commun. Math. Comput. Chem.* (*MATCH*), submitted.
17. G. G. Cash, Relationship between the Hosoya polynomial and the hyper - Wiener index. *Appl. Math. Lett*. **2001**, *14*, in press.

18. I. Gutman, S. Klavžar, M. Petkovšek, and P. Žigert, On Hosoya polynomials of benzenoid graphs. *Commun. Math. Chem.* (*MATCH*), 2001, *43*, 49-66.

19. E. V. Konstantinova and M. V. Diudea, The Wiener polynomial derivatives and other topological indices in chemical research. *Croat. Chem. Acta,* **2000**,*73*, 383-403.

20. H. Wiener, Structural determination of paraffin boiling points. *J. Am. Chem. Soc.,* **1947***, 69*, 17-20.

21. M. Randić, Novel molecular descriptor for structure-property studies. *Chem. Phys. Lett.*, **1993**, *211*, 478-483.

22. M. V. Diudea, Wiener and hyper-Wiener numbers in a single matrix. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 833-836.

23. M. V. Diudea, Cluj matrix invariants. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 300- 305.

24. L. Jäntschi, G. Katona, and M. V. Diudea, Modeling Molecular Properties by Cluj Indices. *Commun. Math. Comput. Chem.* (*MATCH*), **2000**, *41*, 151-188

25. M. V. Diudea and Randić, Matrix Operator, $W_{(M1,M2,M3)}$ and Schultz-Type Numbers. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 1095-1100.

26. S. S. Tratch, M. I. Stankevitch, and N. S. Zefirov, Combinatorial models and algorithms in Chemistry. The expanded Wiener number – a novel topological index. *J. Comput. Chem*., **1990**, *11*, 899-908.