

Graph Invariants for Fullerenes

A. T. Balaban,^{*,†,‡} X. Liu,[†] D. J. Klein,^{*,†} D. Babić,[§] T. G. Schmalz,[†] W. A. Seitz,[†] and M. Randić[⊥]

Texas A&M University, Galveston, Texas 77553-1675, Polytechnic University, Organic Chemistry Department, 77206 Bucharest, Roumania, Rugjer Bošković Institute, 41001 Zagreb, Croatia, and Drake University, Department of Mathematics and Computer Science, Des Moines, Iowa 50311

Received January 20, 1995[®]

A comparison is made among several theoretical invariants in characterizing classes of fullerenes with up through 96 carbon atoms, namely all isomers of C_{60} , and all preferable higher fullerenes (i.e., fullerenes without abutting pentagons) with 60–96 carbon atoms. Four classes of fullerene structures were considered: first, the 1812 C_{60} isomers; second, their duals; third, the 558 isolated-pentagon fullerenes with 60–96 carbon atoms; and fourth their duals. The invariants that were investigated are either purely graph-theoretical (including several topological indices, the number of Hamiltonian circuits, and the number of spanning trees), or associated with quantum-chemical interpretations (including Hückel π -electron energy and HOMO–LUMO gap). In addition to such invariants that are expressed as a single numerical value (integer, rational, or real number), associations of several invariants have also been explored, such as the numbers of Kekulé structures with those of conjugated circuits. Some of the above invariants succeeded in uniquely characterizing the studied fullerenes. This finding has implications for a rational coding and nomenclature of fullerenes.

1. INTRODUCTION

Graph invariants have a broad use in characterizing molecular structures, not only with respect to correlations between structure and physical and chemical properties but also to nomenclature and chemical documentation. Thus an investigation of graph invariants is of value for nontrivial classes of molecular graphs. Indeed such studies have been reported for several classes of graphs, but little seems to have been done with graphs representing what are termed "cluster" molecules. Their frequent high regularity makes them especially sensitive candidates to test nomenclatural and documentational features of such structures. Here we study one such class of clusters of much current interest, namely the *fullerenes*, which we identify to be represented by trivalent graphs embeddable (without crossing bonds) on the surface of a sphere such that all the faces form five- and six-cycles. Indeed, fullerenes opened a new direction in chemistry. The interest in such species is so intense that it has given rise to the journal "*Fullerene Science and Technology*",¹ several books,^{2–5} special issues of journals,^{6–9} book chapters,^{10–12} reviews,^{13–20} and an exponentially increasing flow of papers.²¹

In the present contribution, a range of single- and multivalued graph invariants are studied for two sets of fullerenes. From all possible constitutional isomers of carbon cages C_n with $n = 20–60$, in the first set (Table 1) we consider all 60-atom fullerenes which we have previously comprehensively generated,^{22–25} finding exactly 1812 such graphs, each in correspondence with a different conceivable constitutional isomer of C_{60} (Table 2). This is obviously a nontrivial and interesting set of structures of which only a single one (the truncated icosahedron or buckminster-

Table 1. Counts for Constitutional Isomers of Carbon Cages C_n with 20–60 Vertices (Stereoisomerism Is Ignored)

n	no. of isomers	no. of isomers with $q = A = 0$	n	no. of isomers	no. of isomers with $q = A = 0$
20	1	0	42	45	5
22	0	0	44	89	16
24	1	0	46	116	20
26	1	0	48	199	43
28	2	0	50	271	62
30	3	0	52	437	123
32	6	0	54	580	174
34	6	0	56	924	300
36	15	2	58	1205	442
38	17	1	60	1812	704
40	40	5			

fullerene) is experimentally well characterized, though there is some suggestion²⁶ for the existence of others. The parameters A and B of Table 2 are explained in section 4, and entries in Table 2 are the same numbers as in a previously published²⁷ table of the 1812 isomers in terms of parameters p (the number of abutting pairs of pentagons sharing an edge) and q (the numbers of pentagon triplets sharing a vertex).

Apparently the most useful theoretical restriction, which is chemically well motivated, is that which focuses on those fullerenes with no abutting pairs of pentagonal rings ($p = q = 0$).^{27–34} Thence our second set is that of *all such isolated-pentagon fullerenes with 96 or fewer sites*, i.e., carbon atoms. We have comprehensively generated these structures via the same algorithm as used for the first set. There are 558 such fullerene graphs of various sizes, as indicated in Table 3; they have also been called *preferable fullerenes*.

For each of these two sets of fullerenes we have computed several graph invariants and used these to classify the fullerenes. We place all the fullerenes with the same value for a corresponding set X of invariants X into a single X -equivalence class. For several different choices of X we have investigated the consequent equivalence relations.

[†] Texas A&M University.

[‡] Polytechnic University.

[§] Rugjer Bošković Institute.

[⊥] Drake University.

[®] Abstract published in *Advance ACS Abstracts*, April 15, 1995.

Table 2. Numbers of C_{60} Isomers in Terms of A and B , where $A = q$ is the Number of Points Common to Three Pentagons, and B Is the Number of Points Common to Two Pentagons and One Hexagon: $A(5^3) = q$; $B(5^2,6) = 2p - 3q$; $C(5,6^2) = 60 - 3A - 2B = 60 - 4p + 3q$; $D(6^3) = 2A + B = 2p - q$

B	A										
	0	1	2	3	4	5	6	7	8	9	10
0	1										
1											
2	—										
3											
4	1										
5											
6	3										
7		5									
8	17		6		4						
9		39		11		2					
10	81		54		11		2		—		1
11		147		42				—		—	
12	215		131		25		2		1		
13		214		54		5		—			
14	210		116		10		1				
15		132		16		—					
16	145		31		3						
17		28		2							
18	23		6								
19		1									
20	7										
21											
22	—										
23											
24	1										

Whenever the numbers of X -equivalence classes equals the numbers of the corresponding fullerenes, the invariants achieve a unique characterization. The various invariants studied are described in sections 2–4, while the results for the equivalence classes of the two sets of fullerenes are described in sections 5 and 6.

2. SINGLE-VALUE INVARIANTS

Several classes of single-value numerical invariants for molecular (constitutional) graphs have been tested. The first class is constituted by *topological indices*, which can be (i) integers, such as the Wiener index (see below);³⁵ (ii) rational numbers, such as the Kirchhoff numbers³⁶ (K_f , the half-sum of all entries in the resistance-distance matrix of the graph): on multiplying all these entries by the number of spanning trees, one obtains integers;³⁷ or (iii) real numbers, such as the average distance-sum connectivity (*vide infra* index J).^{38,39}

Topological indices^{40–45} are typically based either upon the *adjacency matrix* with entries being 1 for adjacent vertices, and 0 otherwise, or upon the *graphical distance matrix* \mathbf{D} with the (i,j) th element being the number of edges in a shortest path between vertices i and j . If one sums over rows or columns of \mathbf{D} , then one obtains a vertex invariant d_i which is called the distasum (or the status) of the corresponding vertex i .

The Wiener index is the half-sum of all distasums, or of all entries in the distance matrix \mathbf{D}

$$W = \sum_i d_i / 2 \quad (2.1)$$

From distasums one can also obtain the average distasum connectivity index J , which is typically a real, noninteger, number

$$J = \frac{e}{n - e + 2} \sum_{\text{edges } ij} (d_i d_j)^{-1/2} \quad (2.2)$$

Here e and n are the numbers of edges and vertices, respectively.

Table 3. Counts for Constitutional Isomers of Preferable Cages, Ignoring Enantiomerism (Isolated Pentagon Fullerenes C_n)

n	60	70	72	74	76	78	80	82	84	86	88	90	92	94	96	
counts	1	1	1	1	2	5	7	9	24	19	35	46	86	134	187	sum = 558

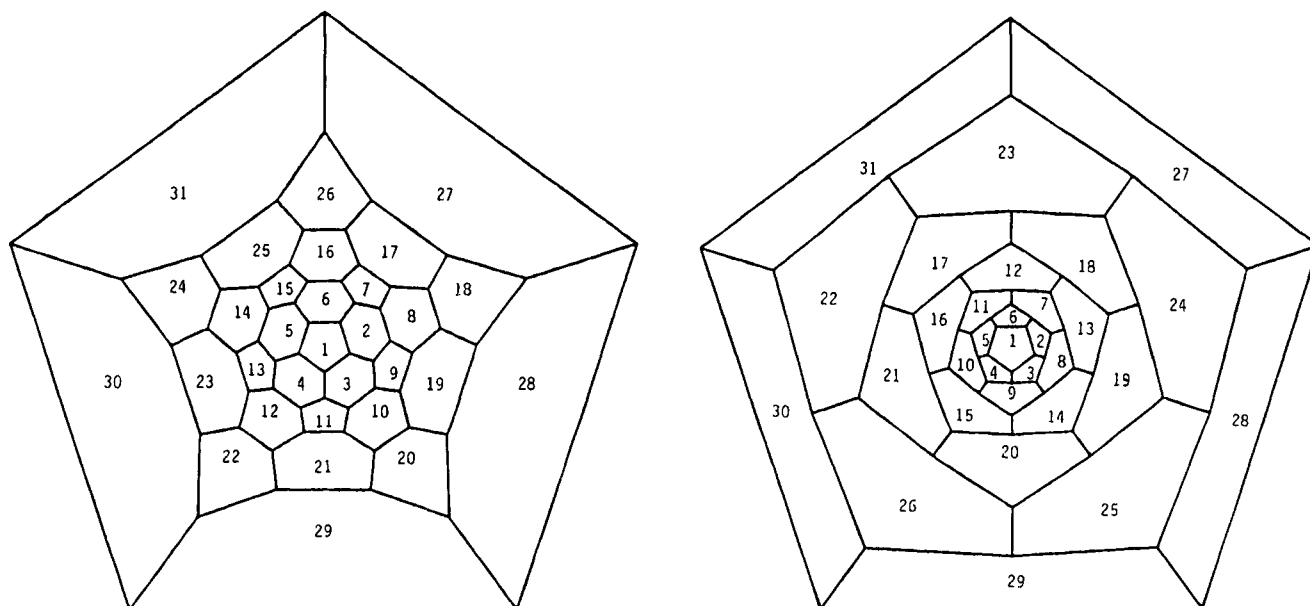


Figure 1. Schlegel diagrams for the two extremal C_{60} isomers discussed in the text; the numberings of pentagons are those for which the distance matrices of reduced duals will be shown in Figures 3 and 5. See also Scheme 1.

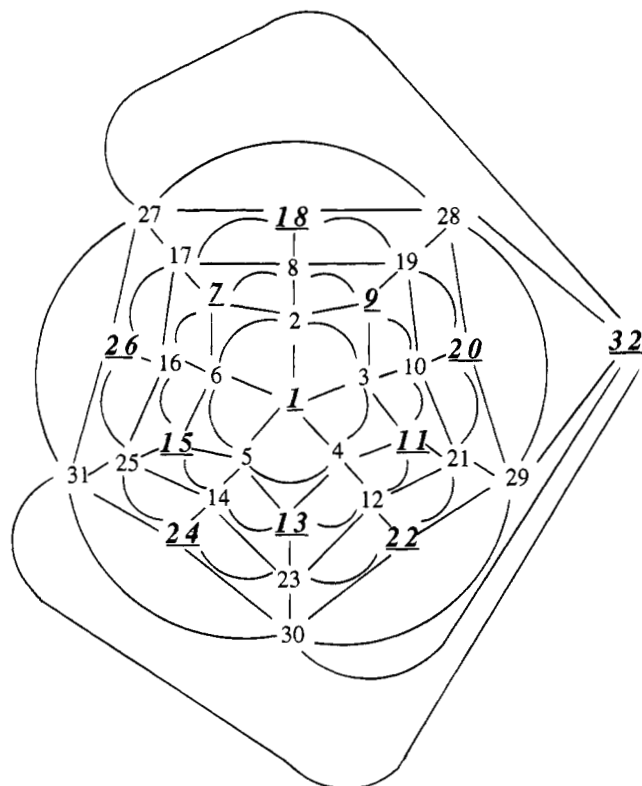


Figure 2. The dual graph of buckminsterfullerene. Vertices with their numberings in underlined italics correspond to five-membered rings.

	1	7	9	11	13	15	18	20	22	24	26	32	Distasum
1	0	2	2	2	2	2	3	3	3	3	3	5	30
7	2	0	2	2	2	2	3	3	5	3	3	3	30
9	2	2	0	2	2	2	3	3	3	5	3	3	30
11	2	2	2	0	2	2	3	3	3	3	5	3	30
13	2	2	2	2	0	2	5	3	3	3	3	3	30
15	2	2	2	2	2	0	3	5	3	3	3	3	30
18	3	3	3	3	5	3	0	2	2	2	2	2	30
20	3	3	3	3	3	5	2	0	2	2	2	2	30
22	3	5	3	3	3	3	2	2	0	2	2	2	30
24	3	3	5	3	3	3	2	2	2	0	2	2	30
26	3	3	3	5	3	3	2	2	2	2	0	2	30
32	5	3	3	3	3	3	2	2	2	2	2	0	30

Distance distribution vector :

Distance : 2 3 5

Number of occurrences : 30 30 6

Figure 3. Distance matrix for the reduced dual of buckminsterfullerene (second in Scheme 1).

As shown by Bonchev and Trinajstić,⁴⁶ information-theoretic formulas applied to a set of non-negative numbers often lead to "better" global indices than the result of simple summation. Thus, if instead of contracting together all distance sums by simple summation into the Wiener index W , one utilizes Shannon's formula^{45,46} with this set of distance sums, then one obtains an information-theoretic index I_D^W

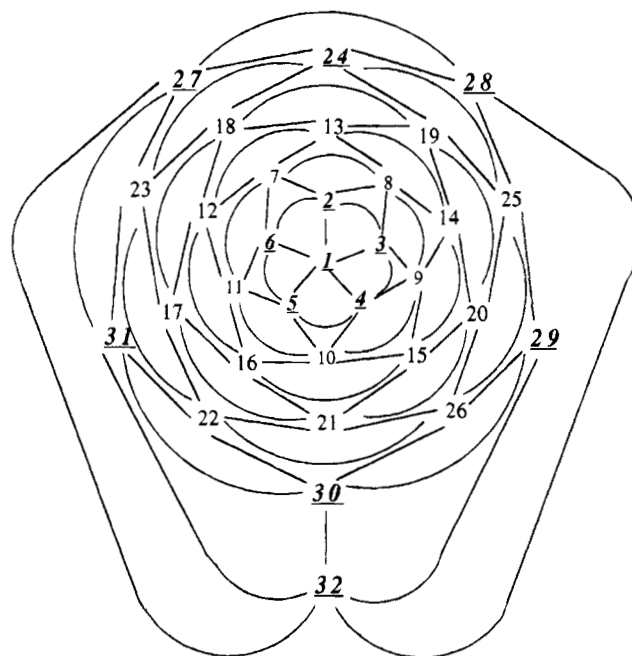


Figure 4. The 32-vertex dual of C_{60} , $p = 20$, $q = A = B = 10$ (*italic underlined numbers indicate vertices corresponding to five-membered faces*).

	1	2	3	4	5	6	27	28	29	30	31	32	Distasum
1	0	1	1	1	1	1	5	5	5	6	6	7	42
2	1	0	1	2	2	1	5	5	5	5	5	6	38
3	1	1	0	1	2	2	5	5	5	5	5	6	38
4	1	2	1	0	1	2	5	5	5	5	5	6	33
5	1	2	2	1	0	1	5	5	5	5	5	6	38
6	1	1	2	2	1	0	5	5	5	5	5	6	38
27	6	5	5	5	5	5	0	1	2	2	1	1	38
28	6	5	5	5	5	5	1	0	1	2	2	1	38
29	6	5	5	5	5	5	2	1	0	1	2	1	38
30	6	5	5	5	5	5	2	2	1	0	1	1	38
31	6	5	5	5	5	5	1	2	2	1	0	1	38
32	7	6	6	6	6	6	2	1	1	1	1	0	42

Distance distribution vector :

Distance : 1 2 5 6 7

Number of occurrences : 20 10 25 10 1

Figure 5. Distance matrix of the reduced dual of the C_{60} isomer with $p = 20$, $q = A = B = 10$ (first in Scheme 1).

with better discrimination ability than W .

The second type of invariant is also graph-theoretical but is usually given a quantum-chemical interpretation. In Hückel molecular orbital (HMO) theory, the eigenvalues λ_i of the adjacency matrix correspond to the π -MO energies, $E_i = \alpha + \beta\lambda_i$. The total π -electron energy (E_π^{total}) is minus twice the sum of the highest $n/2$ values λ_i . In the framework of valence-bond theory, a few integer-number parameters play an important part: the number K of Kekulé structures (or the Kekulé structure count) and the numbers of conjugated circuits C_i for $i = 6-, 8-,$ and 10 -membered circuits. The (conjugated-circuit) resonance energy is reported as $\text{RE/g} = \{R_1C_6 + Q_2C_8 + R_2C_{10}\}/nK(\text{graphite result})$, where the

parameters $R_1, R_2 > 0$, $Q_2 < 0$, are as defined elsewhere,^{48,50} with reference being taken against the resonance energy of graphite.

Finally there are other, perhaps less commonly used, invariants. These include the number H of Hamiltonian circuits, the number t of spanning trees, and invariants expressed in terms of the resistance distance matrix Ω with elements Ω_{ij} giving the effective resistance of the graph network with unit resistors on each edge when a constant voltage is connected across vertices i and j . Since Ω_{ij} is a distance function³⁶ (as is d_{ij}), invariants involving the Ω matrix, in ways analogous to those defined in terms of the D matrix, are possible. In particular, the *Kirchhoff number*, K_f , which is the half-sum of all the Ω_{ij} values is analogous to W . An information-theoretic index I_D^K may be devised, analogously to I_D^W .

Further we report the *mean inertia* of the polyhedron graph as given in terms of "topological coordinates" determined from the three lowest "p-wave" eigenvalues to the Laplacian matrix of the graph. If we denote by λ_x , λ_y , and λ_z these three eigenvalues, then the inertia is proportional to $\lambda_x^{-1} + \lambda_y^{-1} + \lambda_z^{-1}$.

3. DUAL GRAPH INVARIANTS

Instead of using graph G and its associated matrices, one can consider its dual G^* . Several graph invariants have been computed which typically may be viewed as the same as those presented in the preceding section but applied to the "duals" of fullerene graphs. Since each fullerene graph G may be viewed as corresponding to a polyhedron P_G (with vertices and edges of the graph and polyhedron identified to one another), one may define a dual G^* with vertices corresponding to faces of P_G and edges corresponding to adjacent pairs of faces of P_G . Then each graph invariant for G^* is in turn a type of graph invariant for G . That is, for each invariant X we define a new invariant X^* via

$$X^*(G) = X(G^*)$$

Thus we obtain *dual graph invariants*. For a proper fullerene C_n with 12 pentagons and h hexagons, we have $h = -10 + n/2$ hexagons, and $f = 12 + h = 2 + n/2$ faces; therefore, the dual graph has f vertices (by definition).

Besides the invariants of section 2 applied to G^* , we consider a few additional ones. In particular given a distance matrix D for G^* we may select from it the submatrix D_5 associated with the five-valent sites of G^* . Since each fullerene G necessarily has exactly 12 pentagons, the dual G^* must have only 12 five-valent vertices and D_5 is a smaller (12×12) matrix. Evidently, D_5 encodes interesting information about the singular features (the pentagons) of P_G . The resulting invariants are called *reduced dual invariants*. The submatrix D_5 may be considered as the distance matrix of an edge-weighted complete graph K_{12} with weights corresponding to the distances in the dual graph of the fullerene; in other words, the entries in this submatrix represent graphical distances between vertices associated with pentagons in the dual graph of the fullerene (Figures 1–5 exemplify the above definitions for two isomers of C_{60}); alternatively, one might consider inverse distances as edge weights.

Now, whatever sorts of graph invariants $X(G)$ are computed using a distance matrix for G , they lead not only to $X^*(G) = X(G^*)$ but also to invariants $X_5^*(G)$ for which D_5 takes the part of the whole distance matrix.

Of course, for preferable fullerenes the important condition that G be an isolated-pentagon fullerene is that D_5 contain no elements equal to 1, i. e., no two pentagons are adjacent.

4. MULTIVALUE SUBSTRUCTURAL INVARIANTS

The more-or-less standard sets S of invariants include the following:

- the set of adjacency-matrix eigenvalues (corresponding to the Hückel MO energy levels)
- the distasum set (consisting of the row sums of the usual graph distance matrix)
- the resdisum set, as in (b), but using the resistance distances
- a conjugated circuit set (vector V , consisting of K , C_6 , C_8 , and C_{10})
- the eigenvalue-eigenspace tabular invariant⁴⁷ utilized in the isomorphism checking of the multiply generated fullerene cages. In fact, we have found this invariant to be faithful for all of the over 20 000 distinct fullerenes we have generated.

Another general kind of multivalued invariant is based on substructure identification. More precisely, it consists of counts of subgraphs isomorphic to members of a given set Γ of abstract graphs. For instance, such counts for the single-point graph and the single-edge two-point graph yield a classification based upon vertex and edge counts. For a meaningful classification one desires a "natural" set Γ of graphs. One way to choose a natural set is to pay attention to the "environment" of a selected type of subgraph. A natural meaning for environment in the present context refers to the rings surrounding the basic type of subgraph. For instance, if we choose the basic type of subgraph to be a single vertex, the consequent set Γ of graphs including the surrounding rings are as indicated in Figure 6. This categorization of cages has already been suggested (though explained slightly differently) and utilized.⁴⁸ The number A of points surrounded by three pentagons is the same as the parameter q used previously; however, the second parameter p , i.e., the number of pentagon pairs with a common edge, used previously, leads now to different descriptors. In Figure 6 and Table 2, a second parameter B is defined as the number of points surrounded by two pentagons and one hexagon. Two other types of vertices, C and D , defined in Figure 6, can be deduced from A and B , as shown in the heading of Table 2.

A second choice for Γ (substructural set I) is based on the different ring environments for the subgraph of a single site along with its nearest neighbors, as indicated in Figure 7. The 60 vertices of fullerenes in the first set are partitioned into 20 types denoted by letters a–t. Clearly, this idea could be extended to environments based on subgraphs of sites, neighbors and next neighbors (or farther), as seen in Figure 8, but this potentially complete series is not investigated to any higher order (beyond that of Figure 7) here.

Instead, as indicated in Figure 9, we consider a third choice based upon environments for single-ring subgraphs. This particular choice has some affinity with the standard IUPAC

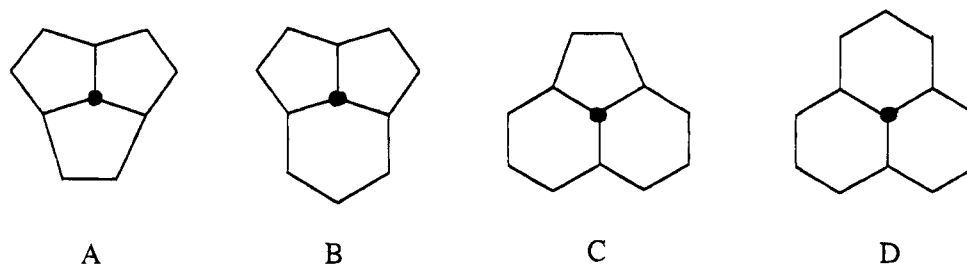


Figure 6. Vertices in fullerenes are classified into types A–D according to the sizes of the three rings to which they belong.

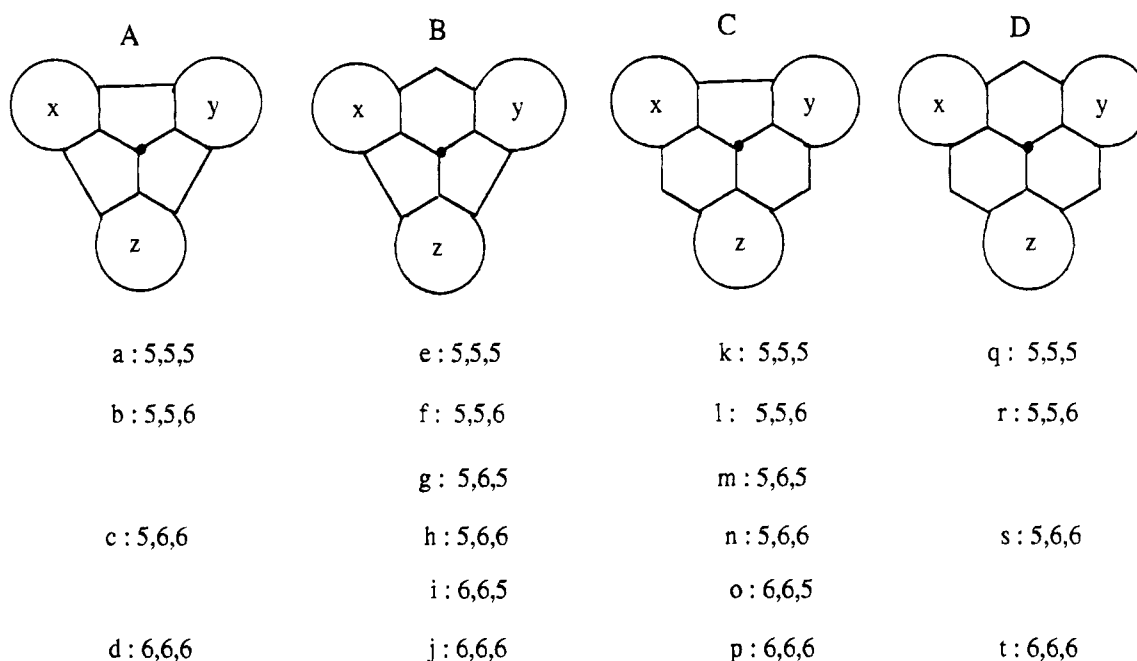


Figure 7. The 20 types of vertices (central points in the diagrams below) according to the rings surround them: the values for ring sizes x , y , and z are indicated for each of the types a–t. Vertices of type A and D give rise to four types each, whereas vertices B and C give rise to six types each (substructural set I). When cage enantiomerism is taken into account, types g, h, m, and n lead to nonsuperimposable mirror images (substructural set II).

or CAS nomenclature for benzenoids: for a given hexagonal ring (or a longer acene), one identifies first the edges where fusion to other hexagonal rings occurs. In the present scheme, as shown in Figure 9, the different manners of fusion of hexagonal rings to the central ring are subject to enumeration. If instead of the present approach (abbreviation through enumeration) one were to label these surrounding rings, repeating the process for every choice of central ring, a comprehensive IUPAC-related nomenclature would presumably result. The most frequently employed notation system for fullerenes now uses the spiral code,^{49,50} though fullerene cages without a spiral have been shown to be possible.^{51,52}

Another point of interest is that the types of substructural invariants (as indicated in Figures 7–9) may be easily extended to distinguish a fullerene and its mirror image. Since a structure and its mirror image have the same graph, these extended invariants clearly must go beyond the abstract graph. Here this extension is taken to be the graph embeddings in a plane or better a spherical surface. As shown by Whitney for polyhedral graphs⁵³ (i.e., graphs which are three-connected and graph-theoretically planar) there are either one or two embeddings on a spherical surface. The case of two embeddings corresponds exactly to polyhedra which are chiral (i.e., in a combinatorial equivalence class which is chiral). In addition to paying attention to the embedding of the parent graph we also pay attention to the

embeddings of the subgraph structures (associated to Γ). That is, the substructural invariant described in Figure 7 (cases $x \neq y$ for B and C) would upon extension distinguish between the embeddings on the left and right sides of Figures 7 (parts B and C). Similarly the extension of the ring-environment invariant of Figure 9 would now distinguish between the embeddings on the left and right sides of this figure. Evidently a graph embedding G_+ and its mirror image G_- give rise to counts $n_{\gamma+}$ and $n_{\gamma-}$ for mirror-image pairs γ_+ and γ_- of substructures such that:

$$n_{\gamma+}(G_+) = n_{\gamma-}(G_-)$$

The ideas of the previous paragraph lead to true graph invariants more refined than originally described in connection with Figures 7–9. Given a graph G with embeddings G_+ and G_- , one obtains two graph-embedding invariants as explained in the previous paragraph, but the first of these two invariants when lexicographically ordered can be used as an invariant for G .

In more specific terms, the above substructural sets can be summarized as follows. Taking into account that all n -atom fullerenes C_n under discussion consist of 12 pentagons and a variable number $h = (n - 20)/2$ of hexagons, the topology of ring fusion can give rise to 20 types of carbon atom environments in terms of nearest ring sizes, as shown in Figure 7 (substructural set I). One can therefore partition

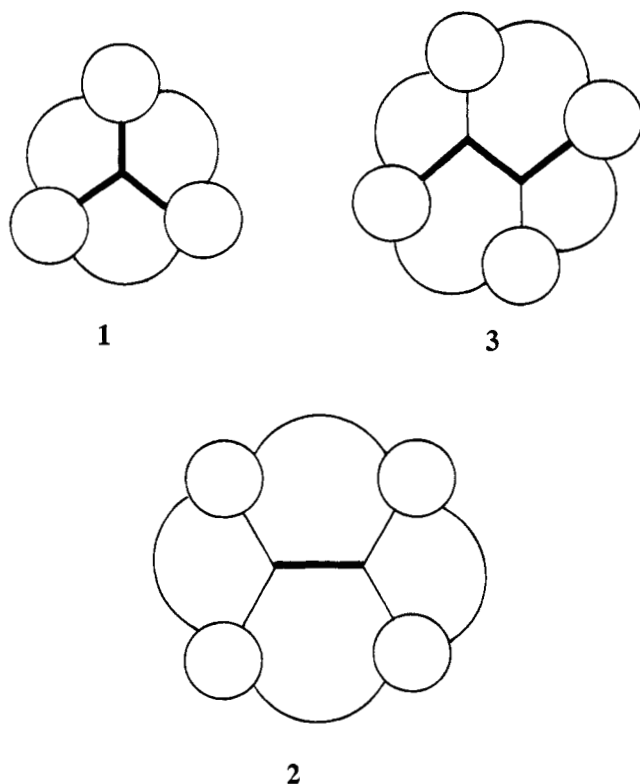


Figure 8. Various possibilities of partitioning vertices (1, substructural set I), edges (2, substructural set III), or sequences of edges according to their adjoining ring sizes (3).

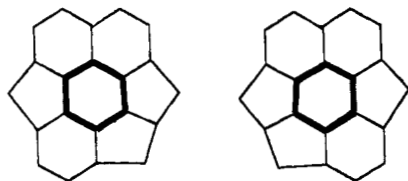


Figure 9. Two mirror images of a portion of a fullerene with a hexagon surrounded by three hexagons and three pentagons.

all n carbon atoms into these types arranged in the alphabetical order shown in Figure 7 and characterize fullerenes by a 40-digit number: two digits indicate the number of atoms for each type, the default value being 00. In this scheme, the designation for buckminsterfullerene ($p = q = A = B = 0$), as will be seen in Table 5, shows that all 60 carbon atoms are of type o, whereas for the isomer having two halves of a pentagonal dodecahedron separated by a "belt" or "tube" of 20 hexagons ($p = 20, q = A = B = 10$) there are 20 vertices of type q, and ten vertices each for type b, i, l, r, again as will be seen in Table 5.

It may be observed that four of these 20 types (g, h, m, n) have no symmetry axes or planes. Therefore if we adopt by convention a sense of rotation around each face, e.g., clockwise rotation, each of these four types will give rise to two distinct subtypes leading now to 24 types of vertices, which may discriminate among enantiomeric fullerenes when partitioning their n vertices (Figure 8, substructural set II).

Finally, one can analyze the carbon atom types according to another topological pattern: instead of considering one vertex and its three neighboring vertices, one can consider an edge, i.e., a pair of adjacent vertices with their four nearest neighbors. In this case there can be a total of 80 types of topological patterns around an edge in a fullerene,

Table 4. Data for the 1812 C_{60} Cages and Their Duals

invariant X	no. of classes	information deficit (D_X)	$ C_\alpha _{\max}$
60-Vertex Cages			
substructural set I	1558	0.3040	5
substructural set II	1693	0.1370	4
substructural set III	1803	0.0099	2
$V(K, C_6, C_8, C_{10})$	1812	0	1
J	1812	0	1
I_D^K	1812	0	1
I_D^W	1812	0	1
Kf: Kirchhoff no.	1812	0	1
W: Wiener no.	116	5.1062	82
no. of spanning trees	1812	0	1
no. of Hamilt. circuits	656	1.7397	10
$\lambda_{\text{HOMO}} - \lambda_{\text{LUMO}}$	1805	0.0097	5
π -electron spectra	1812	0	1
$\lambda_{\text{HOMO}} + \lambda_{\text{LUMO}}$	1809	0.0033	2
E_π^{total}	1812	0	1
mean inertia	1812	0	1
32-Vertex Dual Graphs			
Kf: Kirchhoff no.	1812	0	1
J	1812	0	1
W: Wiener no.	31	7.2934	303
$\lambda_{\text{HOMO}} - \lambda_{\text{LUMO}}$	1808	0.0044	2
I_D^K	1812	0	1
I_D^W	1808	0.0044	2
E_π^{total}	1810	0.0022	2
$\lambda_{\text{HOMO}} + \lambda_{\text{LUMO}}$	1809	0.0033	2
mean inertia	1812	0	1
12-Vertex Reduced Duals			
Kf: Kirchhoff no.	1812	0	1
W: Wiener no.	34	6.8207	210
dist. distrib. vector	819	1.6562	21
J	1811	0.0011	2

and this number offers better discrimination when partitioning the $3n/2$ edges of $[n]$ fullerenes (Figure 7, substructural set III).

5. RESULTS FOR C_{60} FULLERENES

For the single-valued graph invariants of sections 2 and 3 we have computed the various consequent X -equivalence classes of 60-atom fullerenes. We report in Table 2 various features of these different X -equivalence classes, including the number of the classes and the maximum size of such a class. We also give the values for selected invariants for the two extremal cages shown in Table 5. The first structure, buckminsterfullerene, is the standard truncated icosahedron with a minimum of abutting pairs of pentagonal faces (namely no such pairs). The second structure is presumably the most strained and most "antiaromatic" (hence least stable) 60-atom fullerene with a maximum of abutting pairs of pentagonal faces (two halves of a regular dodecahedron separated by a "belt" of 20 hexagons). The corresponding reduced dual distance matrices are shown in Figures 5 and 6. Often these two C_{60} cages yield opposite extremes for the value of an invariant X , as indicated in Table 5 (see Scheme 1).

Finally for each given equivalence relation we report the corresponding *information deficit*, defined as

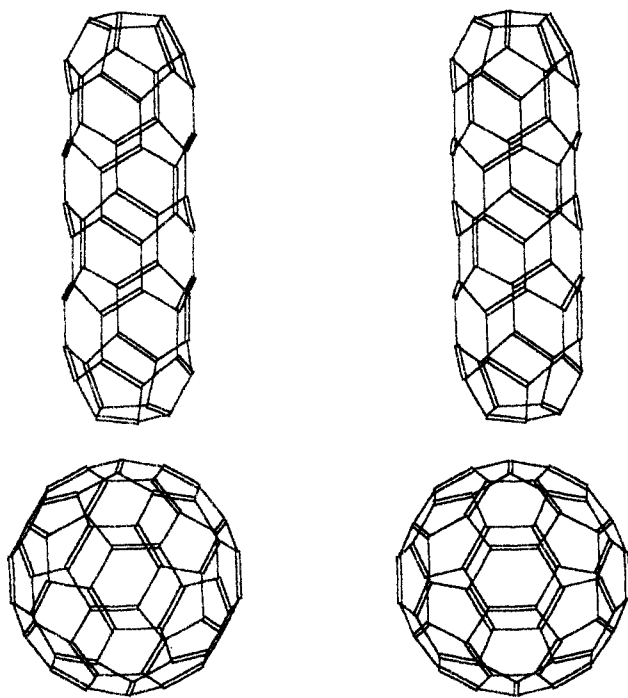
$$D_X = \log_2 N + \sum_{\alpha} \frac{|C_\alpha|}{N} \log_2 \frac{|C_\alpha|}{N} \quad (5.1)$$

where $|C_\alpha|$ is the order of the α th X -class, \log_2 is the

Table 5. Data for the Two Extremal C₆₀ Cages

C60, $p = q = 0$			C60, $p = 20, q = 10$		
60-Vertex Cages					
substructural set I	00 00 00 60 00 00 00 00 00 00		20 10 00 00 00 00 00 00 10 00		
	00 00 00 00 00 00 00 00 00 00		10 00 00 00 00 00 00 10 00		
substructural set II	12 00 00 00 00 00 00 00 00 00		00 00 00 00 00 00 00 02 10 00 10		
	00 00 00 00 00 00 20 00 00 00 00		00 00 00 00 00 00 00 00 00 00 00		
substructural set III	000 000 000 120 060 000 000 000		060 020 020 000 000 040 000 000		
	000 000 000 000 000 000		000 020 000 000 020 000		
$V(K, C_6, C_8, C_{10})$	(12 600, 83 160, 0, 59 760)		(16 501, 83 500, 95 380)		
Kirchhoff no. Kf	4588.3445	(min)	5118.2177		(max)
Wiener no. W	8340	(min)	8880		(max)
no. of spanning trees	375, 291, 866, 372, 898, 816, 000	(max)	270, 116, 969, 877, 332, 883, 360		(min)
reson energy/graphite	0.9619	(max)	0.9171		(min)
gap _{HOMO-LUMO}	0.7566	(max)	0.0000		(min)
32-Vertex Duals					
Kirchhoff no. Kf	470.6475	(min)	524.5240		(max)
I_D^K	4.9989	(max)	4.9931		(min)
Weiner no. W	1280	(neither ^a)	1337		(max)
I_D^W	5.0000	(max)	4.9852		(min)
Gap _{HOMO-LUMO}	1.3028	(max)	0.0000		(min)
HOMO + LUMO	0.0000	(max)	-2.1751		(min)
12-Vertex Reduced Duals					
Kirchhoff no. Kf	70.8517	(min)	74.2820		(neither ^b)
Weiner no. W	180	(min)	195		(neither ^c)

^a Minimum value: $W = 1256$. ^b Maximum value: $Kf = 84.4097$. ^c Maximum value: $W = 232$.

Scheme 1. Stereoview of the Two Extremal Cages with 60 Carbon Atoms

logarithm function to base 2, and

$$N = \sum_{\alpha} |C_{\alpha}| \quad (5.2)$$

Here in (5.1) the α -sum over X -classes gives the negative of the information (or entropy) content for the probability distribution

$$P_{\alpha} = |C_{\alpha}|/N \quad (5.3)$$

while the $\log_2 N$ -term in (5.1) gives the information (or entropy) content for the completely discerning probability

distribution with N nonzero probabilities all equal to $1/N$. Thus D_X is a measure of the information in the maximally discerning distribution compared against that of (5.3). That is, the nearer to zero the information deficit D_X is, the better (or more discerning on the average) is the X -equivalence relation.

Complete discrimination by an invariant X is indicated in boldface characters in Tables 4 and 6. It is seen that several invariants (indices) show complete discrimination of all 1812 60-atom fullerenes. The new multivalued substructural invariants tested here do not characterize uniquely all these fullerenes.

A comment is appropriate for the unique pair of C₆₀ isomers sharing the same J value for their duals: as seen in Figure 10, they have two identical hemispherical moieties connected in two ways differing by a rotation of 90° (the thick line delineates the two moieties). No other such type of degeneracy was found among higher fullerenes.

6. RESULTS FOR ISOLATED-PENTAGON FULLERENES

From Tables 4 and 6, it may be seen that the same invariants that discriminated completely the 1812 isomers of C₆₀ also show complete discrimination among the 558 preferable cages with 60–96 carbons. The multivalued substructural invariants do not characterize uniquely these cages, but for the experimentally realizable cages, the multivalued substructural invariants are likely to characterize them completely. A more detailed discussion is presented below.

7. DISCUSSION OF THE RESULTS

On examining Tables 4 and 6, it emerges that none of the substructural sets I–III is able to provide complete discrimination for the fullerene cages examined here. Among the single value invariants, the J index, the information-theoretic indices, and the Kirchhoff number have always been found

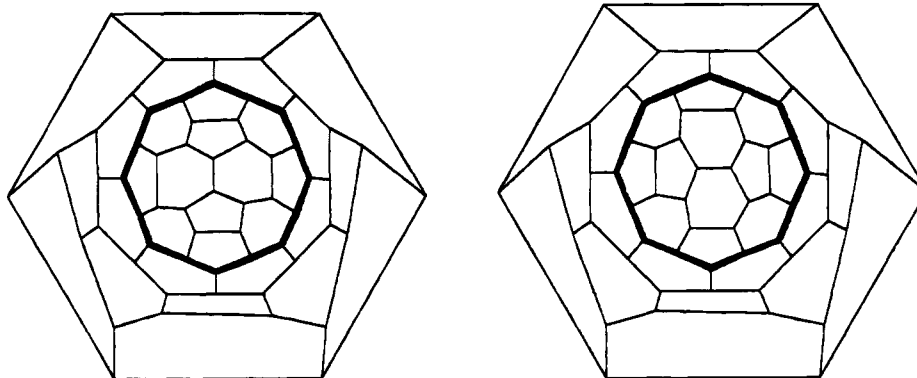


Figure 10. The unique pair of 60-vertex isomers sharing the same J index for their reduced duals (Schlegel diagrams).

Table 6. Invariants for the 558 Preferable Cages with $n = 60-96$ Carbon Atoms

invariant X	no. of classes	information deficit (D_X)	$ C_n _{\max}$
<i>n</i> -Vertex Cages			
substructural set 1	297	1.2312	11
substructural set 2	309	1.1644	9
substructural set 3	265	1.5126	16
$V(K, C_6, C_8, C_{10})$	558	0	1
no. K	558	0	1
no. C_6	558	0	1
no. C_{10}	558	0	1
Kf: Kirchhoff no.	558	0	1
J	558	0	1
I_D^K	558	0	1
I_D^W	558	0	1
W: Wiener no.	351	0.8730	6
no. of spanning trees	558	0	1
E_{π}^{total}	558	0	1
$\lambda_{\text{HOMO}} - \lambda_{\text{LUMO}}$	508	0.2017	9
$\lambda_{\text{HOMO}} + \lambda_{\text{LUMO}}$	558	0	1
mean inertia	558	0	1
<i>f</i> -Vertex Dual Graphs			
Kf: Kirchhoff no.	558	0	1
J	558	0	1
I_D^K	558	0	1
I_D^W	557	0.0036	2
W: Wiener no.	93	3.3864	29
E_{π}^{total}	554	0.0143	2
$\lambda_{\text{HOMO}} - \lambda_{\text{LUMO}}$	518	0.2043	16
E_{π}	536	0.0789	2
$\lambda_{\text{HOMO}} + \lambda_{\text{LUMO}}$	528	0.1075	2
mean inertia	558	0	1
12-Vertex Reduced Duals			
Kf	558	0	1

to provide full discrimination for the cages. However, among the dual graphs of the 558 preferable cages, one pair gives rise to degenerate I_D^W information-theoretic descriptors, and only I_D^K provides complete discrimination.

Other invariants that have zero information deficit, i.e., full discrimination ability, are the integer numbers of spanning trees and the multivalued invariant V (the vector of counts for Kekulé structures and conjugated circuits). Among the quantum-chemical invariants, the total π -electron energy (and some associated invariants such as the delocalization energy, and the ratio between resonance energies per electron in fullerenes and in graphite which is shown in Table 5) provides a complete discrimination for the 1812 C_{60} cages and the 558 preferable cages $C_{60}-C_{96}$. However, these invariants lead to degeneracies for their duals. On the other hand, the mean inertia is (together with Kf and I_D^K) an invariant which always provides complete discrimination at

all levels, i.e., the $n \times n$ matrix associated with the cage, the $f \times f$ matrix associated with the dual graph (there are $f = 2 + n/2$ faces), or the 12×12 matrix of the reduced dual graph.

The Hamiltonian circuit count (of interest in relation to IUPAC nomenclature) and the simple pairs $\{p, q\}$ or $\{A, B\}$ are quite incomplete in characterizing fullerenes. The highest degeneracy is associated with the Wiener number, both for the cages and for their duals or reduced duals. Without entering into further details, one should also mention that the partition⁵⁴ of carbon cages according to their 20 odd symmetry point groups is even less discriminating.

8. CONCLUSIONS

As the methods for obtaining and purifying fullerenes advance, the numbers of preferable isomeric structures for carbon cages with more than 76 atoms are bound to increase. Even carbon cages with abutting pentagons will probably be detected, when methods are found for obtaining fullerenes under suitable kinetic rather than thermodynamic control. As these cages have appreciable chemical reactivity, more and more derivatives are being prepared, by addition of hydrogen or halogen atoms, of carbenes, nitrenes, or their analogs, or by a variety of other chemical processes. Therefore a need exists for a firm basis on which one can base a systematic search and labeling for all possible isomeric cages with given numbers of carbon atoms, and in some cases with certain restrictions such as the absence of type A or B carbon atoms. Fullerene invariants examined in the present paper may assist in this task.

The reduced dual (12×12 matrix for any fullerene, no matter how many carbon atoms it contains) raises a challenge for invariants because of the reduction in information. Within the two classes of fullerenes examined by us, the Kirchhoff number, the associated index I_D^K , and the mean inertia achieve full discrimination even in this difficult case and *a fortiori* in the subsequent easier tasks.

The f -vertex duals of the fullerenes with n carbon atoms and $f = 2 + n/2$ faces (12 pentagons and $n/2 - 10$ hexagons) allow discrimination by a wider range of invariants such as topological indices J and I_D^K . Finally, for the $n \times n$ matrix corresponding to the cage itself, many invariants qualify for complete discrimination among constitutional isomers. The fullerenes are seen to provide a unique data base for sensitive testing of various graph-theoretical invariants for their discriminating power.

ACKNOWLEDGMENT

The support of the Welch Foundation of Houston, TX is gratefully acknowledged. Thanks are expressed to Dr. C. Ruecker for his helpful comments in refereeing this paper.

REFERENCES AND NOTES

- (1) *Fullerene Science and Technology*; Braun, T., Ed., Marcel Dekker Inc., New York, 1994.
- (2) *Buckminsterfullerenes*; Billups, W. E.; Ciufolini, M. A., Eds.; VCH Publishers: New York, 1993.
- (3) *Fullerenes. Synthesis, Properties, and Chemistry of Large Carbon Clusters*; Hammond, G. S.; Kerch, V. J., Eds.; ACS Symposium Series 481; American Chemical Society: Washington, DC, 1992.
- (4) *Novel Forms of Carbon*; Renschler, C. L.; Pouch, J. J.; Cox, D. M., Eds.; Proceedings of a Symposium held April 27–May 1, 1992; San Francisco, CA, Materials Research Society: Pittsburgh, PA (Materials Research Society Symposium Proceedings, 1992; Vol. 270).
- (5) Fowler, P. W.; Manolopoulos, D. E. *An Atlas of Fullerenes*; Oxford University Press: Oxford, 1994.
- (6) *Acc. Chem. Res.* Special Issue on Fullerenes, **1992**, 25, March issue.
- (7) *Carbon*; Kroto, H., Guest Ed.; Special Issue on Fullerenes, **1992**, 30, 1139–1285.
- (8) Renschler, C. L.; Pouch, J. J.; Cox, D. M., Eds., *Novel Forms of Carbon, Part. III: Fullerenes, Fullerides and Fulleroids*; Material Research Society Symposium Proceedings, Material Research Society: Pittsburgh, PA, 1992; Vol. 270, pp 115–314.
- (9) Proceedings of the First Croatian Meeting on Fullerenes: Graovac, A., Klein, D. J., Pisanski, T., Guest Eds.; *Fullerene Sci. Technol.* **1994**, 2, 333–497.
- (10) Kroto, H. W.; Walton, D. R. M. Postfullerene organic chemistry. In *Carbocyclic Cage Compounds. Chemistry and Applications*; Osawa, E.; Yonemitsu, O., Eds.; VCH Publishers: New York, 1992; pp 91–100.
- (11) Klein, D. J.; Schmalz, T. G. In *Quasicrystals, Networks and Molecules with Fivefold Symmetry*; Hargittai, I., Ed.; VCH Publishers: New York, 1990; p 239. Chapter 14: Buckminsterfullerene. Part A: Introduction; see also subsequent chapters (15–17, parts B–D).
- (12) Dresselhaus, M. S.; Dresselhaus, G.; Saito, R.; Eklund, P. C. C_{60} -Related balls and fibers. In *Elementary Excitations in Solids*; Birman, J. L., Ebenene, C., Wallis, R. F., Eds.; Elsevier, Amsterdam, 1992; pp 387–417.
- (13) Weltner, W., Jr.; Van, Zee, R. J. Carbon molecules, ions, and clusters. *Chem. Rev.* **1989**, 89, 1713–1747.
- (14) Kroto, H. W.; Allaf, W.; Balm, S. P. C_{60} -Buckminsterfullerene. *Chem. Rev.* **1991**, 91, 1213–1235.
- (15) Sokolov, V. I.; Stankevich, I. V. The fullerenes — new allotropic forms of carbon: molecular and electronic structure, and chemical properties. *Russ. Chem. Rev.* **1993**, 62, 419–435.
- (16) Dresselhaus, M. S.; Dresselhaus, G.; Eklund, P. C. Fullerenes. *J. Mater. Res.* **1993**, 8, 2054–2097.
- (17) Hebard, A. F. Buckminsterfullerene. In *Annu. Rev. Mater. Sci.* **1993**, 23, 159–191.
- (18) Kroto, H. W. C_{60} : Buckminsterfullerene, the celestial sphere that fell to earth. *Angew. Chem., Int. Ed. Engl.* **1992**, 31, 111–129.
- (19) Hirsch, A. The chemistry of the fullerenes: an overview. *Angew. Chem., Int. Ed. Engl.* **1993**, 32, 1138–1141.
- (20) Taylor, R.; Walton, D. R. M. The chemistry of fullerenes. *Nature* **1993**, 363, 685–693.
- (21) Braun, T. The epidemic spread of fullerene research. *Angew. Chem., Int. Ed. Engl.* **1992**, 31, 588–589.
- (22) Liu, X.; Klein, D. J.; Schmalz, T. G.; Seitz, W. A. Generation of carbon-cage polyhedra. *J. Comput. Chem.* **1991**, 12, 1252–1259.
- (23) Manolopoulos, D. E.; May, J. C.; Down, S. E. Comment on "Favorable structures for higher fullerenes". *Chem. Phys. Lett.* **1992**, 192, 330.
- (24) Manolopoulos, D. E. Comment on "Favorable structures for higher-fullerenes". *Chem. Phys. Lett.* **1992**, 192, 331. Liu, X.; Schmalz, T. G.; Klein, D. J. Reply to comment on "Favorable structures for higher fullerenes". *Chem. Phys. Lett.* **1992**, 192, 331.
- (25) Manolopoulos, D. E.; Fowler, P. W. Molecular graphs, point groups, and fullerenes. *J. Chem. Phys.* **1992**, 96, 7603–7614.
- (26) Pope, C. J.; Marr, J. A.; Howard, J. B. Chemistry of fullerenes C_{60} and C_{70} formation in flames. *J. Phys. Chem.* **1993**, 97, 11001–11013. See, also: Raghavachari, K.; McMichael-Rolfing, C. Imperfect fullerene structures: isomers of C_{60} . *J. Phys. Chem.* **1992**, 96, 2463–2466.
- (27) Klein, D. J.; Liu, X. Elemental carbon isomerism. *Int. J. Quantum Chem.: Quantum Chem. Symp.* **1994**, 28, 501–523.
- (28) Kroto, H. W. The stability of the fullerenes C_n with $n = 24, 28, 32, 36, 50, 60$, and 70. *Nature* **1987**, 329, 529–531.
- (29) Manolopoulos, D. E.; May, J. C.; Down, S. E. Theoretical studies of the fullerenes: C_{34} to C_{70} . *Chem. Phys. Lett.* **1991**, 181, 105–111.
- (30) Fowler, P. W.; Manolopoulos, D. E. Magic numbers and stable structures for fullerenes, fullerides and fullerenium ions. *Nature* **1992**, 355, 428–430.
- (31) Taylor, R. Rationalization of the most stable isomer of a fullerene C_n . *J. Chem. Soc., Perkin Trans. 2* **1992**, 3–4.
- (32) Liu, X.; Klein, D. J.; Schmalz, T. G. Preferable fullerenes and Clar-sextet cages. *Fullerene Sci. Technol.* **1994**, 2, 405–422.
- (33) Austin, S. J.; Fowler, P. W.; Orlandi, G.; Manolopoulos, D. E.; Zerbetto, F. Relative stability of C_{76} isomers. A numerical test of the fullerene isolated pentagon rule. *Chem. Phys. Lett.* **1994**, 226, 219–225.
- (34) Murry, R. L.; Scuseria, G. E. Theoretical Study of C_{90} and C_{96} fullerene isomers. *J. Am. Chem. Soc.* **1994**, 98, 4212–4214.
- (35) Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **1947**, 69, 17–20.
- (36) Klein, D. J.; Randić, M. Resistance distance. *J. Math. Chem.* **1993**, 12, 81–95.
- (37) Shapiro, L. W. An electrical lemma. *Math. Magazine*, **1987**, 60, 36–38.
- (38) Balaban, A. T. Highly discriminating distance-based topological index. *Chem. Phys. Lett.* **1982**, 80, 399–404.
- (39) Balaban, A. T. Topological indices based on topological distances in molecular graphs. *Pure Appl. Chem.*, **1983**, 55, 199–206.
- (40) Balaban, A. T.; Motoc, I.; Bonchev, D.; Mekenyan, O. Topological indices for structure–activity correlations. In *Steric Effects in Drug Design*; Charton, M., Motoc, I., Eds.; Springer: Berlin. *Top. Curr. Chem.* **1983**, 114, 21–55.
- (41) Trinajstić, N. Chemical Graph Theory; CRC Press: Boca Raton, FL, 1983; Chapter 10, pp 225–274.
- (42) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity Studies*. Research Studies Press, Letchworth, 1986. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
- (43) Balaban, A. T. Using real numbers as vertex invariants for third generation topological indexes. *J. Chem. Inf. Comput. Sci.* **1992**, 32, 23–28.
- (44) Balaban, A. T. Chemical graphs: looking back and glimpsing ahead. *J. Chem. Inf. Comput. Sci.*, in press.
- (45) Bonchev, D. *Information Theoretic Indices for Characterization of Chemical Structure*; Research Studies Press-Wiley: Chichester, 1993.
- (46) Bonchev, D.; Trinajstić, N. Information theory, distance matrices, and molecular branching. *J. Chem. Phys.* **1977**, 67, 4517–4533.
- (47) Liu, X.; Klein, D. J. The graph isomorphism problem. *J. Comput. Chem.* **1991**, 12, 1243–1251.
- (48) Herndon, W. C.; Ellzey, M. L., Jr. Resonance Theory. V. Resonance Energies of Benzenoid and Nonbenzenoid π Systems. *J. Am. Chem. Soc.* **1974**, 96, 6631–6642.
- (49) Randić, M. On the characterization of fullerenes. *Fullerene Sci. Technol.* **1994**, 2, 427–444.
- (50) Schmalz, T. G.; Seitz, W. A.; Klein, D. J.; Hite, G. E. Elemental carbon cages. *J. Am. Chem. Soc.* **1988**, 110, 1113–1127. Liu, X.; Schmalz, T. G.; Klein, D. J. Favorable structures for higher fullerenes. *Chem. Phys. Lett.* **1992**, 188, 550–554. Klein, D. J.; Liu, X. Theorems for carbon cages. *J. Math. Chem.* **1992**, 11, 199–205.
- (51) Babic, D.; Balaban, A. T.; Klein, D. J. Nomenclature and coding of fullerenes. *J. Chem. Inf. Comput. Sci.*, in press.
- (52) Manolopoulos, D. E.; Fowler, P. W. A fullerene without a spiral. *Chem. Phys. Lett.* **1993**, 224, 1–7 and references therein.
- (53) Whitney, H. Congruent graphs and the connectivity of graphs. *Amer. J. Math.* **1932**, 54, 150–168.
- (54) Babic, D.; Klein, D. J.; Sah, C. H. Symmetry of fullerenes. *Chem. Phys. Lett.* **1993**, 211, 325–241.

CI950011P