Indian Journal of Chemistry
Vol. 37A, July 1998, pp. 574 - 585

New topological matrices and their polynomials

Ovidiu Ivanciuc*§, Mircea V. Diudea** & Padmakar V. Khadikar‡
§Department of Organic Chemistry, Polytechnic University of Bucharest, Office 12,
P.O.Box 243, 78100 Bucharest, Romania
Revised 12 June 1997; revised 10 November 1997

New polynomials and molecular graph spectra are derived on the basis of recently proposed matrices, D_A, D_P and RD_p, and of the previously defined RD_e (reciprocal distance) matrix. The new molecular graph descriptors are exemplified on a collection of path-graphs L_2-L_8 and cyclic-graphs C_3-C_8. The polynomials for the last two matrices (both based on a reciprocal relationship) are the first defined polynomials with non-integer coefficients. The sum of their absolute values, denoted as SumCh(RD_e) and SumCh(RD_p) can be viewed as global descriptors (*i.e.* topological indices), by analogy with the Hosoya's number *Z*. Values of the two indices as well as the minimum, MinSp(M), and maximum, MaxSp(M), eigenvalues are listed for nonane isomers. The new structural descriptors have been tested for correlation with some physico-chemical properties and are found to give good results in modeling the properties of alkanes.

A molecular graph can be represented by either a matrix, a polynomial or a single number (i.e. topological index). Each of them is aimed to be unique for a given graph, G. However, in quantitative structure-property relationships (QSPR) studies this requirement is not so drastic since there are physico-chemical properties which are quite interrelated; on the other hand, there are different chemical structures which show almost identical values of a given property. All of the above mentioned representations of molecular graphs are considered here; the resulting indices are tested for correlation with seven of the most important physico-chemical properties of nonanes.

The Wiener index[1], *W*, has been introduces fifty years ago (by Harold Wiener) as a descriptor for explaining the boiling points of paraffins on the basis of their structures. It was looked upon as a measure of compactness of molecules but only recently[2] its relation with the molecular van der Waals areas was demonstrated. Many other properties (heats of formation, molar refraction, critical pressure, surface tension) as well as biological properties have found correlation with this structural descriptor[3-5].

The Wiener index, and its extension, the hyper-Wiener index[6], *WW*, can be defined as *edge/path (e/p) contributions* to a global number, *I*

$$I = I(G) = \sum_{e/p} I_{e/p} = \sum_{e/p} N_{i,e/p} N_{j,e/p} \quad ...(1)$$

N_i, N_j being the number of vertices lying on the two sides of edge/path, e/p, (having the vertices i and j as endpoints) and the summation runs over all edges/paths in the graph.

The edge/path contributions $I_{e/p}$ are just the entries in the Wiener matrices[7,8], W_e and W_p. Thus, *I* can be calculated by the relation:

$$I = (1/2)\sum_i \sum_j [W_{e/p}]_{ij} \quad ...(2)$$

The above relations hold only in acyclic structures. There are also relations which extend the "bond contribution" definition for cycle-containing graphs[9-11].

W can be obtained, for any graph, as the half-sum of entries in the distance matrix[12], D_e,

$$W = (1/2) \sum_i \sum_j [D_e]_{ij} \quad ...(3)$$

The Wiener index represents the number of all edges included in all shortest paths in G. For the hyper- Wiener index, an equivalent definition was

** Department of Chemistry, "Babes-Bolyai" University, 3400 Cluj, Romania; e-mail: diudea@chem.ubbcluj.ro
‡ Devy Ahilya University, Indore (M.P.), Indore 453007, India
§ e-mail: o_ivanciuc@chim.upb.ro

given by Diudea[13], as presented in Eq. (7).

New matrix invariants

Diudea[13] has recently defined two new combinatorial matrices, namely:

$$[D_\Delta]_{ij} = \left[\frac{[De]_{ij}}{2}\right] \quad \ldots (4)$$

and

$$[D_p]_{ij} = \left[\frac{[De]_{ij} + 1}{2}\right] \quad \ldots (5)$$

A matrix which collects the reciprocal elements of the D_p matrix, denoted RD_p, was proposed by analogy with the early RD_e matrix, of reciprocal distances[14,15]

$$[RD_p]_{ij} = 1 / [D_p]_{ij} \quad \ldots (6)$$

It is easy to reconstruct the graph from the matrices D_Δ and D_p: the adjacency is given by nondiagonal zeros and unity entries in rows, respectively. We present below the D_Δ, D_p, RD_e and RD_p matrices for the molecular graph of 3-methylhexane (see Fig. 1).

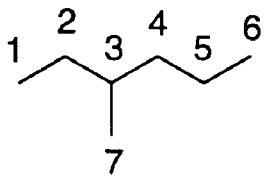

Fig. 1— The molecular graph of 3-methylhexane

The matrix D_p allows another definition[13] of the hyper-Wiener index, with the following equation:

$$WW = (1/2) \Sigma_i \Sigma_j [D_p]_{ij} \quad \ldots (7)$$

representing all paths included in all shortest paths in G.

The expansion of the right member in Eq. (7) enables the decomposition of WW into two terms

$$WW = \sum_{i<j} [De]_{ij} + \sum_{i<j} \left(\frac{[De]_{ij}}{2}\right) \quad \ldots (8)$$

where the former term is just the Wiener number W The latter term is the "non-Wiener" part of the hyper-Wiener number, which is denoted by D_Δ (after the D_Δ matrix)

D_Δ

0	0	1	3	6	10	3
0	0	0	1	3	6	1
1	0	0	0	1	3	0
3	1	0	0	0	1	1
6	3	1	0	0	0	3
10	6	3	1	0	0	6
3	1	0	1	3	6	0

D_p

0	1	3	6	10	15	6
1	0	1	3	6	10	3
3	1	0	1	3	6	1
6	3	1	0	1	3	3
10	6	3	1	0	1	6
15	10	6	3	1	0	10
6	3	1	3	6	10	0

RD_e

0	1	1/2	1/3	1/4	1/5	1/3
1	0	1	1/2	1/3	1/4	1/2
1/2	1	0	1	1/2	1/3	1
1/3	1/2	1	0	1	1/2	1/2
1/4	1/3	1/2	1	0	1	1/3
1/5	1/4	1/3	1/2	1	0	1/4
1/3	1/2	1	1/2	1/3	1/4	0

RD_p

0	1	1/3	1/6	1/10	1/15	1/6
1	0	1	1/3	1/6	1/10	1/3
1/3	1	0	1	1/3	1/6	1
1/6	1/3	1	0	1	1/3	1/3
1/10	1/6	1/3	1	0	1	1/6
1/15	1/10	1/6	1/3	1	0	1/10
1/6	1/3	1	1/3	1/6	1/10	0

$$D_\Delta = \sum_{i<j} [D_\Delta]_{ij} = \sum_{i<j} \binom{[De]_{ij}}{2} \quad \ldots (9)$$

This molecular descriptor represents the number of all paths (larger than unity) included into all shortest paths (larger than unity) in graph. Thus, *WW* can be written as

$$WW = W + D_\Delta \quad \ldots (10)$$

and the corresponding matrix relation

$$D_p = D_e + D_\Delta \quad \ldots (11)$$

For other considerations about these indices the reader can consult refs. 13 and 16.

Polynomials of combinatorial and reciprocal distance matrices

Combinatorial matrices

The polynomial $Ch(D_\Delta,G,x)$ is the characteristic polynomial of the D_Δ matrix:

$$Ch(D_\Delta,G,x) = \det (xI - D_\Delta) \quad \ldots (12)$$

The polynomials $Ch(D_\Delta,G,x)$ for the collection of path-graphs L_2-L_8 and cyclic-graphs C_3-C_8 are presented below:

$Ch(D_\Delta,L_2) = x^2$

$Ch(D_\Delta,L_3) = x^3 - x$

$Ch(D_\Delta,L_4) = x^4 - 11x^2 + 1$

$Ch(D_\Delta,L_5) = x^5 - 57x^3 - 12x^2 + 29x - 6$

$Ch(D_\Delta,L_6) = x^6 - 203x^4 - 144x^3 + 357x^2 - 156x + 20$

$Ch(D_\Delta,L_7) = x^7 - 574x^5 - 906x^4 + 2675x^3 - 1866x^2 + 515x - 50$

$Ch(D_\Delta,L_8) = x^8 - 1386x^6 - 4020x^5 + 14443x^4 - 14328x^3 + 6365x^2 - 1324x + 105$

$Ch(D_\Delta,C_3) = x^3$

$Ch(D_\Delta,C_4) = x^4 - 2x^2 + 1$

$Ch(D_\Delta,C_5) = x^5 - 5x^3 + 5x - 2$

$Ch(D_\Delta,C_6) = x^6 - 33x^4 - 4x^3 + 252x^2 - 96x - 320$

$Ch(D_\Delta,C_7) = x^7 - 70x^5 - 42x^4 + 833x^3 - 882x^2 - 175x - 8$

$Ch(D_\Delta,C_8) = x^8 - 224x^6 - 240x^5 + 9844x^4 - 9120x^3 - 70992x^2 + 38016x + 145152$

By inspecting the above polynomials we can conclude that the coefficients of $Ch(D_\Delta,G,x)$ polynomials are integer numbers showing large absolute values, of little use as structural descriptors in QSPR studies. The spectrum of the matrix D_Δ, $Sp(D_\Delta,G)$, represents the eigenvalues of the matrix $D_\Delta(G)$ (or the solutions of the $Ch(D_\Delta,G,x)$ polynomial). For the above collection of graphs, the spectra are presented in Appendix 1. Two new molecular descriptors are defined on the basis of the $Sp(D_\Delta,G)$ spectrum, namely the maximum- and minimum-values, $MaxSp(D_\Delta,G)$ and $MinSp(D_\Delta,G)$, respectively. Their values for the set of nonanes are listed in Table 1.

The $Ch(D_p,G,x)$ polynomial is the characteristic polynomial of the D_p matrix:

$$Ch(D_p,G,x) = \det (xI - D_p) \quad \ldots (13)$$

The polynomials $Ch(D_p,G,x)$ for the above collection of graphs are given below:

$Ch(D_p,L_2) = x^2 - 1$

$Ch(D_p,L_3) = x^3 - 11x - 6$

$Ch(D_p,L_4) = x^4 - 57x^2 - 84x - 20$

$Ch(D_p,L_5) = x^5 - 203x^3 - 582x^2 - 325x - 50$

$Ch(D_p,L_6) = x^6 - 574x^4 - 2760x^3 - 2655x^2 - 916x - 105$

$Ch(D_p,L_7) = x^7 - 1386x^5 - 10230x^4 - 14827x^3 - 8502x^2 - 2135x - 196$

$Ch(D_p,L_8) = x^8 - 2982x^6 - 31812x^5 - 64295x^4 - 53944x^3 - 22085x^2 - 4376x - 336$

$Ch(D_p,C_3) = x^3 - 3x - 2$

$Ch(D_p,C_4) = x^4 - 22x^2 - 24x + 45$

$Ch(D_p,C_5) = x^5 - 50x^3 - 120x^2 + 65x - 8$

$Ch(D_p,C_6) = x^6 - 168x^4 - 576x^3 + 2448x^2 + 2304x - 7168$

$Ch(D_p,C_7) = x^7 - 322x^5 - 1806x^4 + 4781x^3 + 3066x^2 - 5551x - 3380$

$Ch(D_p,C_8) = x^8 - 768x^6 - 5712x^5 + 49780x^4 + 106400x^3 - 674000x^2 - 560000x + 2400000$

The spectrum $Sp(D_p,G)$ represents the eigenvalues of the matrix $D_p(G)$ (or the roots of the $Ch(D_p,G,x)$ polynomial). For the above collection

Table 1 — Spectral descriptors MaxSp and MinSp of the matrices D_A and D_p

No	Molecular Graph	MaxSp(D_A)	MinSp(D_A)	MaxSp(D_p)	MinSp(D_p)
1	Nonane	56.68815	-51.81249	83.35180	-68.20982
2	2-Methyloctane	48.68512	-44.63692	74.03016	-60.00202
3	3-Methyloctane	44.26889	-39.18169	68.78980	-53.27627
4	4-Methyloctane	42.45764	-35.95544	66.55177	-49.09627
5	3-Ethylheptane	37.36589	-28.75330	60.64732	-40.33791
6	4-Ethylheptane	35.27423	-23.00905	58.12122	-33.00663
7	2,2-Dimethylheptane	37.94818	-34.37718	61.14140	-47.73147
8	2,3-Dimethylheptane	35.61607	-30.43627	58.40137	-42.62526
9	2,4-Dimethylheptane	35.30182	-29.41071	58.07530	-41.35017
10	2,5-Dimethylheptane	36.80535	-32.02677	60.01022	-44.88383
11	2,6-Dimethylheptane	40.91240	-37.34808	64.94966	-51.57304
12	3,3-Dimethylheptane	32.21724	-26.59047	54.14271	-37.77032
13	3,4-Dimethylheptane	31.70309	-24.84470	53.62206	-35.47194
14	3,5-Dimethylheptane	32.98510	-26.94466	55.34213	-38.37089
15	4,4-Dimethylheptane	30.54991	-23.00905	52.02506	-33.00663
16	3-Ethyl-2-methylhexane	29.13677	-18.55416	50.66128	-27.48829
17	4-Ethyl-2-methylhexane	30.52782	-21.72990	52.48670	-31.85817
18	3-Ethyl-3-methylhexane	26.31162	-16.54081	46.93289	-24.80257
19	3-Ethyl-4-methylhexane	27.24071	-17.07561	48.31999	-25.63550
20	2,2,3-Trimethylhexane	26.20312	-20.97830	46.79922	-30.79345
21	2,2,4-Trimethylhexane	27.18616	-22.31566	48.21817	-32.82234
22	2,2,5-Trimethylhexane	30.70623	-27.18539	52.59434	-39.20412
23	2,3,3-Trimethylhexane	24.56413	-18.22328	44.72221	-27.10287
24	2,3,4-Trimethylhexane	25.40360	-19.06643	46.01053	-28.40729
25	2,3,5-Trimethylhexane	28.68509	-23.57654	50.16178	-34.39356
26	2,4,4-Trimethylhexane	25.59678	-19.83292	46.19878	-29.51904
27	3,3,4-Trimethylhexane	22.58951	-15.57933	42.30581	-23.73147
28	3,3-Diethylpentane	22.36227	-7.40512	42.11714	-12.60328
29	2,2-Dimethyl-3-ethylpentane	22.07382	-12.63598	41.82152	-20.03194
30	2,3-Dimethyl-3-ethylpentane	20.66945	-10.40278	39.95890	-16.90465
31	2,4-Dimethyl-3-ethylpentane	23.20912	-12.55744	43.41597	-19.94987
32	2,2,3,3-Tetramethylpentane	17.60267	-11.51963	35.99196	-18.52280
33	2,2,3,4-Tetramethylpentane	20.11501	-14.85031	39.41154	-23.18257
34	2,2,4,4-Tetramethylpentane	21.58928	-17.62414	41.31022	-27.11066
35	2,3,3,4-Tetramethylpentane	18.79551	-12.55744	37.64755	-19.94987

of graphs, the spectra are given in Appendix 1. The corresponding structural descriptors MaxSp(D_p,G) and MinSp(D_p,G) for nonanes are listed in Table 1.

Reciprocal matrices

The polynomial Ch(RD_e,G,x) of a graph G is the characteristic polynomial of the RD_e matrix.[14,15]

$$Ch(RD_e,G,x) = \det (xI - RD_e) \quad ...(14)$$

The polynomial Ch(RD_e,G,x) is the first representant of the class of polynomials with real number coefficients. The coefficients of distance polynomial, Ch(D_e,G,x), show large values and thus are improper for generating topological descriptors, useful in QSPR and QSAR. But, to the contrary, the coefficients of Ch(RD_e,G,x) are suitable for such a purpose. By summing the absolute values of the coefficients a_k of the polynomial Ch(RD_e,G,x) we obtain a descriptor analogous to the Hosoya index[12], Z:

$$SumCh(RD_e) = \sum_k |a_k (Ch(RD_e,G,x))| \quad ... (15)$$

Values of this index for nonanes are given in Table 2. A list of Ch(RD_e,G,x) polynomials for a collection of path- and cyclic-graphs is presented below.

$Ch(RD_e,L_3) = x^3 - 2.25000x - 1$

$Ch(RD_e,L_4) = x^4 - 3.61111x^2 - 2.66667x - 0.15972$

$Ch(RD_e,L_5) = x^5 - 5.03472x^3 - 4.79167x^2 - 0.11092x + 0.61227$

$Ch(RD_e,L_6) = x^6 - 6.49833x^4 - 7.25000x^3 + 0.52844x^2 + 2.88479x + 0.64451$

$Ch(RD_e,L_7) = x^7 - 7.98972x^5 - 9.96204x^4 + 2.07622x^3 + 7.94928x^2 + 2.97405x + 0.06688$

$Ch(RD_e,L_8) = x^8 - 9.50152x^6 - 12.87407x^5 + 4.78665x^4 + 16.87871x^3 + 8.06698x^2 - 0.11232x - 0.46868$

$Ch(RD_e,C_3) = x^3 - 3x - 2$

$Ch(RD_e,C_4) = x^4 - 4.50000x^2 - 4.50000x - 0.93750$

$Ch(RD_e,C_5) = x^5 - 6.25000x^3 - 7.50000x^2 - 2.18750x - 0.18750$

$Ch(RD_e,C_6) = x^6 - 7.83333x^4 - 10.50000x^3 - 1.90046x^2 + 1.65278x - 0.16804$

$Ch(RD_e,C_7) = x^7 - 9.52778x^5 - 14.38889x^4 - 1.73920x^3 + 6.45988x^2 + 1.33936x - 1.00351$

$Ch(RD_e,C_8) = x^8 - 11.13889x^6 - 17.88889x^5 + 0.91387x^4 + 17.06327x^3 + 6.45222x^2 - 4.06499x - 2.13145$

The spectrum Sp(RD_e,G) represents the eigenvalues of the matrix RD_e(G) (or the solutions of Ch(RD_e,G,x)=0 equation). For the above collection of graphs, the spectrum is given in Appendix 1. The values for nonanes of the two new molecular descriptors, MaxSp(RD_e,G) and MinSp(RD_e,G), are given in Table 2.

The polynomial Ch(RD_p,G,x) of a graph G is the characteristic polynomial of the matrix RD_p:

$$Ch(RD_p,G,x) = \det (xI - RD_p) \quad ... (16)$$

This is another polynomial with real number coefficients. Its coefficients are used to define a new molecular descriptor:

$$Sum\ Ch(RD_p) = \sum_k |a_k (Ch(RD_p,G,x))| \quad ... (17)$$

The values of this index for nonanes are given in Table 2. Ch(RD_p,G,x) polynomials for the above collection of graphs are presented below.

$Ch(RD_p,L_3) = x^3 - 2.11111x - 0.66667$

$Ch(RD_p,L_4) = x^4 - 3.25x^2 - 1.55556x + 0.44753$

$Ch(RD_p,L_5) = x^5 - 4.39889x^3 - 2.53333x^2 + 1.79164x + 0.84720$

$Ch(RD_p,L_6) = x^6 - 5.55222x^4 - 3.55259x^3 + 4.24793x^2 + 3.20421x + 0.11116$

$Ch(RD_p,L_7) = x^7 - 6.70782x^5 - 4.59355x^4 + 7.91997x^3 + 7.46495x^2 + 0.01529x - 0.69143$

$Ch(RD_p,L_8) = x^8 - 7.86470x^6 - 5.646856x^5 + 12.86059x^4 + 13.85582x^3 - 1.029420x^2 - 3.72491x - 0.52987$

$Ch(RD_p,C_3) = x^3 - 3x - 2$

$Ch(RD_p,C_4) = x^4 - 4.22222x^2 - 2.66667x - 0.43210$

$Ch(RD_p,C_5) = x^5 - 5.55556x^3 - 4.44444x^2 + 0.80247x - 0.03292$

$Ch(RD_p,C_6) = x^6 - 6.75000x^4 - 5.48148x^3 + 4.44676x^2 + 2.85185x - 1.446181$

$Ch(RD_p,C_7) = x^7 - 7.97222x^5 - 6.87037x^4 + 8.80401x^3 + 8.86523x^2 - 2.62755x - 2.86272$

$Ch(RD_p,C_8) = x^8 - 9.15111x^6 - 7.97037x^5 + 15.21332x^4 + 18.37315x^3 - 4.38180x^2 - 10.54136x - 2.96519$

The spectrum Sp(RD_p,G) represents the eigenvalues of the matrix RD_p(G) (or the roots of Ch(RD_p,G,x) polynomial). For the above collection of graphs, the

Table 2 — Spectral descriptors MaxSp and MinSp and SumCh(M) indices of the matrices RD_e and RD_p

No	MaxSp(RD_e)	MinSp(RD_e)	MaxSp(RD_p)	MinsSp(RD_p)	SumCh(RD_e)	SumCh(RD_p)
1	3.74049	-1.35934	2.86405	-1.50438	89.71628	76.80166
2	3.83088	-1.37996	2.95005	-1.53602	93.88263	72.45017
3	3.88949	-1.39550	3.00846	-1.55719	85.31952	66.37881
4	3.91810	-1.39977	3.03669	-1.56405	78.07320	69.57511
5	3.97103	-1.41138	3.08549	-1.57901	69.78597	65.63738
6	3.99693	-1.41380	3.10919	-1.58341	67.74463	69.21294
7	4.01806	-1.42805	3.14395	-1.61411	96.32293	71.21244
8	4.02525	-1.42656	3.14098	- 1.60122	79.90156	65.35204
9	4.01378	-1.41047	3.12405	- 1.58370	79.57659	69.41800
10	3.97924	-1.40690	3.08773	- 1.57368	90.35920	68.49660
11	3.91897	-1.38971	3.02879	- 1.55314	101.00045	74.19637
12	4.11610	-1.44856	3.23864	- 1.63937	80.88253	62.36811
13	4.08704	-1.43564	3.19876	- 1.61449	72.26710	60.10992
14	4.04390	-1.41791	3.15142	- 1.59188	78.56807	63.31859
15	4.14666	-1.44904	3.26671	- 1.64242	73.40243	63.28952
16	4.11002	-1.43757	3.21809	- 1.61754	68.47534	62.61426
17	4.06796	-1.42181	3.17265	-1.59690	73.47414	63.59971
18	4.21215	-1.46425	3.32759	-1.66013	67.15264	54.60532
19	4.14491	-1.44343	3.25077	-1.62582	65.11322	58.30437
20	4.22683	-1.46809	3.34175	-1.66547	83.51070	61.19526
21	4.17365	-1.44043	3.28144	-1.63412	93.75835	66.38074
22	4.10574	-1.43649	3.21592	-1.62286	107.81940	75.70775
23	4.26454	-1.47565	3.37855	-1.67556	76.32096	58.05071
24	4.19587	-1.45225	3.29952	-1.63949	76.90603	58.24704
25	4.11950	-1.43469	3.22289	-1.61458	87.20227	66.55531
26	4.21307	-1.45241	3.32182	-1.64899	85.75659	63.07368
27	4.29734	-1.47908	3.40732	-1.68051	73.21446	52.94883
28	4.28067	-1.47715	3.39074	-1.67632	61.26824	53.04559
29	4.28732	-1.47211	3.39470	-1.67303	72.46244	56.75668
30	4.33247	-1.48654	3.44033	-1.68978	67.05941	49.70956
31	4.22361	-1.45385	3.32419	-1.64346	69.86495	58.96288
32	4.44311	-1.50871	3.55113	-1.72322	77.44882	50.00178
33	4.33528	-1.47691	3.43782	-1.68096	86.60200	58.77651
34	4.30235	-1.45211	3.40460	-1.66039	108.76769	78.82451
35	4.38245	-1.49436	3.48702	-1.70125	76.27681	52.44864

spectra are presented in Appendix 1. Using the above spectrum, we define two new molecular descriptors, MaxSp(RD_e,G) and MinSp(RD_e,G); their values for nonanes are given in Table 2.

Structure-property relationships with the new descriptors

The spectral data obtained from the four matrices were tested for correlation with seven of the most useful physico-chemical properties of nonanes: boiling point (BP), heat of vaporization (HV), molar volume (MV), critical pressure (CP), surface tension (ST) and critical temperature (CT). The experimental values were taken from ref. 17 and are listed in Table 3. Their intercorrelation matrix for the values of the seven properties is presented in Table 4.

A correlation coefficient of over 0.80 suggests a common causal determination, although not quite transparent. Best correlated are the properties MR and MV, a result already known; as a consequence, the two properties will be best described basically by the same descriptors (see below).

In *single variable regression*, none of the spectral descriptors gives good correlations. The maximum correlation are as follows: BP, 0.580 with SumCh(RD_e); HV, 0.695 with MaxSp(RD_e); MR, 0.919 with SumCh(RD_p); MV, 0.913 with SumCh(RD_p); CP, 0.840 with SumCh(RD_p); ST, 0.843 with SumCh(RD_e) and CT, 0.782 with SumCh(RD_p). The intercorrelation coefficients between the ten spectral descriptors are presented in the Table 5.

In *two variable regression* only the heat of vaporization could be adequately explained: $r = 0.969$; $s = 0.408$ with MaxSp(D_Δ) and SumCh(RD_e).

Because the spectral descriptors are highly intercorrelated, except the two SumCh(M), (see Table 5), it is not possible to investigate QSPR models with a higher number of parameters using only spectral descriptors. In the following models we will use the spectral descriptors together with the walk numbers ${}^eW_A$ as non-spectral parameters. They are the classical molecular walk counts[18] of length e defined as:[3]

$$ {}^eW_A = (1/2)\sum_i {}^eW_{A,i} = (1/2)\sum_i \sum_j [A^e]_{ij} \qquad \ldots (18) $$

where ${}^eW_{A,i}$ is the row sum in the adjacency matrix raised to the power e. For the more general descriptors ${}^eW_M$, and procedures of evaluation that avoid the matrix power calculation, the reader can consult ref. 19. Table 6 lists the ${}^eW_A$ values for nonanes.

In *three variable regression*, the best correlations obtained are: BP, $r = 0.967$; $s = 1.600$, with MaxSp(D_Δ), $1/({}^2W_A)$ and ${}^7W_A$; HV, $r = 0.974$; $s = 0.382$, with MaxSp(D_Δ), SumCh(RD_e) and SumCh(RD_p); MR, $r = 0.965$; $s = 0.067$, with MaxSp(RD_p), $1/({}^2W_A)$ and ${}^3W_A$; MV, $r = 0.986$; $s = 0.568$, with MaxSp(RD_p), $1/({}^2W_A)$ and ${}^7W_A$; CP, $r = 0.912$; $s = 0.590$, with MaxSp(RD_p), $1/({}^2W_A)$ and ${}^3W_A$; ST, $r = 0.976$; $s = 0.226$, with MaxSp(RD_p), $1/({}^2W_A)$ and ${}^7W_A$; and CT, $r = 0.916$; $s = 4.632$, with *MaxSp*(RD_p), $1/({}^2W_A)$ and ${}^7W_A$.

It is obvious that the walk numbers, particularly ${}^2W_A$ (as $1/({}^2W_A)$) and ${}^7W_A$ account for the majority of correlation: BP, 0.933; HV, 0.976; MR, 0.936; MV, 0.976; CP, 0.906; ST, 0.973 and CT, 0.916.

Among the investigated matrices, the most useful is RD_p, both as MaxSp(RD_p) and SumCh(RD_p) molecular descriptors. The two types of indices were successfully tested in discriminating a pair of graphs (2;3) built up so that they show identical pairwise distance degree sequence, DDS (see Fig. 2); also, for this pair of graphs the descriptors based on distances are degenerated.[20] The indices SumCh(M), M = RD_e and RD_p discriminate between the two nonisomorphic graphs. We present below the structural descriptors of the two graphs 2 and 3: DDS, Ch(RD_e), SumCh(RD_e), Ch(RD_p), SumCh(RD_p).

2

3

Fig. 2 — Pairs of graphs with pairwise identical distance degree sequence, DDS

Table 3 — Physico-chemical properties of nonanes.

No	BP	HV	MR	MV	PC	ST	TC
1	150.798	46.44	43.8423	178.713	22.74	22.92	322
2	143.26	44.65	43.8795	179.773	23.6	21.88	315
3	144.18	44.75	43.7296	177.952	23.7	22.34	318
4	142.48	44.75	43.7687	178.15	23.06	22.34	318.3
5	143.00	44.81	43.642	176.41	23.98	22.81	318.
6	141.20	44.81	43.4907	175.685	23.98	22.81	318.3
7	132.69	42.28	43.9138	180.507	22.8	20.8	302
8	140.50	43.79	43.6369	176.653	23.79	22.34	315
9	133.50	42.87	43.7393	179.12	22.7	21.3	306
10	136.00	43.87	43.8484	179.371	22.7	21.3	307.8
11	135.21	42.82	43.9258	180.914	23.7	20.83	306
12	137.30	42.66	43.687	176.897	24.19	22.01	314
13	140.60	43.84	43.5473	175.349	24.77	22.8	322.7
14	136.00	42.98	43.6378	177.386	23.59	21.77	312.3
15	135.20	42.66	43.6022	176.897	24.18	22.01	317.8
16	138.00	43.84	43.655	175.445	24.77	22.8	322.7
17	133.80	42.98	43.6472	177.386	25.56	21.77	330.3
18	140.60	43.04	43.268	173.077	25.66	23.22	327.2
19	140.40	43.95	43.3746	172.844	23.59	23.27	312.3
20	133.60	41.91	43.6226	175.878	25.07	21.86	318.1
21	126.54	40.57	43.7638	179.22	23.39	20.51	301
22	124.084	40.17	43.9356	181.346	22.41	20.04	296.6
23	137.68	42.23	43.4347	173.78	25.56	22.41	326.1
24	139.00	42.93	43.4917	173.498	25.46	22.8	324.2
25	131.34	41.42	43.6474	177.656	23.49	21.27	309.4
26	130.648	40.84	43.6598	177.187	23.79	21.17	309.1
27	140.46	42.28	43.3407	172.055	26.45	23.27	330.6
28	146.168	43.36	43.1134	170.185	26.94	23.75	342.8
29	133.83	42.02	43.4571	174.537	25.96	22.38	322.6
30	142.00	42.55	42.9542	170.093	26.94	23.87	338.6
31	136.73	42.93	43.4037	173.804	25.46	22.8	324.2
32	140.274	41	43.2147	169.495	27.04	23.38	334.5
33	133.016	41	43.4359	173.557	25.66	21.98	319.6
34	122.284	38.1	43.8747	178.256	24.58	20.37	301.6
35	141.551	41.75	43.2016	169.928	26.85	23.31	334.5

Table 4 — Intercorrelation coefficients for the values of the physico-chemical properties of nonanes

	HV	MR	MV	CP	ST	CT
BP	0.8267	0.3931	0.4130	0.2324	0.8142	0.6545
HV	1	0.0376	0.0570	0.2334	0.4826	0.2611
MR		1	0.9643	0.8520	0.8102	0.8301
MV			1	0.8880	0.8341	0.8461
CP				1	0.6535	0.8475
ST					1	0.8832

Table 5 — Intercorrelation coefficients for the structural descriptors defined on the basis of molecular spectra

	MinSp (RD_e)	MaxSp (D_Δ)	MinSp (D_Δ)	MaxSp (D_p)	MinSp (D_p)	MaxSp (RD_p)	MinSp (RD_p)	SumCh (RD_e)	SumCh (RD_p)
MaxSp(RD_e)	- 0.980	- 0.977	0.931	- 0.981	0.928	0.999	- 0.986	- 0.269	- 0.729
MinSp(RD_e)	1.000	0.945	- 0.911	0.950	- 0.911	- 0.985	0.996	0.351	0.799
MaxSp(D_Δ)		1.000	- 0.968	1.000	- 0.962	- 0.972	0.949	0.309	0.708
MinSp(D_Δ)			1.000	- 0.967	1.000	0.925	- 0.904	- 0.500	- 0.766
MaxSp(D_p)				1.000	- 0.962	- 0.976	0.954	0.310	0.713
MinSp(D_p)					1.000	0.923	- 0.903	- 0.518	- 0.777
MaxSp(RD_p)						1.000	- 0.991	- 0.271	- 0.733
MinSp(RD_p)							1.000	0.290	0.753
SumCh(RD_e)								1.000	0.737

DDS(2;3) = 43, 45, 49, 49, 55, 61, 63, 63, 63, 65, 65, 69, 77, 77, 79, 79, 79, 85.

$Ch(RD_e, 2) = x^{18} - 29.64728\,x^{16} - 86.52238\,x^{15} + 10.86007\,x^{14} + 423.10927\,x^{13} + 703.98001\,x^{12} + 188.11898\,x^{11} - 715.92046\,x^{10} - 904.33453\,x^{9} - 327.36651\,x^{8} + 130.24328x^{7} + 135.60344\,x^{6} + 18.51769\,x^{5} - 13.63601\,x^{4} - 4.09625\,x^{3} + 0.34336\,x^{2} + 0.18546\,x^{1} + 0.00493$

$SumCh(RD_e, 2) = 3693.48989$

$Ch(RD_e, 3) = x^{18} - 29.64728\,x^{16} - 86.52238\,x^{15} + 10.87284\,x^{14} + 423.30087\,x^{13} + 705.29232\,x^{12} + 193.54512\,x^{11} - 700.86841\,x^{10} - 874.95926\,x^{9} - 286.36122\,x^{8} + 170.81936\,x^{7} + 162.80754\,x^{6} + 29.26616\,x^{5} - 12.58471\,x^{4} - 5.19546\,x^{3} - 0.11363\,x^{2} + 0.15804\,x^{1} + 0.01503$

$SumCh(RD_e, 3) : 3693.32964$

$Ch(RD_p, 2) = x^{18} - 20.90141\,x^{16} - 28.32566\,x^{15} + 110.69769\,x^{14} + 264.62231\,x^{13} - 62.04211\,x^{12} - 626.35739\,x^{11} - 397.89036\,x^{10} + 405.45258\,x^{9} + 512.44210\,x^{8} - 12.15070x^{7} - 217.60485\,x^{6} - 61.65536\,x^{5} + 31.16778\,x^{4} + 16.28853\,x^{3} + 0.24539\,x^{2} - 0.87565\,x^{1} - 0.11548$

$SumCh(RD_p, 2) = 2769.83536$

$Ch(RD_p, 3) = x^{18} - 20.90141\,x^{16} - 28.32566\,x^{15} + 110.69981\,x^{14} + 264.65957\,x^{13} - 61.74455\,x^{12} - 624.92563\,x^{11} - 393.29342\,x^{10} + 415.66413\,x^{9} + 528.01536\,x^{8} + 3.00077x^{7} - 210.96757\,x^{6} - 65.31885\,x^{5} + 24.40092\,x^{4} + 13.34204\,x^{3} + 0.57945\,x^{2} - 0.32099\,x^{1} - 0.01248$

$SumCh(RD_p, 3) = 2767.17264$

The input of molecular graph was made by using HyperChem 4.5 [21]. The characteristic polynomi-

Table 6 — Walk numbers, ${}^{c}W_A$, in nonanes

No	${}^{1}W_A$	${}^{2}W_A$	${}^{3}W_A$	${}^{4}W_A$	${}^{5}W_A$	${}^{6}W_A$	${}^{7}W_A$	${}^{8}W_A$
1	16	30	56	106	200	380	720	1370
2	16	32	60	120	228	456	872	1742
3	16	32	62	124	244	490	968	1948
4	16	32	62	126	248	506	1002	2046
5	16	32	64	130	264	538	1096	2236
6	16	32	64	132	268	554	1128	2332
7	16	36	68	154	296	674	1304	2980
8	16	34	68	142	292	608	1260	2624
9	16	34	66	124	278	600	1178	2544
10	16	34	66	138	274	568	1138	2346
11	16	34	64	134	256	534	1024	2134
12	16	36	72	162	332	748	1544	3482
13	16	34	70	148	310	656	1380	2920
14	16	34	68	144	292	618	1256	2658
15	16	36	72	166	336	780	1584	3684
16	16	34	70	150	314	670	1410	3002
17	16	34	68	146	296	632	1288	2788
18	16	36	76	172	368	834	1788	4054
19	16	34	72	154	328	702	1496	3202
20	16	38	78	180	382	872	1874	4258
21	16	38	74	176	346	822	1622	3850
22	16	38	72	168	328	752	1496	3388
23	16	38	80	184	400	912	2000	4548
24	16	36	76	166	356	776	1668	3634
25	16	36	72	158	324	702	1456	3132
26	16	38	76	180	364	862	1748	4148
27	16	38	82	188	414	946	2090	4772
28	16	36	80	180	400	900	2000	4500
29	16	38	80	188	400	936	2000	4672
30	16	38	84	192	432	982	2216	5030
31	16	36	76	170	360	840	1704	3804
32	16	42	92	222	508	1200	2780	6522
33	16	40	84	200	428	1014	2176	5152
34	16	42	80	210	400	1050	2000	5250
35	16	40	88	204	464	1064	2432	5568

als were computed by the LeVerier-Frame-Fadeev method[22,23] implemented in an original program.

Conclusions

The molecular graph matrices discussed here represent a convenient and rigorous base for devising topological descriptors. In the present investigation we have explored the characteristic polynomial and spectra invariants. The new structural descriptors were tested in correlations with physico-chemical properties of alkanes, and were found to provide good results. The molecular descriptors similar to the well known Z index of Hosoya, SumCh(M), deserve subsequent attention, both in QSPR and isomer coding studies.

Acknowledgement

One of the authors (Ovidiu Ivanciuc) acknowledges a software grant offered by Hypercube, and the partial financial support of this research by the Ministry of Research and Technology under Grant 310 T-A69 and the Ministry of Education under Grant 7001 T-34.

Appendix 1

1. Spectra of the $Ch(D_{\Delta},G,x)$ polynomials for molecular graphs representing the paths L_3 to L_8 and the cycles C_4 to C_8

$Sp(D_{\Delta},L_3) = \{1, 0, -1\}$

$Sp(D_{\Delta},L_4) = \{3.30278, 0.30278, -0.30278, -3.30278\}$

$Sp(D_{\Delta},L_5) = \{7.62160, 0.40512, 0.28837, -0.90997, -7.40512\}$

$Sp(D_{\Delta},L_6) = \{14.53505, 0.53787, 0.36256, 0.26931, -1.89761, -13.80718\}$

$Sp(D_{\Delta},L_7) = \{24.62725, 0.69750, 0.45700, 0.31155, 0.26527, -3.34952, -23.00905\}$

$Sp(D_{\Delta},L_8) = \{38.48321, 0.88290, 0.56855, 0.36764, 0.29896, 0.26027, -5.35072, -35.51081\}$

$Sp(D_{\Delta},C_4) = \{1, 1, -1, -1\}$

$Sp(D_{\Delta},C_5) = \{2, 0.61803, 0.61803, -1.61803, -1.61803\}$

$Sp(D_{\Delta},C_6) = \{5, 2, 2, -1, -4, -4\}$

$Sp(D_{\Delta},C_7) = \{8, 1.93900, 1.93900, -0.08815, -0.08815, -5.85086, -5.85086\}$

$Sp(D_{\Delta},C_8) = \{14, 4, 4, 2, -1.75736, -1.75736, -10.24264, -10.24264\}$

2. Spectra of the $Ch(D_p,G,x)$ polynomials for molecular graphs representing the paths L_3 to L_8 and the cycles C_3 to C_8

$Sp(D_p,L_2) = \{1, -1\}$

$Sp(D_p,L_3) = \{3.56155, -0.56155, -3\}$

$Sp(D_p,L_4) = \{8.21699, -0.29844, -1.21699, -6.70156\}$

$Sp(D_p,L_5) = \{15.54948, -0.28385, -0.39672, -2.26563, -12.60328\}$

$Sp(D_p,L_6) = \{26.14393, -0.26890, -0.35316, -0.52612, -3.79078, -21.20497\}$

$Sp(D_p,L_7) = \{40.58558, -0.26467, -0.31062, -0.44350, -0.68275, -5.87741, -33.00663\}$

$Sp(D_p,L_8) = \{59.45975, -0.26018, -0.29753, -0.36626, -0.55141, -0.86532, -8.61080, -48.50824\}$

$Sp(D_p,C_3) = \{2, -1, -1\}$

$Sp(D_p,C_4) = \{5, 1, -3, -3\}$

$Sp(D_p,C_5) = \{8, 0.23607, 0.23607, -4.23607, -4.23607\}$

$Sp(D_p,C_6) = \{14, 2, 2, -2, -8, -8\}$

$Sp(D_p,C_7) = \{20, 1.63102, 1.63102, -0.73125, -0.73125, -10.89977, -10.89977\}$

$Sp(D_p,C_8) = \{30, 4, 4, 2, -2.92893, -2.92893, -17.07107, -17.07107\}$

3. Spectra of the $Ch(RD_e,G,x)$ polynomials for molecular graphs representing the paths L_3 to L_8 and the cycles C_3 to C_8.

$Sp(RD_e,L_3) = \{1.68614, -0.50000, -1.18614\}$

$Sp(RD_e,L_4) = \{2.20326, -0.06574, -0.86992, -1.26759\}$

$Sp(RD_e,L_5) = \{2.61684, 0.30328, -0.56081, -1.05328, -1.30603\}$

$Sp(RD_e,L_6) = \{2.96093, 0.62047, -0.27855, -0.82466, -1.14905, -1.32914\}$

$Sp(RD_e,L_7) = \{3.25530, 0.89725, -0.02402, -0.60368, -0.97150, -1.21024, -1.34311\}$

$Sp(RD_e,L_8) = \{3.51238, 1.14216, 0.20578, -0.39636, -0.79270, -1.06925, -1.24927, -1.35275\}$

$Sp(RD_e,C_3) = \{2.00000, -1.00000, -1.00000\}$

$Sp(RD_e,C_4) = \{2.50000, -0.50000, -0.50000, -1.50000\}$

$Sp(RD_e,C_5) = \{3.00000, -0.19098, -0.19098, -1.30902, -1.30902\}$

$Sp(RD_e,C_6) = \{3.33333, 0.16667, 0.16667, -1.16667, -1.16667, -1.33333\}$

$Sp(RD_e,C_7) = \{3.66667, 0.42381, 0.42381, -0.93035, -0.93035, -1.32680, -1.32680\}$

$Sp(RD_e,C_8) = \{3.91667, 0.69281, 0.69281, -0.75000, -0.75000, -1.19281, -1.19281, -1.41667\}$

4. Spectra of the $Ch(RD_p,G,x)$ polynomials for molecular graphs representing the paths L_3 to L_8 and the cycles C_3 to C_8

$Sp(RD_p,L_2) = \{1, -1\}$

$Sp(RD_p,L_3) = \{1.59067, -0.33333, -1.25733\}$

$Sp(RD_p,L_4) = \{1.98025, 0.20283, -0.81359, -1.36950\}$

$Sp(RD_p,L_5) = \{2.25741, 0.62479, -0.39810, -1.05813, -1.42598\}$

$Sp(RD_p,L_6) = \{2.46525, 0.96196, -0.03651, -0.73595, -1.19541, -1.45934\}$

$Sp(RD_p,L_7) = \{2.62724, 1.23669, 0.27400, -0.43524, -0.94005, -1.28240, -1.48024\}$

$Sp(RD_p,L_8) = \{2.75725, 1.46465, 0.54121, -0.16319, -0.68938, -1.07608, -1.34003, -1.49443\}$

$Sp(RD_p,C_3) = \{2, -1, -1\}$

$Sp(RD_p,C_4) = \{2.33333, -0.33333, -0.33333, -1.66667\}$

$Sp(RD_p,C_5) = \{2.66667, 0.07869, 0.07869, -1.41202, -1.41202\}$

$Sp(RD_p,C_6) = \{2.83333, 0.50000, 0.50000, -1.16667, -1.16667, -1.50000\}$

$Sp(RD_p,C_7) = \{3, 0.79831, 0.79831, -0.83786, -0.83786, -1.46045, -1.46045\}$

$Sp(RD_p,C_8) = \{3.10000, 1.07851, 1.07851, -0.56667, -0.56667, -1.27851, -1.27851, -1.56667\}$

References

1 Wiener H, *J Am chem Soc*, 69 (1947) 17.

2 Gutman I & Körtvélyesi T, *Z Naturforsch*, 50a (1995) 669.

3 Nikoliæ S, Trinajstiæ N & Mihaliæ Z, *Croat chem Acta*, 68 (1995) 105.

4 Gutman I, Yeh Y N, Lee S L & Luo Y L, *Indian J Chem*, 32A (1993) 651.

5 Nikoliæ S, Mediæ-Šariæ M & Matijeviæ-Sosa J, *Croat chem Acta*, 66 (1993) 151.

6 Randiæ M, *Chem Phys Lett*, 211 (1993) 478

7 Randiæ M, Guo X, Oxley T & Krishnapriyan H, *J chem Inf Comput Sci*, 33 (1993) 700.

8 Randiæ M, Guo X, Oxley T, Krishnapriyan H & Nayor L, *J chem Inf Comput Sci*, 34 (1994) 361.

9 Lukovits I & Linert W, *J chem Inf Comput Sci*, 34 (1994) 899.

10 Lukovits I, *Croat Chem Acta*, 68 (1995) 99.

11 Lukovits I & Gutman I, *Commun Math Comput Chem (MATCH)*, 31 (1994) 133.

12 Hosoya H, *Bull chem Soc Jpn*, 44 (1971) 2332.

13 Diudea M V, *J chem Inf Comput Sci*, 36 (1996) 535.

14 Ivanciuc O, *Rev Roum Chim*, 34 (1989) 1361; Ivanciuc O, Balaban T S & Balaban A T, *J math Chem*, 12 (1993) 309.

15 Plavšiæ D, Nikoliæ S, Trinajstiæ N & Mihaliæ Z, *J math Chem*, 12 (1993) 235.

16 Diudea M V & Ivanciuc O, *Molecular topology*, (Comprex, Cluj), 1995 (in Romanian).

17 Needham D E, Wei I C & Seybold P G, *J Am chem Soc*, 110 (1988) 4185.

18 Rücker G & Rücker Ch, *J chem Inf Comput Sci*, *33* (1993) 683.

19 Diudea M V, Topan M & Graovac A, *J chem Inf Comput Sci*, 34 (1994) 1072.

20 Ivanciuc O, Balaban T S & Balaban A T, *J math Chem*, 14 (1993) 21.

21 HyperChem 4.5 produced by Hypercube, Inc., 419 Philip St., Waterloo, Ontario, Canada, N2L 3X2, Tel. (519)-725-4040; Fax:(519)-725-5193; E-mail: info@hyper.com; WWW: http: //www.hyper.com.

22 Balasubramanian K, *J comput Chem*, 5 (1984) 387.

23 Barysz M, Nikoliæ S & Trinajstiæ N, *Commun Math Comput Chem (MATCH)*, 19 (1986) 117.