match no.1 pp. 159-170 1975

GRAPHS IN CHEMISTRY

A. Graovac and N. Trinajstić
The Rugjer Bosković Institute, P.O.B. 1016,
41001 Zagreb, Croatia, Yugoslavia

Although very sophisticated methods for computing the electronic structure of rather complicated molecules are available today, it is appealing to have a simple method for qualitative predictions of the structure and reactivity of molecules without going through the mathematical manipulations of solving the eigenvalue problem of the Hamiltonian matrix, preferably using just a pencil and paper or a pictorial approach. Such an opportunity is offered by the application of graph theory in chemistry[1].

The basic description of a molecule is given by its structural formula i.e. by stating which atoms are connected to each other by chemical bonds. On the other hand a graph G is defined by a given set of vertices and a binary relation defined on this set of vertices (namely, two vertices are either connected by an "edge" or not). The analogy between the structural formulae (atoms, bonds) and the corresponding graphs (vertices, edges) is obvious. But, in quantum chemistry this analogy turned out to be a nontrivial one. In the 1930's E. Hückel described[2] the π-electrons in planar conjugated hydrocarbons by the following effective Hamiltonian matrix H:

$$H_{rs} = \begin{cases} \alpha_C & \text{if } r=s \\ \beta_{CC} & \text{if atoms r and s are bonded} \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

where α_C and β_{CC} are the standard (benzene) values of the Coulomb and the resonance integrals, respectively. On the other hand the so called adjacency (or topological) matrix A can be assigned to the corresponding molecular graph. It is defined as follows[3]:

$$A_{rs} = \begin{cases} 1 & \text{if } r=s \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

By using β as an energy unit (β =1) and taking α as a zero-energy reference point (α =0) matrices H and A become identical (if the numbering of the atoms and the vertices is the same). This means that in the simple MO-theory the orbital energies, the MO's and other molecular properties are directly shaped by the molecular topology[4]. A number of general results concerning the dependence of several physical and chemical properties (like stabilities and reactivities, total π -electron energies, HOMO-LUMO separations, distribution of orbital energies, number of non-bonding orbitals, resonance energies, charge density distributions, dipole moments, redox behaviour, boiling points, dependence of

vapour pressures on temperature, chromatographic retention volumes and retention times and others) on the molecular topology have been established[1,5], mainly, for the class of alternant hydrocarbons and in some cases for the non-alternant hydrocarbons, the positional isomers and others.
It is interesting to see whether this fruitful analogy between the simple π -electron Hamiltonian matrix H and the adjacency matrix A of the corresponding (molecular) graph can be extended in order to cover the heteroconjugated molecules.
The introduction of heteroatoms will cause, in the first step, appropriate changes in the diagonal elements of the Hamiltonian matrix. The values α_X of the Coulomb integral of heteroatom X can be written as:

$$\alpha_X = \alpha_C + h_X \beta_{CC} \tag{3}$$

A non-zero value of h is a device for description of the difference between the heteroatoms and the carbon atoms.
The graphs, which can be used to represent heteroconjugated molecules, are called rooted graphs[6]. These are graphs which contain one or more vertices which are considered to differ, in some way, from the others. These vertices of different "type" are called the roots or the rooted vertices. The adjacency matrix A, assigned to the rooted graph, is defined as follows:

$$A_{rs} = \begin{cases} h_r & \text{if } r=s, \text{ for rooted vertices} \\ 1 & \text{if } r=s, \text{ for other vertices} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

h may be thought of as some kind of "weighting" of the rooted vertices[7] or, alternatively, as some kind of assignment of "self-loops" to the rooted vertices[8].
The second step in treating heteroconjugated molecules includes, beside changes in the diagonal elements, also changes in the off-diagonal elements of the Hamiltonian matrix. The value β_{XY} of the resonance integral for the bond (edge) connecting atoms X and Y can be written as:

$$\beta_{XY} = b_{XY}\, \beta_{CC} \tag{5}$$

The b's describe the different types of bonds in the heteromolecules (but they can be used in treating the hydrocarbons as well).
The graphs which can be used now are called weighted rooted (or generalized) graphs[9]. They are graphs which contain one or more edges (and vertices) which are considered to differ in some way from the others. The edges of different "type" are called the weighted edges. The adjacency matrix assigned to the weighted rooted graph is defined as follows:

$$A_{rs} = \begin{cases} h_r & \text{if } r=s, \text{ for rooted vertices} \\ 1 & \text{if } r=s, \text{ for other vertices} \\ b_{rs} & \text{if } r \text{ and } s \text{ are connected, for weighted edges} \\ 1 & \text{if } r \text{ and } s \text{ are connected, for other edges} \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

b's may be thought of as some kind of "weighting" of the bonds[9].

Until now, we have treated the π-electron systems composed of orbital arrays in which there is no sign inversion (or more generally, in which there is an even number of sign inversions) among the adjacent 2p π-orbitals. In 1964, E. Heilbronner[10] discussed the stability of Möbius-type conformations of higher annulenes $(CH)_n$, $n > 20$. Möbius-type structures may also appear as transition states in electrocyclic closures of linear polyenes. Such systems are called Möbius or anti-Hückel systems (because their stability is governed by the rules opposite to those of Hückel) and they are defined[10] as cyclic arrays of orbitals in which there is one sign inversion, or more generally, in which there is an odd number of sign inversions resulting from the negative overlaps between the adjacent 2p π-orbitals of different

sign. Approximately, the whole effect of the consecutive twisting of π -orbitals can be assigned only to the pairs of orbitals of different sign. As a result, the resonance integrals, corresponding to such pairs, will change the sign (in comparison to the standard β_{CC} values).

In oreder to extend the graph-theoretical approach to Möbius systems we need to define a new type of graphs which are called Möbius[11,12], generalized[11,12], anti-Hückel or extended graphs[13]. The weight of edges in these graphs is either +1 or -1 depending whether two $2p^{\pi}$-atomic orbitals in a Möbius molecule are in the positive-positive (+1) or in the positive--negative (-1) overlap relationship, respectively. The adjacency matrix A assigned to the Möbius graph is defined as follows:

$$A_{rs}=\begin{cases} 1 & \text{if there is a positive edge between adjacent vertices r and s} \\ -1 & \text{if there is a negative edge between adjacent vertices r and s} \\ 0 & \text{otherwise} \end{cases} \qquad (7)$$

The whole concept can be generalized to polycyclic π -electron systems[11,12].

Formally, this matrix is a special case od adjacency matrix for weighted rooted graphs.

In the same manner as for the Hückel graphs matrices H and A

can be made identical, thus showing that the molecular topology (in the framework of simple MO-theory) will shape the chemical properties of heteroconjugated and Möbius systems, too.

It is interesting to note that the extended graphs appeared also in the attempts to estimate the energy gap of hydrocarbon polymers[14]. When the solid-state physics methods, based on the "translational symmetry" of a polymer are used, one can reduce[15] the problem of electronic band structure of a polymer to the diagonalization of the matrix:

$$T(k) = H + B \exp(ik) + B \exp(-ik), \quad k \in [0, 2\pi] \tag{8}$$

where H is the Hamiltonian matrix of an isolated unit, matrix B describes the interaction between this unit and, for example, its right neighbour, k corresponds to a crystal momentum and $i = \sqrt{-1}$. When the units are connected by one bond only, it can be shown[14,15] that the energy gap is determined by the matrices $T(k)$ for $k=0$ and $k=\pi$. In the framework of the simple MO-theory two graphs G_1 and G_2 can be associated with matrices $T(0)$ and $T(\pi)$, respectively. Let the graph G correspond to the polymer unit. The graph G_1 (or G_2) is obtained by addition (or subtraction) of the edge between the vertices in graph G which

are connected with the neighbouring units. The negative edge, and therefore the extended graph, appears in the case in which the vertices, which are connected with the neighbouring units, are not linked together in the graph G.

For all the systems mentioned previously, the problem of the orbital energies and the molecular orbitals (in the framework of simple MO-theory) can be completely reduced to the eigenvalue problem of the corresponding adjacency matrix. The eigenvalues (the orbital energies) are given as the roots of the characteristic polynomial P(G;x):

$$P(G;x) = \det\left|x\ I - A\right| = \sum_{n=0}^{N} a_n x^{N-n} \qquad (9)$$

where N is the number of vertices in the graph G and I is the unit matrix. Since the characteristic polynomial is uniquely defined by the graph G, the coefficients a_n can be found without going through the procedure of solving the determinant, but solely from the knowledge of the topological structure of the graph. Such a procedure has been suggested by Coulson[16] and put into a concise mathematical form by Sachs[17]. For the Hückel graphs the Sachs formula for the coefficients a_n may be very conveniently used for evaluating the polynomial coefficients and on this basis the structure-reactivity relationship can be developed[18,19,20]. Recently, Sachs' formula was extended to the rooted graphs, relating the structure of

a given rooted graph to the characteristic polynomial as a whole[7], or alternatively, to the individual coefficients a_n of the characteristic polynomial[8]. Sachs' formula for the weighted rooted graphs, relating the structure of a given graph to the characteristic polynomial as a whole[9] was established too. This formula is close to Coates' topological formula[21] which is in use in circuit theory. Sachs' formula for Möbius graphs, relating the structure of a given graph to the individual coefficients a_n, is a direct generalization of the original Sachs formula with a dramatic effect on the stability of the corresponding molecular systems[11,12]. The parity of algebraic structures[18,22] (Kekulé structures of Hückel molecules are special cases of algebraic structures) for such systems was defined[11,12].

The philosophy behind all the discussed Sachs-type formulae is based on the well known fact that the coefficients a_n of the characteristic polynomial $P(G;x)$ equals the arithmetic sum of all principal nth order minors of the adjacency matrix determinant under consideration. Graph-theoretically, this means that for evaluation of a_n the subgraphs of G with n vertices needs to be taken into account. Furthermore, from the definition of determinant it follows that only those subgraphs which have no other components but (isolated) edges and rings need to be considered. For such subgraphs the term Sachs' graphs of the graph G has been introduced[14]. In the case of rooted

(or weighted rooted) graphs the definition of Sachs' graphs was extended[8] to include graphs comprising self-loops.
The Sachs' theorem becomes a cumbersome method for enumeration of the coefficients a_n for large molecules. However, a real virtue of the method is in providing a single coefficient a_n (or a few of them) which may be the only ones needed. In particular, the value of a_N coefficient is very often required, because the structure-reactivity relationship shows that this coefficient can ba of value in predicting the stability of an investigated species[18]. a_N coefficient is simply related to the difference of the number of even and odd Kekulé (or more generally algebraic) structures[23]. Similarly, very good upper and lower bounds for the total π -electron energy of conjugated molecules are derived[24] which depend only on the coefficients a_N, the number of bonds (a_2) and the number of atoms (N) in a molecule.
The list of molecular systems, which can be treated by graph-theoretical methods, is far from being exhausted by those mentioned here. There is a priori no reason why topological MO's should not be able to describe every class of molecules. The best results to date have been obtained for inorganic complexes[25] and for boron hydrides[26]. Very large molecules, even infinite ones, can be investigated by use of graph-theoretical methods[14,15,27]. A very interesting application in this area may be an attempt to study the σ -electronic systems and to develop the topological theory of chemical reactivity.

Literature

1. see, for example: I. Gutman and N. Trinajstić, Topics Curr. Chem. 42, 49 (1973)
2. E. Hückel, Z. Physik 70, 204 (1931); 72, 310 (1931), etc.
3. F. Harary, Graph Theory. Reading, Mass. Addison-Wesley (1969)
4. K. Ruedenberg, J. Chem. Phys. 34, 1861 (1961)
5. M. Randić, J. Amer. Chem. Soc., in press
6. F. Harary and E. Palmer, Graphical Enumeration. Academic Press, New York (1973), p.6.
7. R.B. Mallion, N. Trinajstić and A.J. Schwenk, Z. Naturforsch. 29a, 1481 (1974)
8. R.B. Mallion, A.J. Schwenk and N. Trinajstić, Croat. Chem. Acta 46, 171 (1974)
9. A. Graovac and N. Trinajstić, to be published
10. E. Heilbronner, Tetrahedron Lett., 1923 (1964)
11. A. Graovac and N. Trinajstić, J. Mol. Structure, in press
12. A. Graovac and N. Trinajstić, Croat. Chem. Acta, in press
13. M. Randić, to be published
14. A. Graovac, I. Gutman, M. Randić and N. Trinajstić, to be published
15. D.A. Bochvar and J.V. Stankevich, Zhur. Strukt. Khim. 8, 943 (1967)
16. C.A. Coulson, Proc. Cambridge Phil. Soc. 46, 202 (1950)

17. H. Sachs, Publ. Math. (Debrecen) 11, 199 (1963)

18. A. Graovac, I. Gutman, N. Trinajstić and T. Živković, Theoret. Chim. Acta 26, 67 (1972)

19. I. Gutman and N. Trinajstić, Croat. Chem. Acta 45, 423 (1972)

20. I. Gutman and N. Trinajstić, Chem. Phys. Lett. 17, 535 (1972); 20, 257 (1973)

21. C.L. Coates, IRE Trans. on Circuit Theory CT-6, 170 (1959)

22. D. Cvetković, I. Gutman and N. Trinajstić, J. Chem. Phys. 61, 2700 (1974)

23. C.F. Wilcox, Jr., Tetrahedron Lett., 795 (1968)

24. I. Gutman, Chem. Phys. Lett. 24, 283 (1974)

25. H.H. Schmidtke, Coord. Chem. Rev. 2, 3 (1967)

26. S.F.A. Kettle, Theoret. Chim. Acta 3, 211 (1965); 4, 150 (1966)

27. I. Gutman, Naturwissenschaften 61, 216 (1974)