

Complexity, spanning trees and relative energies in fullerene isomers¹

Patrick W. Fowler

School of Chemistry, University of Exeter, Stocker Road, Exeter EX4 4QD, UK

(Received August 20, 2002)

Abstract

Within a set of isomers, the number of spanning trees of a fullerene graph (the *complexity*) correlates negatively with the number of pentagon adjacencies, and hence correlates positively with relative stability. The most stable isomers of C_{40} , C_{60} and C_{70} each have the greatest complexity within their respective sets of 40, 1812 and 8149 candidates. A similar sorting pattern, but with less marked separation of isomers, is evident for Randić's modified definition of complexity, which is based on augmented vertex degree.

1. Introduction

Every connected graph G has at least one *spanning tree*, *i.e.* a subgraph that connects all the vertices of G but has no cycles. Spanning trees of weighted and unweighted graphs have applications in operational research, in design of communications and distribution networks [1], and in classical physics in the computation of currents in complex electrical circuits [2]. They appear in chemistry as intermediaries

¹Dedicated to Professor Horst Sachs on the occasion of his 75th birthday.

in McWeeny's formulation of the Hückel-London model for ring currents in π systems [3]. Well known algorithms exist for finding a spanning tree, or one with specified minimal properties, and also for counting the total number of spanning trees associated with a given graph, a quantity known as the *complexity* of the graph, $C(G)$ [4-7]. Several authors have dealt with this graph-invariant property in the specific case of *fullerene* graphs, *i.e.* those cubic polyhedral graphs on n vertices comprising 12 pentagonal and $(n/2 - 10)$ hexagonal faces [8]. Analytical evaluations of complexity have been carried out for highly symmetric cases C_{60} [9] and C_{140} [10], and computed values have been published for several fullerenes in the range C_{20} to C_{90} [11]. The values are unwieldy integers, with the $\sim n/3$ digits expected from the degree of these cubic graphs, which have high discriminating power, and are capable of uniquely labelling all 1812 isomers of C_{60} , for example [12]. The target of the current investigation is a different one: to check for the existence of a link between the mathematical complexity of a fullerene and its physical energetics.

2. Method

The Matrix-Tree theorem, implicit in the work of Kirchhoff in 1847 [2], gives the number of spanning trees $C(G)$ as the value of (every) cofactor of the Laplacian matrix of the graph, $L(G)$. The Laplacian is

$$L(G) = D(G) - A(G) \quad (1)$$

where A is the conventional adjacency matrix with $A_{ij} = 1$ if i and j are vertices joined by an edge, $A_{ij} = 0$ otherwise, and D is the (diagonal) matrix of vertex degrees, with $D_{ij} = 0$ if $i \neq j$, $D_{ii} = \sum_j A_{ij} = d_i$, the degree or valency of the i th vertex ($= 3$ for fullerenes). The characteristic polynomial of $L(G)$ is a generating function for the numbers of spanning trees of subgraphs of G , and expressions for the first few of its coefficients in terms of structural features have been given for fullerenes [11]. For connected graphs, $C(G)$ is readily evaluated from the spectrum of the Laplacian [13, 7], as

$$C(G) = \left(\frac{1}{v_G} \right) \prod_{i>1}^{v_G} \lambda_i^L, \quad (2)$$

where v_G is the number of vertices of G and $\{\lambda_i^L\}$ are the eigenvalues of $L(G)$ arranged in non-decreasing order, and the product omits the zero eigenvalue of $L(G)$, which by construction is singular. In a regular graph with all vertices of degree r and $v_G = n$, the spectra of $L(G)$ and $A(G)$ are related by

$$\lambda_i^L = r - \lambda_i^A \quad (3)$$

with,

$$0 = \lambda_1^L < \lambda_2^L \leq \dots \leq \lambda_n^L \leq 2r \quad (4)$$

and

$$r = \lambda_1^A > \lambda_2^A \geq \dots \geq \lambda_n^A \geq -r \quad (5)$$

In a fullerene, C_n , which is a regular graph with $r = 3$, a useful formula for the complexity is [7, 13]

$$C(G) = \left(\frac{1}{n}\right) \prod_{i>1}^n (3 - \lambda_i^A). \quad (6)$$

More-efficient algorithms for evaluation of $C(G)$ exist, but the form (6) has the attractions that it requires only a trivial modification of a standard Hückel molecular orbital program and that it has good numerical stability properties. In most previous applications, the emphasis has been on exact evaluation of the very long integers $C(G)$. To achieve this for large n by taking products of eigenvalues of the Laplacian, or by any other real-arithmetic formula, requires special, and eventually prohibitive, attention to precision [6, 7, 11, 9]. The number of multiplications for the fullerene C_n could be reduced further by taking advantage of the fact that the complexity of a planar graph is equal to that of its dual [14], and hence applying the Laplacian formula (2) to the spectrum of the (non-regular) deltahedral dual of the fullerene, a graph with 12 vertices of degree 5 and $(n/2 - 10)$ of degree 6 and hence $v_G = (n/2 + 2)$. Quadruple-precision FORTRAN was used here, although the highest absolute precision is not in fact necessary, as all the isomers considered here can be distinguished by comparison within the 8 or 9 most significant digits, so that double-precision real-number approximations to $C(G)$ are sufficient for the present purpose. The complexity for all fullerenes of reasonable size is easily determined to greater precision than any experimental quantity with which it is likely to be correlated.

3. Results

Adjacency information for fullerenes is coded by the face spiral [8, 15] construction. Lists of spirals were used to generate $A(G)$, and hence $C(G)$ for three sets of candidates: the 40, 1812 and 8149 classical fullerene isomers of C_{40} , C_{60} and C_{70} .

The identity of the isomer of lowest total energy in each set is well known [8]. Within the C_{40} set, isomer 40 : 38, which has D_2 symmetry and the topology of a tennis ball, with a seam of 10 hexagons surrounding two isolated crescents of 6 pentagons each, is consistently predicted by quantum-mechanical methods to have the lowest total energy. Within the C_{60} and C_{70} sets, the unique isolated-pentagon isomers (60 : 1812 and 70 : 8149, approximating the football and rugby ball, respectively) have the lowest relative energies. These three most stable isomers are also those with the unique highest complexity within each set. Significantly, in each case a low- $C(G)$ extreme is found for energetically poorest isomer, the first in the spiral order of isomers, the cylinder with 10 pentagon adjacencies in each of its two hemi-dodecahedral caps. The ranges spanned by the computed values are from the minima

$$C(40 : 1) = 37\,483\,980\,220\,840 = 2^3 \times 5 \times 751^2 \times 1289^2 \sim 3.748 \times 10^{13};$$

$$\begin{aligned} C(60 : 1) &= 270\,116\,969\,877\,332\,883\,360 = 2^5 \times 3^8 \times 5 \times 11^2 \times 31^2 \times 47\,041^2 \\ &\sim 2.701 \times 10^{20}; \end{aligned}$$

$$\begin{aligned} C(70 : 1) &= 725\,097\,819\,324\,933\,982\,208\,000 = 2^{14} \times 5^3 \times 19^2 \times 101^2 \times 149^2 \times 2081^2 \\ &\sim 7.251 \times 10^{23}, \end{aligned}$$

to the maxima

$$\begin{aligned} C(40 : 38) &= 41\,404\,929\,874\,560 = 2^7 \times 3^5 \times 5 \times 7 \times 43 \times 347 \times 2549 \\ &\sim 4.140 \times 10^{13}; \end{aligned}$$

$$\begin{aligned} C(60 : 1812) &= 375\,291\,866\,372\,898\,816\,000 = 2^{25} \times 3^4 \times 5^3 \times 11^5 \times 19^3 \\ &\sim 3.753 \times 10^{20}; \end{aligned}$$

$$\begin{aligned} C(70 : 8149) &= 1\,136\,544\,737\,068\,261\,950\,000\,000 = 2^7 \times 3 \times 5^8 \times 11^2 \times 13 \times 59^2 \times 37199^2 \\ &\sim 1.137 \times 10^{24}. \end{aligned}$$

The values $C(60 : 1812)$ and $C(70 : 8149)$ agree with previously published values [6, 7, 9, 12].

As an index of relative total $(\sigma + \pi)$ energy, the value of N_p , the number of pentagon-pentagon edges of the fullerene isomer, is a good choice. It is easy to determine, and has been shown to correlate well with quantum mechanically calculated isomer energies; the penalty of $\sim 1\text{eV}$ per adjacency is the driving force behind the isolated-pentagon rule for stable fullerenes [16]. As Figure 1 shows, N_p also gives a good prediction of the different values of fullerene-graph complexities: all three fullerene isomer sets, totalling 10,001 molecular graphs, give scatterplots of $C(G)$ vs. $N_p(G)$, each with a well defined envelope pointing to a maximum $C(G)$ for minimum N_p .

In view of the equality of complexity between a planar graph and its dual [14], the plots in Figure 1 can also be re-interpreted simply by re-labelling the axes as correlations between complexity and the number of adjacent five-valent vertices in the deltahedral fullerene duals. The chemical systems modelled by fullerene duals are (hypothetical) closo-boranes $B_nH_n^{2-}$; those with more than 12 boron vertices remain only a theoretical possibility [17, 18]. According to Figure 1, the trend of maximum complexity for minimum adjacency of low-coordinate boron vertices, and hence for energetic stability, carries over from the fullerenes to the boranes.

It is interesting to compare the performance of the standard definition of complexity with others that have been proposed. In another paper in this special issue, Mallion and Trinajstić [19] consider refinements of the raw count of spanning trees as a measure of the complexity or ‘intricacy’ of a graph. Their *spanning-tree density* $\text{STD}(G)$ measures the probability that a randomly chosen set of $v_G - 1$ edges form the total of e_G edges of the graph G will be a spanning tree, i.e.

$$\text{STD}(G) = C(G) \binom{e_G}{v_G - 1}^{-1}. \quad (7)$$

Within isomer sets $\text{STD}(G)$ is of course strictly proportional to $C(G)$, and so all the scatterplots on the left of Figure 1 apply equally well to $\text{STD}(G)$.

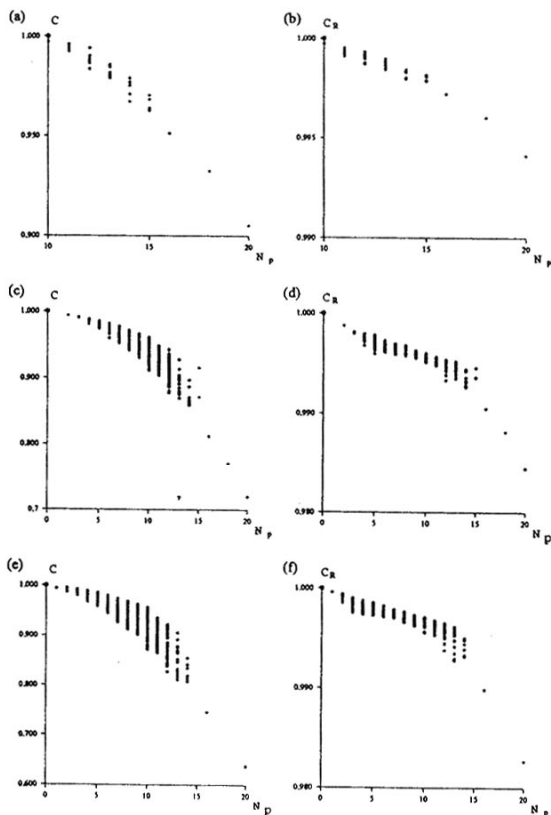


Figure 1: Correlation of fullerene complexity and pentagon adjacency. The left-hand panels show scatterplots of the relative complexity, defined as the ratio of $C(G)$ to the complexity of the most complex isomer, against N_p , the number of pentagon fusions. The right-hand panels show the equivalent plots for the C_R , again normalised to the most Randić-complex isomer. The three rows refer to (a),(b) the 40 fullerene isomers of C_{40} , (c),(d) the 1812 fullerene isomers of C_{60} , (e),(f) the 8149 fullerene isomers of C_{70} .

In a different approach, Randić suggests that it may be useful to define complexity through the *augmented valences* [20, 21]. By taking each vertex in turn as a starting point, the valencies of all its neighbouring vertices in successive shells are accumulated, weighted by a decreasing function of their distance from the starting vertex. Averaging over all possible starting vertices yields

$$C_R(G) = \left(\frac{1}{n}\right) \sum_{i=1}^n \sum_{j=1}^n d_i/2^{d_{ij}} \quad (8)$$

where d_i is the degree of vertex i and d_{ij} is the distance in edges between vertices i and j . The subscript R for Randić is used here to distinguish this quantity from the usual C .

As Figure 1 shows, C_R and C give broadly the same trends with pentagon adjacency. Maximum C_R is found for an isomer with minimum N_p in all three cases though, for C_{40} , the Randić complexity favours the less stable [16] of the two isomers that have $N_p = 10$, with

$$C_R(40 : 39) = 695.156, \quad C_R(40 : 38) = 694.969$$

The isolated-pentagon isomers have

$$C_R(60 : 1812) = 1110.586, \quad C_R(70 : 8149) = 1317.656.$$

Minimum values of C_R are again found for the isomer of each set that has the most pentagon adjacencies:

$$C_R(40 : 1) = 691.055, \quad C_R(60 : 1) = 1093.257, \quad C_R(70 : 1) = 1294.823.$$

$C_R(G)$ yields a narrower range of relative values than does $C(G)$ for the fullerene sets examined, and it seems that $C(G)$ is superior from the point of view of isomer discrimination.

4. Discussion

It is becoming clear that a variety of invariants show a pattern of correlation with the computed total quantum-mechanical energies of fullerenes, reinforcing the chemical intuition that the energy of a carbon cage has a significant strain/curvature

component from the adaptation of the carbon atoms to the non-ideal bond angles imposed by the wrapping of the graphite sheet onto the sphere. Wiener index, Balaban index [22], resistance distance [23], smallest eigenvalue [24], and now complexity all show a qualitative tendency to distinguish isomers of very high and very low energies, which may be a pointer to underlying family relationships between the mathematical invariants themselves. From a practical point of view, these invariants will not replace the much simpler count of pentagon adjacencies, which gives a first qualitative picture of the total energy, but they may yet be useful in constructing combinations to pick out especially stable candidates from within the isolated-pentagon set.

References

- [1] See any one of a number of teaching texts, e.g.
 - a) B. Kollman, R.C. Busby, S. Ross, *Discrete Mathematical Structures*, Prentice-Hall, Upper Saddle River NJ, USA, 1996;
 - b) R. Johnsonbaugh, *Discrete Mathematics*, Prentice-Hall, Upper Saddle River NJ, USA, 1997;
 - c) E.G. Goodaire, M.M. Parmenter, *Discrete Mathematics with Graph Theory*, Prentice-Hall, Upper Saddle River NJ, USA, 1998.
- [2] G. Kirchhoff, Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird, *Ann. Phys. Chem.* **72** (1847) 497–508. Extracts and a commentary on this paper are given in N.L. Biggs, E.K. Lloyd, and R.J. Wilson, *Graph Theory 1736–1936*, Clarendon Press, Oxford (1998), Chapter 8. A complete English translation is given in J.B. O’Toole, On the solution of the equations obtained from the investigation of the linear distribution of galvanic currents, *I.R.E. Trans. Circuit Theory* **CT5** (1958) 4–7.
- [3] R. McWeeny, Ring currents and proton magnetic resonance in aromatic molecules, *Molec. Phys.* **1** (1958) 311–321.
- [4] R.C. Prim, Shortest connection networks and some generalisations, *Bell Systems Tech. J.* **36** (1957) 1389–1401.
- [5] J.B. Kruskal, Jr., On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Amer. Math. Soc.* **7** (1956) 48–50.

- [6] T.N.J. Brown, R.B. Mallion, P. Pollak and A. Roth, Some methods for counting the spanning trees in labelled molecular graphs, examined in relation to certain fullerenes, *Discr. Appl. Math.* **67** (1996) 51–66 (and erratum **75** (1997) 199–200).
- [7] Z. Mihalić and N. Trinajstić, On the number of spanning trees in fullerenes, *Fullerene Sci. Technol.* **2** (1994) 89–95.
- [8] P.W. Fowler and D.E. Manolopoulos, *An Atlas of Fullerenes*, Clarendon Press, Oxford (1995).
- [9] T.J.N. Brown, R.B. Mallion, P. Pollak, B.R.M. de Castro and J.A.N.F. Gomes, The number of spanning trees in buckminsterfullerene, *J. Comput. Chem.* **12** (1991) 1118–1124.
- [10] C.W. Haigh, Calculation of the complexity of a chiral C_{140} fullerene, *MATCH Commun. Math. Comput. Chem.* **33** (1996) 139–146. The complexity of $I C_{140}$ is quoted as the 50-digit integer

$$18\,693\,322\,896\,120\,604\,727\,745\,322\,410\,126\,395\,778\,265\,745\,166\,336.$$
This number has the factorisation: $2^{12} \times 3^4 \times 29^3 \times 31^3 \times 389^3 \times 503^4 \times 1831^5$.
- [11] K. Balasubramanian, Computation of spanning tree generators of fullerenes, *Molec. Phys.* **85** (1995) 335–348.
- [12] A.T. Balaban, X. Liu, D.J. Klein, D. Babic, T.G. Schmalz, W.A. Seitz, and M. Randic, Graph invariants for fullerenes, *J. Chem. Inf. Comput. Sci.*, **35** (1995) 396–404.
- [13] D. Cvetković, M. Doob and H. Sachs, *Spectra of Graphs: Theory and Application*, Academic Press, New York, 1980, proposition 1.3, p. 39.
- [14] N.L. Biggs, *Algebraic Graph Theory*, Cambridge University Press, Cambridge, 1974, p. 38.
- [15] D.E. Manolopoulos, J.C. May, and S.E. Down, Theoretical studies of the fullerenes- C_{34} to C_{70} , *Chem. Phys. Lett.*, **181** (1991) 105–111.
- [16] E. Albertazzi, C. Domene, P.W. Fowler, T. Heine, G. Seifert, C. Van Alsenoy, and F. Zerbetto, Pentagon adjacency as a determinant of fullerene stability, *Phys. Chem. Chem. Phys.*, **1** (1999) 2913–2918.
- [17] J. Bicerano, D.S. Marynick and W.N. Lipscomb, Large closo boron hydrides, *Inorg. Chem.* **17** (1978) 2041–2042; Molecular orbital studies on large closo boron hydrides, *Inorg. Chem.* **17** (1978) 3443–3448.

- [18] P.W. Fowler, The pairing principle in tensor surface harmonic theory: electron counts of large closo-boranes, *Polyhedron*, **4** (1985) 2051–2057.
- [19] R.B. Mallion and N. Trinajstić, Reciprocal spanning-tree density: a new index for characterising the intricacy of a (poly)cyclic molecular graph, *MATCH Commun. Math. Comput. Chem.*, **48** (2003), this issue.
- [20] M. Randić, On complexity of transitive graphs representing degenerate rearrangements, *Croat. Chem. Acta*, **74** (2001) 683–705.
- [21] M. Randić and D. Plavšić, On the concept of molecular complexity, *Croat. Chem. Acta*, **75** (2002) 107–116.
- [22] P.W. Fowler, G. Caporossi and P. Hansen, Distance matrices, Wiener indices and related invariants of fullerenes, *J. Phys. Chem. A* **105** (2001) 6232–6242.
- [23] P.W. Fowler, Resistance distances in fullerene graphs, *Croat. Chem. Acta* **75** (2002) 401–408.
- [24] P.W. Fowler, P. Hansen and D. Stevanović, A note on the smallest eigenvalue of fullerenes, *MATCH Commun. Math. Comput. Chem.*, **48** (2003), this issue.