

GRAPH THEORY AND MOLECULAR ORBITALS—X

THE NUMBER OF KEKULÉ STRUCTURES AND THE THERMODYNAMIC STABILITY OF CONJUGATED SYSTEMS¹

I. GUTMAN, N. TRINAJSTIĆ* and C. F. WILCOX, JR.††

Institute "Ruder Bošković" P.O.B.1016, 41001 Zagreb, Croatia, Yugoslavia

Abstract—The dependence of Hückel π -electron energies, E_n , on the basic graph theoretical parameters N (the number of vertices), ν (the number of edges) and ASC (the algebraic structure count) is explored. The form with the ASC enters E_n is established and an equation for E_n is developed. It is shown how the early and apparent success of the (resonance) theory rested on the fortunate fact that all Kekulé structures for benzenoid hydrocarbons and acyclic polyenes have the same parity. The significance of ASC in determining chemical stability and reactivity is discussed briefly.

One of the basic postulates of the resonance theory description of conjugated molecules is that their thermodynamic stability is an increasing monotonic function of the number of Kekulé structures that can be drawn for each compound.² This postulate has received support both from the approximate application of rigorous π -electron valence bond theory³ as well as from the broad experience of many organic chemists.² Largely because of the weight of accumulated empirical evidence the postulate continues to be accepted in spite of its known theoretical limitations.⁴

The aim of the present work was to investigate the validity of the resonance theory postulate using a graph-theoretical approach developed in the preceding series.^{1,5-12} It appears that only a partial justification can be obtained and then only under restricted conditions. The nature of these restrictions is important.

The present discussion will be limited to alternant hydrocarbons but the conclusions can be readily transferred to heterocyclic analogs.¹³ The extension to nonalternants is also possible^{8,11} but implies additional difficulties⁶ and will not be considered here.

Although the conceptual background of Hückel molecular-orbital theory¹⁴ is completely different from valence bond theory,¹⁵ a close connection between Hückel and resonance theory has been established.¹⁶ This connection was recently analysed systematically using graph theory.¹¹

Derivation. Let K^+ and K^- be the number of even and odd Kekulé structures, respectively ($K^+ + K^- = K =$ total number of Kekulé structures), as defined previously.^{5,11,16,17} Due to the original definition of Dewar and Longuet-Higgins¹⁶ two Kekulé structures are of the same (different) parity if one can be obtained from the other by cyclic transposition of odd (even) number of double bonds. A simple procedure for enumerating K^+ and K^- has been given.^{5,12} Then:¹⁷

$$\det A = (-1)^{N/2}(K^+ - K^-)^2 \quad (1)$$

where A is the adjacency matrix of the molecular graph^{5,18} and N is the number of vertices. In order to emphasize the distinction between the total number of Kekulé structures, K , and the difference $|K^+ - K^-|$, that appears above, the latter was called the ASC = Algebraic Structure Count.¹⁷ The same quantity has also been called the CSC = Corrected Structure Count.¹⁹

If x_j ($j = 1, 2, \dots, N$) are the eigenvalues of A ("the graph spectrum"),^{5,18}

$$(\text{ASC})^2 = \prod_{j=1}^N |x_j| \quad (2)$$

and

$$2 \ln (\text{ASC}) = \sum_{j=1}^N \ln |x_j| \quad (3)$$

assuming $\text{ASC} \neq 0$, i.e. that there is no zero in the graph spectrum.⁶ The case of $\text{ASC} = 0$ is treated separately in Appendix 1.

For the interval²⁰ $(0, 3)$, $\ln x$ can be approximated by a finite polynomial $U_m(x) = \sum_{i=0}^m (-1)^{i+1} u_i x^i$. When this polynomial is substituted into Eq (3), one obtains simply:

$$2 \ln (\text{ASC}) \approx \sum_{i=0}^m (-1)^{i+1} u_i W_i \quad (4)$$

*Also of Cornell University, Department of Chemistry, Ithaca, N.Y. 14850, U.S.A.

†National Academy of Science Fellow 1974.

§The late Professor Coulson in 1949 was first to show that the elements of the spectrum of a graph associated with a conjugated molecule lie in the range -3 to $+3$. However, this is just an example of a more general proposition that the eigenvalues x_i of a simple graph satisfy $\rho_{\max} > x_i > -\rho_{\max}$ of which the graph theoretical proof was given by L. Collatz and U. Sinogowitz, *Abh. Math. Sem. Univ. Hamburg* **21**, 63 (1957).

where

$$W_t = \sum_{i=1}^N |x_i|^t \quad (5)$$

The advantage of Eq (4) is that W_t 's are easily related to molecular topology.^{7†} Thus, it is well known^{3,21} that

$$W_0 = N \quad (6)$$

$$W_2 = 2\nu \quad (7)$$

where ν is the number of edges in the graph. Analogous but more complicated expressions for W_4 and W_6 have been derived,⁷ e.g.:

$$W_4 = 8n_4 - 2\nu + 2 \sum_{p=1}^N D_p^2 \quad (8)$$

where n_4 is the number of 4-membered rings in the molecule and D_p the degree (valency)²² of the p -th vertex.

It is easily seen that for alternant hydrocarbons W_t is just the total π -electron energy E_π (in β units).^{9,23} In Appendix 2 it is shown that

$$W_3 \approx \frac{2E_\pi}{N} \left(3\nu - \frac{E_\pi^2}{N} \right) \quad (9)$$

A common property of all of the W_t 's for $t \geq 4$ is that they depend on the number and size of the rings contained in the molecule (Eq 8). Ring-size effects, which have been discussed elsewhere,^{7,9,24} will not be considered here since they unduly complicate the resulting expressions. Termination of Eq (4) at the cubic term should not produce any serious error since it has already been demonstrated²⁵ that an expression for E_π in terms of only the topological parameters N , ν and ASC (equivalent to termination at the square terms) is in error by only 5%. This is an upper limit to the possible error since no attempt was made to make a least square fit as is implied in Eq (4). Moreover, inclusion of the extra degree of freedom in Eq (4) must reduce the error still further. A demonstration of this point will be made after the final expression is developed.

The necessary conditions for U_m to be a good approximation for $\ln x$ are (c.f. Ref 9):

$$u_1 > 0 \quad (10)$$

$$u_1 < u_2 < \dots < u_m \quad (11)$$

The numerical values of u_t 's are irrelevant for further discussion and relations (10) and (11) will suffice. Combining Eqs (4)–(9) one obtains

$$aE_\pi - bE_\pi^3 = c \quad (12)$$

where

$$a = u_1 + 6u_3\nu/N \quad (12a)$$

$$b = 2u_3/N^2 \quad (12b)$$

$$c = 2 \ln(\text{ACS}) + u_0N + 2u_2\nu \quad (12c)$$

Because of Eqs (10) and (11)

$$b \ll a \quad (13)$$

and therefore a good approximate solution of Eq (12) is

$$E \approx \frac{c}{a} + \frac{b}{a} \left(\frac{c}{a} \right)^3 \quad (14)$$

A topological formula is thus obtained showing the dependence of E_π on the number of carbon atoms (N), carbon-carbon bonds (ν) and the ASC of a conjugated hydrocarbon. Moreover, the ASC and *not the number of Kekulé structures* plays an important role in determining E_π , and thus the resonance theory postulate cannot be confirmed for the general case.

The primary focus of this work is to show the mathematical form with which the ASC enters E_π rather than to develop another approximate expression for E_π . However, it is instructive to explore the quality of fit of Eq (4) and (14) if these equations were exact, both would yield identical values of u_0 , u_1 , u_2 and u_3 for a given set of π -networks. As is well known, the least squares fitting of data which fluctuates from the fitting equation yields equation parameters that depend on the choice of independent variable selected. For an arbitrary set of 30 π -molecules (10 branched and linear acyclics, 10 benzenoid aromatics and 10 aromatics containing a cyclobutadiene ring) fitting of the quantity $\ln(\text{ASC})^2$ as shown in Eq (4) gave the parameters listed below with a standard deviation of 0.041 for the range of $\ln(\text{ASC})^2$ from 0.000 to 4.395. With the same set of 30 molecules the π -electron energy expressed by Eq (14) gave the parameters shown below and a standard deviation in E_π of 0.092 for the range of energies of 2.000 to 27.980. It should be emphasized that the variables in these equations are highly correlated with each other, so that the parameters obtained will be sensitive to the composition of the set of π -molecules employed.

parameter	Eq (4)	Eq (14)
u_0	1.799	0.956
u_1	2.555	1.810
u_2	0.954	1.163
u_3	0.157	0.211

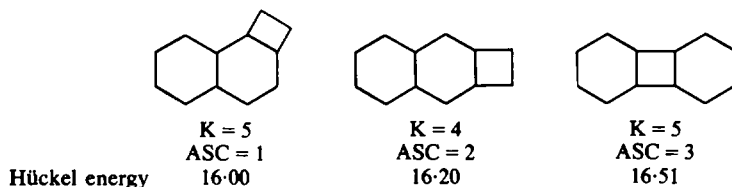
It has recently been demonstrated²⁶ that E_π is linearly related to the measurable total energy, $E_\pi + E_\sigma$. Thus, Eq (14) would also be used to expressed observably total energies in terms of the same graph theoretical parameters.

Eq (14) provides a natural explanation for the fact^{16,17} that the ASC (and not K) is a rough measure for stability

†Similarly, the various coefficients of the different powers of x in the characteristic polynomial of a graph were also first investigated by the late Professor Coulson in 1949. A more specific graph theoretical approach to this problem was by A. Mowshowitz, *J. Combinatorial Theory* 12 (B), 177 (1972).

(and related chemical behaviour) of conjugated systems. Particularly, in a series of isomeric molecules which have the same N and ν values and thus the same number of rings (it is tacitly assumed that they also have same ring sizes and similar branching), E_* is simply proportional to ASC .²⁷

Example



Therefore, the predictions of the traditional resonance theory approach should not be generally valid because the parity of the Kekulé structures is not taken into account. The early apparent success of the theory rested on the fortunate fact that all Kekulé structures for benzenoid hydrocarbons and acyclic polyenes have the same parity.^{11,16} For these molecules $ASC = K$. It is only in recent years that numerous examples of planar nonbenzenoid hydrocarbons have become available.²⁸ Since that number will surely increase relative to the number of benzenoid hydrocarbons it is important that organic chemists realize that the usual application of resonance theory is not based on a general chemical principle and that the more general formulation discussed here should be used.

Finally, we would like to emphasize that the dependence of E_* on ASC is logarithmic, so that ASC gives only a small, second order contribution to E_* . This is, of course, in agreement with the fact²⁹ that E_* is mainly determined by N and ν . Thus if one is interested in thermodynamic stability, the utility of the ASC is within a set of isomeric molecules (see the example above) where N and ν are identical. As will be developed elsewhere,³⁰ however, this limitation does not apply to chemical reactivity since it can be shown that reactivity also depends on related small differences in ASC .

REFERENCES

- ¹Part IX: D. Cvetković, I. Gutman and N. Trinajstić, *Theoret. Chim. Acta* **34**, 129 (1974)
- ²For details on resonance theory and further references see: G. W. Wheland, *The Theory of Resonance and Its Application to Organic Chemistry*, Wiley, New York (1953)
- ³L. Pauling, *J. Chem. Phys.* **1**, 280 (1933); L. Pauling and G. W. Wheland, *Ibid.* **1**, 362 (1933); L. Pauling and J. Sherman, *Ibid.* **1**, 679 (1933)
- ⁴C. A. Coulson, *Proc. Roy. Soc. A* **207**, 91 (1951); P. Daudel and R. Daudel, *J. Chem. Phys.* **16**, 639 (1948); M. B. Oakley and G. E. Kimball, *Ibid.* **17**, 265 (1950)
- ⁵Part I: A. Graovac, I. Gutman, N. Trinajstić and Živković, *Theoret. Chim. Acta* **26**, 67 (1972)
- ⁶Part II: D. Cvetković, I. Gutman and N. Trinajstić, *Croat. Chem. Acta* **44**, 365 (1972)
- ⁷Part III: I. Gutman and N. Trinajstić, *Chem. Phys. Lett.* **17**, 535 (1972)
- ⁸Part IV: I. Gutman and N. Trinajstić, *Croat. Chem. Acta* **45**, 539 (1973)
- ⁹Part V: I. Gutman and N. Trinajstić, *Chem. Phys. Lett.* **20**, 257 (1973)
- ¹⁰Part VI: I. Gutman, N. Trinajstić and T. Živković, *Tetrahedron* **29**, 3449 (1973)
- ¹¹Part VII: D. Cvetković, I. Gutman and N. Trinajstić, *J. Chem. Phys.* in press
- ¹²Part VIII: I. Gutman and N. Trinajstić, *Croat. Chem. Acta* **45**, 539 (1973)
- ¹³J. V. Knop, N. Trinajstić, I. Gutman and L. Klasinc, *Naturwissenschaften* **60**, 475 (1973); J. V. Knop, N. Trinajstić and T. Živković, *Coll. Czech. Chem. Comm.* **39**, 2431 (1974)
- ¹⁴E. Hückel, *Z. Phys.* **70**, 204 (1931)
- ¹⁵W. Heitler and F. London, *Ibid.* **44**, 455 (1927)
- ¹⁶M. J. S. Dewar and H. C. Longuet-Higgins, *Proc. Roy. Soc. A* **214**, 482 (1952)
- ¹⁷C. F. Wilcox, Jr., *Tetrahedron Letters* 795 (1968); *J. Am. Chem. Soc.* **91**, 2732 (1969)
- ¹⁸I. Gutman and N. Trinajstić, *Topics Curr. Chem.* **42**, 49 (1973)
- ¹⁹W. C. Herndon, *Tetrahedron* **29**, 3 (1973); *J. Chem. Educ.* **51**, 10 (1974)
- ²⁰The spectrum of a molecular graph lies in the interval $(-3, +3)$: C. A. Coulson, *Proc. Cambridge Phil. Soc.* **46**, 202 (1949)
- ²¹H. H. Günthard and H. Primas, *Helv. Chim. Acta* **51**, 1675 (1956)
- ²²F. Harary, *Graph Theory* (p. 14) Addison-Wesley, Reading, Mass. (second printing) (1971)
- ²³K. Ruedenberg, *J. Chem. Phys.* **34**, 1884 (1961)
- ²⁴I. Gutman, M. Milun and N. Trinajstić, *Ibid.* **59**, 2772 (1973)
- ²⁵I. Gutman, *Chem. Phys. Lett.* **24**, 283 (1974)
- ²⁶L. J. Schaad and B. A. Hess, Jr., *J. Am. Chem. Soc.* **94**, 3068 (1972)
- ²⁷G. G. Hall, *Internat. J. Math. Educ. Sci. Technol.* **4**, 233 (1973)
- ²⁸C. F. Wilcox, Jr., J. P. Uetrecht and K. K. Grohman, *J. Am. Chem. Soc.* **94**, 2532 (1972); S. Masamune, N. Nakamura, M. Suda and H. Ona, *Ibid.* **95**, 8481 (1973) and refs therein
- ²⁹B. J. McClelland, *J. Chem. Phys.* **54**, 640 (1971)
- ³⁰C. F. Wilcox, Jr., I. Gutman and N. Trinajstić, *Tetrahedron* **31**, 147 (1975)
- ³¹D. Cvetković, *Publ. Fac. Electrotechnique Univ. Belgrade* **354**, 1 (1971)

APPENDIX 1

The characteristic polynomial P_G of the molecular graph G is defined as:¹⁸

$$P_G(x) = \det(xI - A) \quad (15)$$

where I is a unit matrix of order N . P_G is a polynomial of degree N and can be represented as

$$P_G(x) = \sum_{j=0}^N a_j x^{N-j} \quad (16)$$

Hence

$$(ASC)^2 = |P_G(0)| = |a_N| \quad (17)$$

Suppose now that there are n zeros in the graph spectrum. Let asc and a_{N-n} be defined as

$$((asc)^2 = \prod_{i=1}^N |x_i|) \tag{18}$$

$$a_{N-n} = \prod_{i=1}^N |x_i| = (asc)^2 \tag{24}$$

where Π' denotes a product of all non-zero eigenvalues. Since $asc > 0$,

$$2 \ln(asc) \approx \sum_{i=0}^m (-1)^{i+1} u_i W_i \tag{19}$$

where

$$W_i = \sum_{j=1}^N |x_j|^i \tag{20}$$

and Σ' denotes summation over all non-zero eigenvalues. Of course,

$$W_i = W_i \tag{21}$$

and therefore in Eqs (12) and (14) asc should be written instead of ASC .

If $G-p$ is the graph obtained from G after deletion of the vertex p ,

$$\frac{\partial}{\partial x} P_G(x) = \sum_{p=1}^N P_{G-p}(x) \tag{22a}$$

$$\frac{\partial^2}{\partial x^2} P_G(x) = \sum_{p=1}^N \sum_{q=1}^N P_{G-p-q}(x) \tag{22b}$$

Similar relations hold for higher derivatives of P_G also.³¹

If there are n zeros in the graph spectrum,

$$P_G(0) = \frac{\partial}{\partial x} P_G(0) = \dots = \frac{\partial^{n-1}}{\partial x^{n-1}} P_G(0) = 0 \tag{23a}$$

$$\frac{\partial^n}{\partial x^n} P_G(0) = n! \quad a_{N-n} \neq 0 \tag{23b}$$

For reasons of simplicity we consider here the case of $n = 1$ only. Then from (22a):

$$a_{N-n} = \sum_{p=1}^N P_{G-p}(0) \tag{25}$$

But $P_{G-p}(0) = ((ASC)_p)^2$, where $(ASC)_p$ is the ASC value for the graph $G-p$, which finally gives (for $n = 1$)

$$(asc)^2 = \sum_{p=1}^N ((ASC)_p)^2 \tag{26}$$

Therefore asc has a similar but less simple meaning than ASC . Generalization for $n > 1$ follows straightforwardly from Eq (22b) etc). Thus for $n = 2$

$$(asc)^2 = \frac{1}{2} \sum_{p=1}^N \sum_{q=1}^N ((ASC)_{pq})^2 \tag{27}$$

where $(ASC)_{pq}$ is the ASC for the graph $G-p-q$.

APPENDIX 2

In order to derive a topological expression for W_i , note that there necessarily exists a number Q such that

$$\sum_{i=1}^N (|x_i| - Q)^3 = 0 \tag{28}$$

it is natural to expect that Q is near to the mean value of $|x_i|$'s. But

$$Q \approx \frac{1}{N} \sum_{i=1}^N |x_i| = \frac{E_w}{N} \tag{29}$$

After substitution of $Q \approx E_w/N$ in Eq (28), Eq (9) follows immediately.