CROATICA CHEMICA ACTA
CCACAA **80** (2) 271–276 (2007)

ISSN-0011-1643
CCA-3169
Conference Paper

Hosoya Matrices as the Numerical Realization of Graphical Matrices and Derived Structural Descriptors*

Dušanka Janežič,[a] Bono Lučić,[b] Ante Miličević,[c] Sonja Nikolić,[b,] Nenad Trinajstić,[b] and Damir Vukičević[d,**]**

[a]*National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia*

[b]*The Rugjer Bošković Institute, P. O. Box 180, HR-10002 Zagreb, Croatia*

[c]*The Institute of Medical Research and Occupational Health, P. O. Box 291, HR-10002 Zagreb, Croatia*

[d]*Department of Mathematics, University of Split, Nikole Tesle 12, HR-21000, Split, Croatia*

RECEIVED NOVEMBER 6, 2006; REVISED FEBRUARY 7, 2007; ACCEPTED FEBRUARY 9, 2007

Keywords
alkanes
computer program
double invariants
graphical matrices
Hosoya matrices
Hosoya-Wiener indices
numerical matrices
trees

Graphical matrices are used to generate the Hosoya matrices which in turn produce the Hosoya-Wiener indices. The computer program to generate graphical matrices of acyclic structures, the corresponding numerical matrices and double invariants is delineated.

INTRODUCTION

In 1971, Haruo Hosoya introduced a molecular descriptor[1] that became known in the literature as the Hosoya index.[2,3] This descriptor has been amply used in the structure-property-activity modeling.[4] In the same paper, Hosoya also introduced the term topological index that has remained in everyday use[5] and presented a new way to compute the Wiener index of a given structure from its distance matrix.

In 1994, Randić proposed a novel graph-theoretical matrix that he named the Hosoya matrix.[6] He derived two versions of the Hosoya matrix: the sparse and the dense variants of the matrix. Randić used these matrices as sources for two novel molecular descriptors.

Randić and co-workers[7,8] also introduced a new type of graph-theoretical matrices that they named graphical matrices. Graphical matrices are matrices whose elements are subgraphs of the graph rather than numbers. There are a number of ways how to construct these matrices.[9–11] However, they cannot be used in this non-numerical form, they need to be transferred into a numerical form. This is the advantage of graphical matrices, since they offer many possibilities of numerical realizations. In order to obtain a numerical form of a graphical matrix, one needs to select a molecular descriptor and

* Dedicated to Professor Haruo Hosoya in the happy occasion of his 70th birthday.

** Authors to whom correspondence should be addressed. (E-mail: sonja@irb.hr; vukicevi@pmfst.hr)

replace all the graphical elements (subgraphs of some form) by the corresponding numerical values of the selected descriptor. In this way, the numerical form of the graphical matrix is established and can be used to derive the final descriptor – an invariant of the matrix – to be used in the structure-property-activity modeling.

The numerical realization of graphical matrices can be done by various molecular descriptors. We have already pointed out that Hosoya matrices may be regarded as the numerical realization of a given graphical matrix.[9] Here, we will demonstrate the construction of graphical matrices for trees representing alkanes and their numerical realizations by using the Hosoya index.

The aim of the present report is to show how the procedure to generate the graphical matrices can be carried out by a computer instead of, as it was previously done, by hand, which has so far limited the use of double invariants in the structure-property-activity modeling.

The report is structured as follows. After the introductory words, in the second section, we discuss graphical matrices and the computer program, and in the third section the Hosoya matrices and the related double invariants. We end our report with concluding remarks.

GRAPHICAL MATRICES

We demonstrate the construction of grahical matrices, denoted by $\boldsymbol{G}$, using a branched tree representing the carbon skeleton of 2,2-dimethylhexane. The labeled hydrogen-depleted 2,2-dimethylhexane tree, denoted by T, is shown in Figure 1.

There are several ways to construct graphical matrices, which depend on how one selects the subgraphs constituting the matrix. Here, we present four ways of generating graphical matrices of trees. One way is to define the elements of the graphical matrix $[\boldsymbol{G}]_{ij}$ as the subgraphs obtained after consecutive removal of the edges connecting vertices i and j from tree T. We call this matrix the edge-graphical matrix and denote it by ${}^{e}\boldsymbol{G}$, where e stands for the edge. The ${}^{e}\boldsymbol{G}$ matrices are sparse matrices because they contain only a few non-vanishing elements corresponding to the removed edges. The edge-graphical matrix of T is given in Figure 2. We give only the upper triangle of the matrix since it is a square, $V \times V$, symmetrical matrix, where V is the total number of vertices in T. Similarly, all other graphical matrices in this report will be given in the same way.

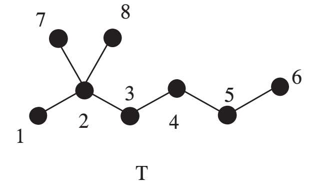

Figure 1. A labeled branched tree T depicting the carbon skeleton of 2,2-dimethylhexane.

If, however, we generate a graphical matrix by consecutive removal of the paths joining vertices i and j instead of edges, the obtained matrix is dense, that is, all its elements but the diagonal elements are non-zero. We call this matrix the path-graphical matrix and denote it by ${}^{p}\boldsymbol{G}$, where p stands for the path. The path-graphical matrix of T is given in Figure 3.

Instead of removing edges, one can remove adjacent vertices i and j, and the incident edges from a tree. The

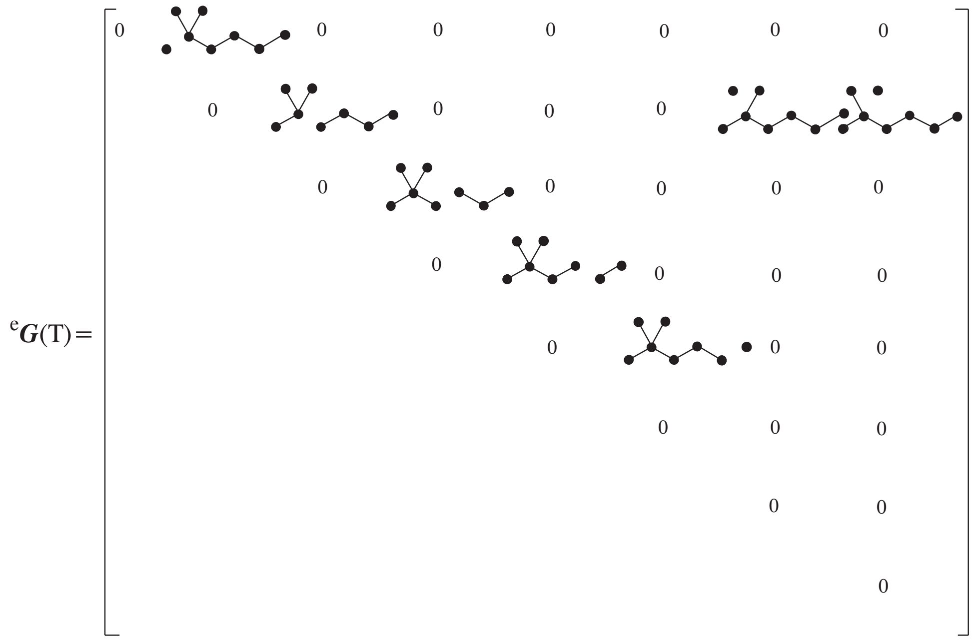

Figure 2. The edge-graphical matrix ${}^{e}\boldsymbol{G}$ of tree T. The removed edges are not shown.

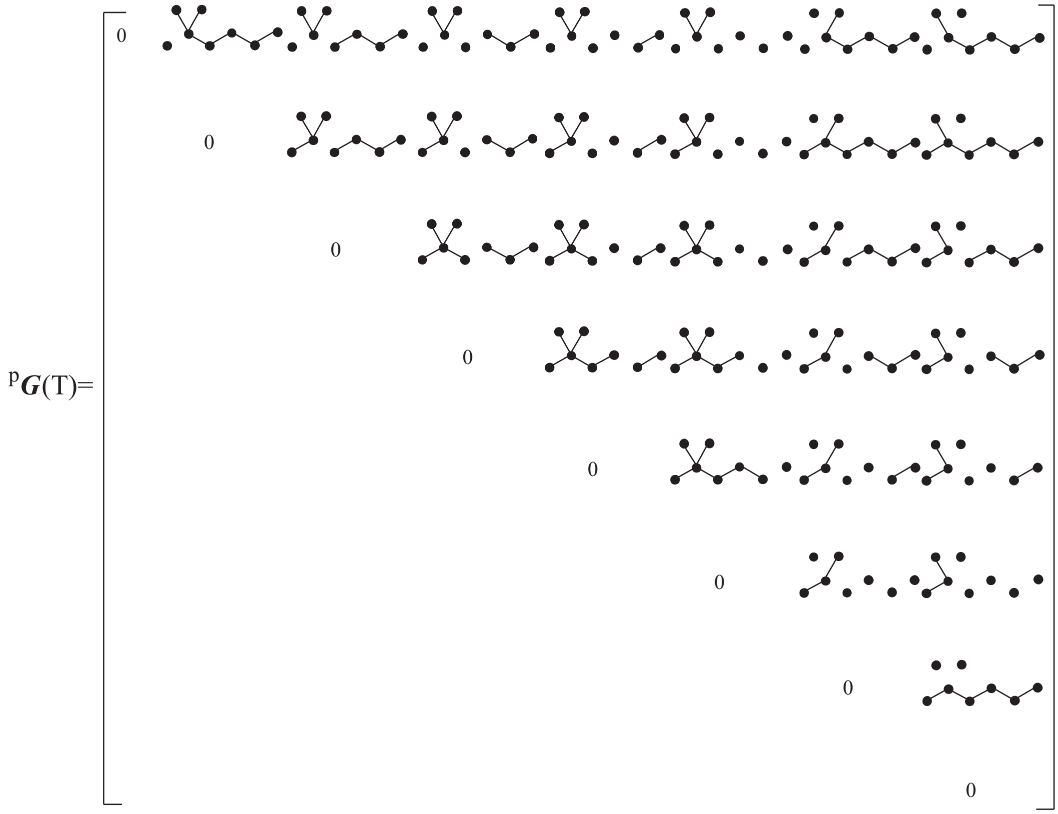

Figure 3. The path-graphical matrix $^{p}\boldsymbol{G}$ of tree T. The removed paths are not shown.

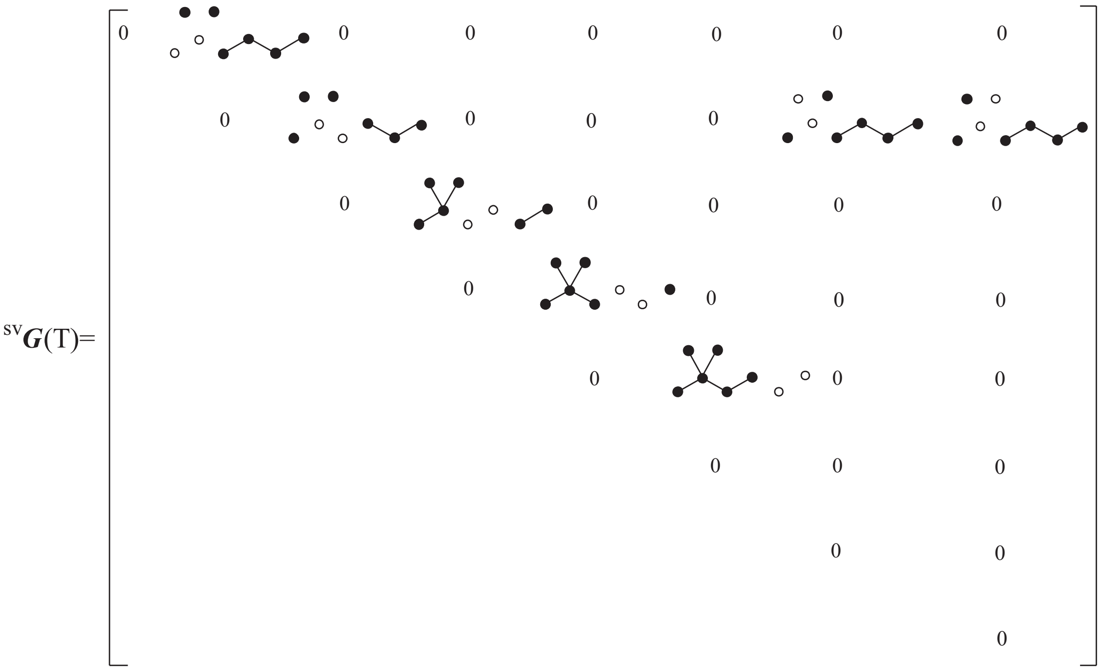

Figure 4. The sparse vertex-graphical matrix $^{sv}\boldsymbol{G}$ of tree T. The removed vertices are shown as empty circles. The removed edges are not shown.

obtained graphical matrix is necessarily sparse because it contains only a few non-vanishing elements corresponding to the removed adjacent vertices. We call this matrix the sparse vertex-graphical matrix and denote it by $^{sv}\boldsymbol{G}$, where s stands for sparse. The sparse vertex-graphical matrix of T is given in Figure 4.

Finally, we can remove pairs of vertices i and j at increasing distances and incident edges. The obtained gra-

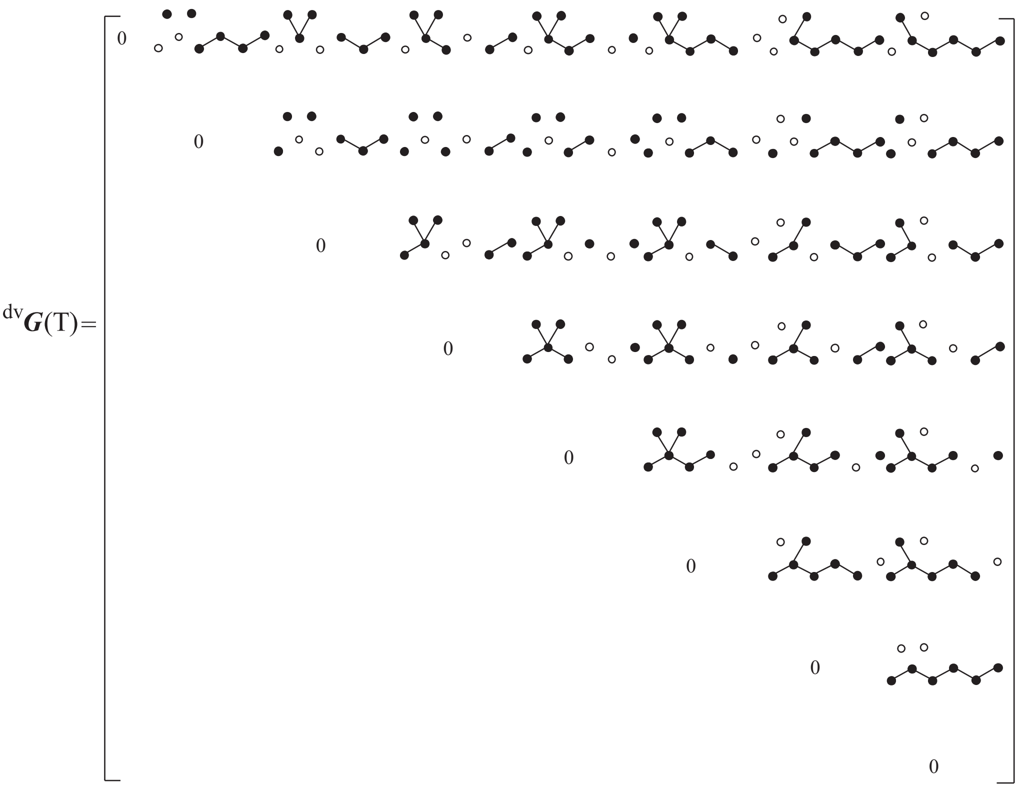

Figure 5. The dense vertex-graphical matrix ${}^{dv}\boldsymbol{G}$ of tree T. The removed vertices are shown as empty circles.

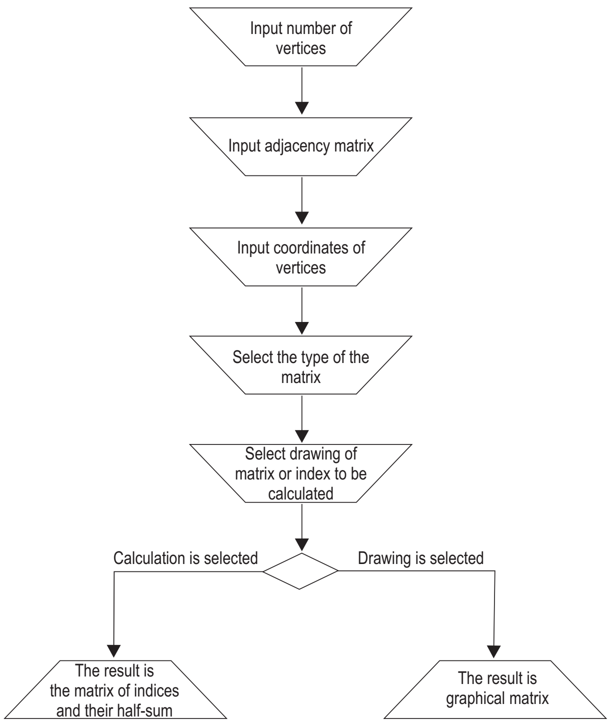

Figure 6. The block-diagram of the program *Dubrovnik Graphical Matrices Calculator.*

phical matrix is dense. We call this matrix the dense vertex-graphical matrix and denote it by ${}^{dv}\boldsymbol{G}$, where d stands for dense. The dense vertex-graphical matrix of T is given in Figure 5.

It should be pointed out that only the edge-graphical matrix and the sparse vertex-graphical matrix can be straightforwardly used for structures containing cycles.

There was a problem with graphical matrices – they were generated by hand. This is perhaps the reason why their use so far was rather limited. However, we developed a computer program that allows the construction of graphical matrices and computation of selected double invariants for trees representing carbon skeletons of acyclic hydrocarbons. The block-diagram of this program is presented in Figure 6.

HOSOYA MATRICES

To get the Hosoya matrices, it is necessary to replace subtrees in graphical matrices by Hosoya indices.[1] The numbers that replace the subgraphs in graphical matrices can be obtained either by summing up or by multiplying their Hosoya indices. Here, we multiplied the Hosoya indices while generating the Hosoya matrices. The Hosoya indices of subtrees are taken from our book on computational chemical graph theory.[12] However, they can be computed by the following formula:

$$Z = \sum_{k=0}^{[V/2]} p(\mathrm{G}; k)$$

where Z denotes the Hosoya index, G stands for a simple connected graph, V is the number of vertices in G and

p(G;k) is the number of independent sets of k edges of G. A set S of k edges is independent if no two edges of set S are adjacent in G. The Gaussian brackets [] above the summation denote the integer part of $V/2$. The empty set and all singleton sets are independent, hence p(G;0) = 1 and p(G;1) = the number of edges in G.

The edge-graphical matrix gives a rise to a numerical matrix that we call the edge-Hosoya matrix and denote it by $^{e}\mathbf{Z}$. As already stated, Randić[6] called this matrix the sparse Hosoya matrix. The $^{e}\mathbf{Z}$ matrix of T is given below.

$$^{e}\mathbf{Z}(\mathrm{T}) = \begin{bmatrix} 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 20 & 0 & 0 & 0 & 18 & 18 \\ & & 0 & 15 & 0 & 0 & 0 & 0 \\ & & & 0 & 18 & 0 & 0 & 0 \\ & & & & 0 & 14 & 0 & 0 \\ & & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 \\ & & & & & & & 0 \end{bmatrix}$$

If we sum up the elements in the above matrix-triangle in the way how Hosoya calculated the Wiener index from the distance matrix,[1] we obtain an index that we should like to call the edge-Hosoya-Wiener index and denote it by ^{e}ZW. Since two topological indices (two graph invariants) are used to generate ^{e}ZW, indices of this type are called by Randić *et al.*[6] double invariants. The ^{e}ZW index of T is 121.

Similarly, the path-Hosoya matrix, denoted by $^{p}\mathbf{Z}$, represents the numerical realization of the corresponding path-graphical matrix. Randić[6] called this matrix complete Hosoya matrix. The path-Hosoya matrix of T is given below.

$$^{p}\mathbf{Z}(\mathrm{T}) = \begin{bmatrix} 0 & 18 & 15 & 9 & 6 & 3 & 13 & 13 \\ & 0 & 20 & 12 & 8 & 4 & 18 & 18 \\ & & 0 & 15 & 10 & 5 & 15 & 15 \\ & & & 0 & 18 & 9 & 9 & 9 \\ & & & & 0 & 14 & 6 & 6 \\ & & & & & 0 & 3 & 3 \\ & & & & & & 0 & 13 \\ & & & & & & & 0 \end{bmatrix}$$

If we sum up the elements in the above matrix-triangle, we obtain an index that we like to call the path-Hosoya-Wiener index and denote it by ^{p}ZW. The ^{p}ZW index of T is 293.

The sparse vertex-graphical matrix gives a rise to a numerical matrix that we call the sparse vertex-Hosoya matrix and denote it by $^{sv}\mathbf{Z}$. The $^{sv}\mathbf{Z}$ matrix of T is given below.

$$^{sv}\mathbf{Z}(\mathrm{T}) = \begin{bmatrix} 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 3 & 0 & 0 & 0 & 5 & 5 \\ & & 0 & 8 & 0 & 0 & 0 & 0 \\ & & & 0 & 5 & 0 & 0 & 0 \\ & & & & 0 & 9 & 0 & 0 \\ & & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 \\ & & & & & & & 0 \end{bmatrix}$$

If we sum up the elements in the above matrix-triangle, we obtain an index that we like to call the sparse vertex-Hosoya-Wiener index and denote it by ^{sv}ZW. The ^{sv}ZW index of T is 40.

Similarly, the dense vertex-Hosoya matrix, denoted by $^{dv}\mathbf{Z}$, represents the numerical realization of the corresponding dense vertex-graphical matrix. The dense vertex-Hosoya matrix of T is given below.

$$^{dv}\mathbf{Z}(\mathrm{T}) = \begin{bmatrix} 0 & 5 & 9 & 6 & 6 & 11 & 13 & 13 \\ & 0 & 3 & 2 & 2 & 3 & 5 & 5 \\ & & 0 & 8 & 4 & 8 & 9 & 9 \\ & & & 0 & 5 & 5 & 8 & 8 \\ & & & & 0 & 9 & 7 & 7 \\ & & & & & 0 & 11 & 11 \\ & & & & & & 0 & 13 \\ & & & & & & & 0 \end{bmatrix}$$

If we sum up the elements in the above matrix-triangle, we obtain an index that we like to call the dense vertex-Hosoya-Wiener index and denote it by ^{dv}ZW. The ^{dv}ZW index of T is 205.

It is of interest to note that the Hosoya-Wiener index of two of the four graphical matrices $^{e}\boldsymbol{G}$, $^{p}\boldsymbol{G}$, $^{sv}\boldsymbol{G}$ and $^{dv}\boldsymbol{G}$, namely of $^{e}\boldsymbol{G}$ and $^{sv}\boldsymbol{G}$, can be calculated without producing graphical matrices by the procedure described as follows. The edge-graphical matrix is defined by $[^{e}\boldsymbol{G}]_{ij} = \begin{cases} G-i-j, & ij \in E(\mathrm{G}) \\ 0, & ij \notin E(\mathrm{G}) \end{cases}$, where E(G) is the set of edges of G. The vertex-graphical matrix is defined by $^{sv}\boldsymbol{G}_{ij}$ = G – i – j.

Let us calculate ^{e}ZW. Denote by a_k the number of partial matchings in G with k edges. Note that each partial matching in some $^{e}\boldsymbol{G}_{ij}$ is also a partial matching in G. Also, note every partial matching with k edges in G is calculated in all subgraphs (*i.e.,* in all $^{e}\mathbf{G}_{ij}$, $i > j$) that do not contain any of its double bonds, hence:

$$^{e}ZW(\mathrm{G}) = \sum_{k=0}^{\lfloor n/2 \rfloor} (e(\mathrm{G}) - k) \cdot a_k = e \cdot h_{\mathrm{G}}(1) - h_{\mathrm{G}}'(1) \,,$$

where e(G) is the number of edges of G, h_{G} is the Hosoya polynomial of G and h'_{G} is the first derivative of h_{G}.

Let us calculate ${}^{sv}ZW$. Note that each partial matching in some ${}^{sv}\boldsymbol{G}_{ij}$ is also a partial matching in G. Also, note every partial matching with k edges in G is calculated in all subgraphs (*i.e.,* in all ${}^{sv}\boldsymbol{G}_{ij}$, $i > j$) that do not contain any of vertices incident to its double bonds. Note that k double bonds are incident with $2k$ vertices; hence, they are not counted in $\binom{2k}{2}$ ${}^{sv}\boldsymbol{G}_{ij}$-graphs such that $i > j$.

Therefore:

$$
{}^{sv}ZW(\mathrm{G}) = \sum_{k=0}^{\lfloor n/2 \rfloor} \left(\binom{v(\mathrm{G})}{2} - \binom{2k}{2} \right) \cdot a_k =
$$

$$
\sum_{k=0}^{\lfloor n/2 \rfloor} \left(\binom{v(\mathrm{G})}{2} - 2k \cdot (k-1) + k \right) \cdot a_k =
$$

$$
\binom{v(\mathrm{G})}{2} \cdot h_{\mathrm{G}}(1) - 2 \cdot h_{\mathrm{G}}''(1) + h_{\mathrm{G}}'(1),
$$

where h_{G}'' is the second derivative of h_{G}.

CONCLUDING REMARKS

In this report, we used the Hosoya index, one of the first descriptors that was proposed to be used in the structure-property-activity studies. The Hosoya index was deliberately chosen for illustrating our computational approach in order to show our appreciation of Haruo Hosoya for his ground-breaking work in mathematical chemistry. Since he has also introduced an easy way to compute the Wiener index, we selected this index as a matrix invariant and thus produced the double invariant that we call the Hosoya-Wiener index, linking in this way the names of the two great mathematical chemistry pioneers Wiener and Hosoya.

We have also demonstrated that the graphical matrices of alkanes can be efficiently generated by using the computer approach and how the numerical form of these matrices can be obtained by using the molecular descriptor of choice. Our computer program is currently limited to acyclic structures.

We have not applied the Hosoya-Wiener index because it has been already demonstrated by Randić and his co-workers[6,7] that the invariants have a future in the modeling of molecular properties and activities. The program for computing double invariants is freely available at the address www.pmfst.hr/~vukicevi.

Acknowledgement. – This work was supported in part by the Ministry of Science, Education and Sports of the Republic of Croatia.

REFERENCES

1. H. Hosoya, *Bull. Chem. Soc. Japan* **44** (1971) 2332–2339.
2. N. Trinajstić, *Chemical Graph Theory*, 2nd revised edition, CRC, Boca Raton, FL, 1992.
3. J. Devillers and A. T. Balaban (Eds.), *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon & Breach, Amsterdam, 1999.
4. R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2000.
5. *e.g.,* A. Miličević and N. Raos, *Polyhedron* **25** (2006) 2800–2808.
6. M. Randić, *Croat. Chem. Acta* **67** (1994) 415–429.
7. M. Randić, D. Plavšić, and M Razinger, *MATCH Commun. Math. Comput. Chem*. **35** (1997) 243–259.
8. M. Randić, N. Basak, and D. Plavšić, *Croat. Chem. Acta* **77** (2004) 251–257.
9. S. Nikolić, A. Miličević, and N. Trinajstić, *Croat. Chem. Acta* **78** (2005) 241–250.
10. A. Miličević and N. Trinajstić, in: A. Hinchliffe (Ed.), *Chemical Modelling: Applications and Theory*, Vol 4, The Royal Society of Chemistry, Cambridge, 2006, pp. 408–472.
11. D. Janežič, A. Miličević, S. Nikolić, and N. Trinajstić, *Graph-Theoretical Matrices in Chemistry*, The University of Kragujevac, Kragujevac, 2007.
12. N. Trinajstić, S. Nikolić, J. V. Knop, W. R. Müller, and K. Syzmanski, *Computational Chemical Graph Theory – Characterization, Enumeration and Generation of Chemical Structures by Computer Methods*, Horwood, New York, 1991, Chapter 9.
13. M. Randić, *Croat. Chem. Acta* **66** (1993) 289–312.

SAŽETAK

Hosoyine matrice kao numerička realizacija grafičkih matrica i izvedeni strukturni deskriptori

Dušanka Janežič, Bono Lučić, Ante Miličević, Sonja Nikolić, Nenad Trinajstić i Damir Vukičević

Grafičke matrice su upotrijebljene za generiranje Hosoyinih matrica, koje su zatim upotrijebljene za računanje Hosoya-Wienerovih indeksa. Prikazan je računalni program za generiranje grafičkih matrica acikličkih struktura, njihovih numeričkih matrica i dvostrukih strukturnih invarijanata.