CROATICA CHEMICA ACTA CCACAA **51** (1) 29—33 (1978)

YU ISSN 0011-1643

CCA-1096

541.57

Original Scientific Paper

Topological Formulas for Free-Valence Index

I. Gutman

Department of Chemistry, Faculty of Sciences, University of Kragujevac, 34000 Kragujevac, R. Domanovića 12, Serbia, Yugoslavia

Received September 19, 1977

Four topological identities (5), (6), (7) and (9) for free-valence index are derived.

The bond number N_r and the closely related free-valence index F_r were defined by Coulson[1,2] as

$$N_r = \sum_s P_{rs}$$

$$F_r = N_{max} - N_r$$

Thereafter, this concept has been widely applied in the theory of conjugated molecules.[3]

In the above formulas, P_{rs} is the bond order between atoms r and s and the constant N_{max} is the maximal possible bound number, usually assumed to be $\sqrt{3}$. Here and later $\sum_s$ denotes the summation over all sites s adjacent to r.

Both indices N_r and F_r contain the same information about the π-electron system of a conjugated molecule. In the following we shall analyse mainly the properties of the bond number, but, of course, our results apply to free-valence index as well.

For the purposes of the present consideration it is convenient to use the language and mathematical formalism of graph theory.[4] Thus G is the molecular graph with n vertices. Its adjacency matrix is A and its characteristic polynomial is P (G, x), where

$$P(G, x) = \det(xI - A) = \Delta \tag{1}$$

Sometimes the characteristic polynomial will be denoted by Δ for reasons which will become clear later. The eigenvalues x_j and the eigenvectors $C_j = (C_{j1}, C_{j2}, \ldots, C_{jn})^T$ of the graph G fulfil the relation

$$A\,C_j = x_j\,C_j$$

If e is an edge and r a vertex of G, then G-e and G-r denote the subgraphs obtained by deletion of e and r, respectively, from G. Consequently, G-r-s is obtained by deletion of both vertices r and s. Note that G-e, G-r and G-r-s have n, n-1 and n-2 vertices, respectively. If Z is a cycle of G, then G-Z is the subgraph obtained by deletion of all vertices of Z from G. Hence, G-Z has n-z vertices, with z being the size of the cycle Z.

Finally, we shall use the abbreviated notation

$$\frac{1}{\pi}\int_{-\infty}^{\infty} F(x)\, dx \equiv < F(x) > \equiv < F > \tag{2}$$

In the above defined graph-theoretical terminology, the total π-electron energy E and the density matrix P of the corresponding conjugated system are introduced as[4]

$$E = \sum_{j=1}^{n} g_j x_j$$

$$P = \| P_{rs} \| ; \qquad P_{rs} = \sum_{j=1}^{n} g_j C_{jr} C_{js}$$

with g_j being the occupation number of the j'th molecular orbital. In the great majority of cases, the bonding orbitals ($x_j > 0$) are occupied by two electrons, whereas the antibonding orbitals ($x_j < 0$) are empty. Then the occupation numbers fulfil the identity[5]

$$g_j x_j = x_j + | x_j |$$

From this relation it follows

$$A P = A + | A | \tag{3}$$

where the matrix $| A |$ is defined via

$$| A | C_j = | x_j | C_j \tag{4}$$

Since $A_{rr} = 0$, it is

$$| A |_{rr} = (A P)_{rr} = \sum_{s} P_{rs}$$

and therefrom we have our *first identity*

$$N_r = | A |_{rr} \tag{5}$$

The *second identity* follows now immediately from the definition (4) of the matrix $| A |$, namely

$$N_r = \sum_{j=1}^{n} | x_j | C_{jr}^2 \tag{6}$$

This relation has an intriguing interpretation, because C_{jr}^2 is just the density of the electron in the j'th orbital at site r. Hence, N_r can be understood as a weighted electron density, with the major contribution coming from the lowest occupied molecular orbital (LOMO) and with negligible contribution from the highest occupied molecular orbital (HOMO). This is seemingly a paradoxial conclusion since the reactivity of a site is known to be mainly determined by the HOMO. But in the case of the free-valence index (which is the real reactivity index), the situation is exactly the opposite. Namely,

$$F_r = \sum_{j=1}^{n} (N_{max} - | x_j |)\, C_{jr}^2$$

If the constant N_{max} is chosen to be sufficiently close to the LOMO level, the contribution of the LOMO to F_r will vanish while the contribution of the HOMO will be the dominant one. This argument explains also the finding that $N_{max} =$

$= \sqrt{3}$ can be applied in the case of both tertiary, secondary and primary carbon atoms.[6]

The trace of a matrix is invariant under similarity transformations. Therefore,

$$\sum_{r=1}^{n} N_r = \mathrm{Tr} \,|\, A \,| = \sum_{j=1}^{n} |\, x_j \,| = E$$

which is just another formulation of the well known fact $E = \Sigma P_{rs}$. Thus, the bond number N_r can be interpreted as that fraction of the total π-electron energy which can be associated with the site r.

Another partitioning of E on sites of the molecule can be realized starting from the Coulson integral formula[7]

$$E = < \mathrm{n} - \mathrm{i}x\, P' (\mathrm{G} - \mathrm{r},\ \mathrm{i}x)/P\,(\mathrm{G},\ \mathrm{i}x) >$$

where $\mathrm{i} = \sqrt{-1}$. Because of the relation

$$P' (\mathrm{G},\ x) = \sum_{r=1}^{n} P\,(\mathrm{G} - \mathrm{r},\ x)$$

it follows simply

$$E = \sum_{r=1}^{n} < 1 - \mathrm{i}x\, P\,(\mathrm{G} - \mathrm{r},\ \mathrm{i}x)/P\,(\mathrm{G},\ \mathrm{i}x) >$$

We proceed now to show that this partition of total π-electron energy is equivalent to the previous one, namely that a *third identity* is valid.

$$N_r = < 1 - \mathrm{i}x\, P\,(\mathrm{G} - \mathrm{r},\ \mathrm{i}x)/P\,(\mathrm{G},\ \mathrm{i}x) > \tag{7}$$

In order to deduce eq. (7), we expand the determinant (1) in terms of the r'th row. Δ_{rs} will denote the minor obtained by deletion of the r'th row and the s'th column from Δ. Note that $\Delta_{rr} = \mathrm{P}\,(\mathrm{G\text{-}r},\ \mathrm{x})$. Hence,

$$P\,(\mathrm{G},\ \mathrm{x}) = \sum_{s=1}^{n} (-1)^{r+s} (\mathrm{x}\, I - A)_{rs}\, \Delta_{rs} = x\, \Delta_{rr} + \sum_{\substack{s=1 \\ (s \neq r)}}^{n} (-1)^{r+s+1} A_{rs}\, \Delta_{rs} =$$

$$= x\, P\,(\mathrm{G} - \mathrm{r},\ x) + \sum_{s} (-1)^{r+s+1} \Delta_{rs}$$

This is further transformed into

$$1 - x\, P\,(\mathrm{G} - \mathrm{r},\ x),/P\,(\mathrm{G},\ x), = \sum_{s} (-1)^{r+s+1} \Delta_{rs}/\Delta$$

Substitution $x \rightarrow \mathrm{i}x$ and integration as indicated in (2) yields

$$< 1 - \mathrm{i}x\, P\,(\mathrm{G} - \mathrm{r},\ \mathrm{i}x)/P\,(\mathrm{G},\ \mathrm{i}x) > = \sum_{s} (-1)^{r+s+1} < \Delta_{rs}/\Delta >$$

But according to the Coulson-Longuet-Higgins formula[8],

$$(-1)^{r+s+1} < \Delta_{rs}/\Delta > = P_{rs}$$

which completes the proof of eq. (7).

The characteristic polynomial of a graph fulfils the following recurrence relation[9]

$$P(\mathrm{G}, x) = P(\mathrm{G} - \mathrm{e}, x) - P(\mathrm{G} - \mathrm{r} - \mathrm{s}, x) - 2 \sum_{\mathrm{Z}}{}_{\mathrm{e}}\, P(\mathrm{G} - \mathrm{Z}, x) \tag{8}$$

where e is the edge between the vertices r and s and Σ_e denotes the summation over all cycles Z which contain the edge e. A succesive application of (8) to all edges incident to the vertex r gives

$$P(\mathrm{G}, x) = x\, P(\mathrm{G} - \mathrm{r}, x) - \sum_{\mathrm{s}} P(\mathrm{G} - \mathrm{r} - \mathrm{s}, x) - 2 \sum_{\mathrm{Z}}{}_{\mathrm{r}}\, P(\mathrm{G} - \mathrm{Z}, x)$$

with Σ_r indicating the summation over all cycles Z which contain the vertex r. Substitution of this latter relation back into eq. (7) yields

$$N_{\mathrm{r}} = - \sum_{\mathrm{s}} < P(\mathrm{G} - \mathrm{r} - \mathrm{s}, \mathrm{i}x)/P(\mathrm{G}, \mathrm{i}x) > - 2 \sum_{\mathrm{Z}}{}_{\mathrm{r}} < P(\mathrm{G} - \mathrm{Z}, \mathrm{i}x)/P(\mathrm{G}, \mathrm{i}x) > \tag{9}$$

which is our *fourth identity* for N_r. According to it, N_r is decomposed into a number of additive contributions of the form $< P(\mathrm{H}, \mathrm{i}x)/P(\mathrm{G}, \mathrm{i}x) >$ with H being a certain subgraph of G. The analytical form of these integrals is relatively simple and enables a deeper insight into the complicated topological relations which determine the bond number and the free-valence index.[10]

In order to illustrate this, let us consider the sign of the integrals in (9) for the case of an alternant hydrocarbon. It can be shown that

$$\operatorname{sign} < P(\mathrm{H}, \mathrm{i}x)/P(\mathrm{G}, \mathrm{i}x) > = (-1)^{(\mathrm{n}-\mathrm{h})/2}$$

with h being the number of vertices of H. Now, G-r-s and G-Z have n-2 and n-z vertices, respectively, and eq. (9) is transformed into

$$N_{\mathrm{r}} = \sum_{\mathrm{s}} \left| < P(\mathrm{G} - \mathrm{r} - \mathrm{s}, \mathrm{i}x)/P(\mathrm{G}, \mathrm{i}x) > \right| + 2 \sum_{\mathrm{Z}}{}_{\mathrm{r}} (-1)^{z/2-1} \,\Big| < P(\mathrm{G} - \mathrm{Z}, \mathrm{i}x)/P(\mathrm{G}, \mathrm{i}x) >$$

In words: The bond number of the site r is a sum of positive contributions coming from edges incident to r, of positive contributions coming from cycles of length 4m + 2 and of negative contributions coming from cycles of length 4m. Hence a Hückel 4m + 2 type regularity is observed also in the case of bond number and free-valence index.

This and related topological properties of N_r and F_r will be analysed in more detail elsewhere.[10]

Acknowledgment. — The author gratefully acknowledges the support of the present research by the Alexander von Humboldt Foundation.

REFERENCES

1. C. A. Coulson, *Trans, Faraday Soc.* **42** (1946) 265.
2. C. A. Coulson, *Disc. Faraday Soc.* **2** (1947) 9.
3. A. Streitwieser, *Molecular Orbital Theory for Organic Chemists,* Wiley, New York 1966, pp. 329—341 .
4. For review see: I. Gutman and N. Trinajstić, *Topics Curr. Chem.* **42** (1973) 49; *Croat. Chem. Acta* **47** (1975) 507.
5. K. Ruedenberg, *J. Chem. Phys.* **34** (1961) 1884.
6. H. H. Greenwood, *Trans. Faraday Soc.* **48** (1952) 677; J. D. Roberts, A. Streitwieser, and C. M. Regan, *J. Amer. Chem. Soc.* **74** (1952) 4579.
7. C. A. Coulson, *Proc. Cambridge Phil. Soc.* **36** (1940) 201; A. Graovac, I. Gutman, and N. Trinajstić, *Chem. Phys. Letters* **35** (1975) 555.

8. C. A. Coulson and H. C. Longuet-Higgins, *Proc. Roy. Soc. (London)* **A 191** (1947) 39; see also: H. Hosoya and K. Hosoi, *J. Chem. Phys.* **64** (1976) 1065.
9. H. Hosoya, *Theoret. Chim. Acta* **25** (1972) 215.
10. I. Gutman, in preparation.

SAŽETAK

Topološke formule za indeks slobodne valencije

I. Gutman

Izvedene su četiri topološke identičnosti za indeks slobodne valencije.

ZAVOD ZA HEMIJU
PRIRODNO-MATEMATIČKOG
FAKULTETA UNIVERZITETA
U KRAGUJEVCU
34000 KRAGUJEVAC, JUGOSLAVIJA

Prispjelo 19. rujna 1977.