
**SIMILARITY APPROACH TO CHEMICAL REACTIVITY.
SPECIFICITY OF MULTIBOND PROCESSES**

Robert PONEC and Martin STRNAD

*Institute of Chemical Process Fundamentals,
Czechoslovak Academy of Sciences, 165 02 Prague 6 - Suchbát*Received March 6, 1990
Accepted March 28, 1990

The recently suggested approach to analysis of chemical reactivity in terms of the similarity index has been generalized by introduction of similarity indices of higher orders. Using an example of concrete application of the third order similarity index some specific features accompanying the process of electron reorganization in the so-called multibond reactions have been analyzed and discussed.

One of general features appearing in the development of all exact natural sciences is the increasing utilization of mathematical methods and procedures with the aim of giving more precision and numerical quantification to a number of intuitively introduced but, nevertheless, immensely useful qualitative notions. One of such notions is that of molecular similarity whose utilization has attracted attention of both some agrochemical and pharmaceutical firms and chemical theory.

Among the various approaches to characterization of molecular similarity¹⁻³ the one based on utilization of the so-called similarity index r_{AB} (refs⁴⁻⁹) has begun to make itself felt increasingly not long ago. The framework of these attempts also includes our report¹⁰ which — by incorporating the original definition by Polansky⁵ and Carbo⁴ into the framework of the overlap determinant method¹¹ — showed a new, very broad field of application of the similarity approach, viz. for an alternative formulation of the Woodward-Hoffmann rules. Some possibilities of such applications are published e.g. in refs¹²⁻¹⁴.

Along with these works making use of both the similarity index itself and some further quantities derived therefrom subsequently we have recently introduced¹⁵ the so-called second-order similarity index g_{AB} as a further alternative increasing the predictive and discriminative abilities of the similarity index. The definition of this g_{AB} index is formally quite analogous to that of the r_{AB} index, and the only (but of course very substantial) difference consists in the fact that for comparison of similarity of electronic structures of molecules A and B it does not use the one-particle density matrices $\rho_A(1)$ and $\rho_B(1)$ but instead uses the second-order density matrices with greater information content.

In connection with such a procedure there arises a quite natural question whether or not the respective generalization could continue further by introduction of third-order and higher-order density matrices and analogous definition of the corresponding similarity indices. In this present communication our purpose is utilization of the above-mentioned philosophy for introduction of the third-order similarity index q_{AB} as the next simplest representative of the series of generalized indices. Using a number of examples of comparison with analogous values of the first-order (r_{AB}) and the second-order (g_{AB}) similarity indices we have discussed possible applications of this index to the theory of pericyclic reactivity. It has been shown that the application of the third-order density matrix in the q_{AB} index really leads to a further, nevertheless still qualitative increase of its information and discrimination content. Inter alia this enables revealing of the specific role which is played by the so-called multibond processes within the whole class of pericyclic reactions.

THEORETICAL

As the generalization suggested is immediately connected in its concepts with a number of earlier papers¹⁰⁻¹⁵ dealing with the problem of similarity indices, we will restrict our text to only a brief presentation of basic relations with the reference to original literature where all the necessary equations can be found. Therefore, on the basis of immediate analogy with earlier papers, we will define – at first – the third-order similarity index q_{AB} by the relation:

$$q_{AB} = \frac{\int \rho_A^{(3)}(1, 2, 3) \rho_B^{(3)}(1, 2, 3) d\tau_1 \dots d\tau_3}{[\int \rho_A^{(3)}(1, 2, 3) d\tau_1 \dots d\tau_3]^{1/2} [\int \rho_B^{(3)}(1, 2, 3) d\tau_1 \dots d\tau_3]^{1/2}}, \quad (1)$$

where $\rho_A^{(3)}(1, 2, 3)$ and $\rho_B^{(3)}(1, 2, 3)$ are the reduced spinless third-order density matrices of the structures A and B, resp., and $d\tau_i$ means the volume element of the i -th electron. This general relation can be further simplified similarly as the previous expressions for the r_{AB} and g_{AB} indices in the case of approximation of wave functions of the structures A and B by the single Slater determinant:

$$\Phi_A = |\varphi_1 \bar{\varphi}_1 \varphi_2 \bar{\varphi}_2 \dots \varphi_n \bar{\varphi}_n|, \quad (2a)$$

$$\Phi_B = |\psi_1 \bar{\psi}_1 \psi_2 \bar{\psi}_2 \dots \psi_n \bar{\psi}_n|. \quad (2b)$$

In such case the general expression for the 3. order density matrices is reduced to the relation:

$$\begin{aligned} \rho_X^{(3)}(1, 2, 3) = & \rho_X^{(1)}(1) \rho_X^{(1)}(2) \rho_X^{(1)}(3) + \frac{1}{2} \rho_X^{(1)}(1, 2) \rho_X^{(1)}(1, 3) \rho_X^{(1)}(2, 3) - \\ & - \frac{1}{2} \rho_X^{(1)}(1) [\rho_X^{(1)}(2, 3)]^2 - \frac{1}{2} \rho_X^{(1)}(2) [\rho_X^{(1)}(1, 3)]^2 - \\ & - \frac{1}{2} \rho_X^{(1)}(3) [\rho_X^{(1)}(1, 2)]^2 \quad (X = A, B), \end{aligned} \quad (3)$$

where $\varrho_X^{(1)}(i)$ and $\varrho_X^{(1)}(i, j)$ are the respective diagonal and nondiagonal elements of the first-order density matrix defined by Eqs (4) in terms of usual charge density bond order matrix \mathbf{P} :

$$\varrho_A(i) = 2 \sum_{k=1}^{\text{occ}} \varphi_k^2(i), \quad (4a)$$

$$\varrho_B(i) = 2 \sum_{k=1}^{\text{occ}} \psi_k^2(i), \quad (4b)$$

$$\varrho_A(i, j) = 2 \sum_{k=1}^{\text{occ}} \varphi_k(i) \varphi_k(j), \quad (4c)$$

$$\varrho_B(i, j) = 2 \sum_{k=1}^{\text{occ}} \psi_k(i) \psi_k(j). \quad (4d)$$

In terms of such an approach and with subsequent utilization of the topological approximation which eliminates the problems of invariance with respect to mutual positions of molecules^{10,15} the original definition expression (1) can be simplified to the final form:

$$q_{AB} = \frac{3 \text{Tr}^3 \mathbf{P}_A \bar{\mathbf{P}}_B - 7 \text{Tr} \mathbf{P}_A \bar{\mathbf{P}}_B \text{Tr} (\mathbf{P}_A \bar{\mathbf{P}}_B)^2 + 4 \text{Tr} (\mathbf{P}_A \bar{\mathbf{P}}_B)^3}{8N(3N^2 - 14N + 16)}. \quad (5)$$

With respect to the fact that the meaning of all the individual symbols in this relation is fully identical with that in the analogous expressions for the similarity indices of the first (Eq. (6)) and the second (Eq. (7)) orders,

$$r_{AB} = \text{Tr} \mathbf{P}_A \bar{\mathbf{P}}_B / 2N, \quad (6)$$

$$g_{AB} = \frac{9 \text{Tr}^2 \mathbf{P}_A \bar{\mathbf{P}}_B - 7 \text{Tr} (\mathbf{P}_A \bar{\mathbf{P}}_B)^2}{4N(9N - 14)}, \quad (7)$$

we will not repeat the respective data here, and we will turn directly to the presentation of the proper results which concern the application of the q_{AB} index to the theory of pericyclic reactivity. In order to maintain continuity with the analogous earlier studies based on applications of the r_{AB} and g_{AB} indices, we will restrict this present work to substantially the same — only slightly extended — reaction series, so that immediate comparison of the calculated values might be possible. This circumstance will also make it possible to restrict the specification of the necessary technical details of the calculations carried out. These data can be found in the original paper¹⁰. Hence we will only give the most basic piece of information, viz. that the calculations of the respective density matrices \mathbf{P}_A and \mathbf{P}_B were carried out by the simple

HMO method. The corresponding values of the individual similarity indices r_{AB} , g_{AB} , and q_{AB} are presented in Table I.

RESULTS AND DISCUSSION

Let us try to discuss some general conclusions following from Table I. First of all it can be seen that all the cases studied show a distinct general trend to a decrease in absolute values of the respective indices on transition to higher-order similarity indices. This result is quite natural, because it is obvious that with a more detailed description of electronic structure in terms of 2. and 3. order density matrices also the probability of sufficiently high coincidence in all the involved aspects of electronic structure will decrease entirely in the nature of a law. However, for our purposes

TABLE I

The values of the first-, second-, and third-order similarity indices calculated for a series of selected pericyclic reactions (a allowed, f forbidden)

Reaction A → B	r_{AB}		g_{AB}		q_{AB}	
	a	f	a	f	a	f
Butadiene → cyclobutene	0.724	0.500	0.524	0.091	—	—
Hexatriene → → cyclohexadiene	0.759	0.659	0.566	0.358	0.419	0.129
Octatetraene → → cyclooctatriene	0.718 ^a	0.661 ^a	0.503	0.401	0.344	0.212
Ethene + ethene → → cyclobutane	0.500	0.500	0.250	0.091	—	—
Ethene + butadiene → → cyclohexene	0.575	0.575	0.298	0.272	0.151	0.085
Butadiene + butadiene → → 1,5-cyclooctadiene	0.612	0.612	0.347	0.341	0.183	0.166
Hexatriene + ethene → → 1,3-cyclooctadiene	0.621	0.621	0.361	0.352	0.197	0.174
Butadiene → bicyclobutane	0.500	0.500	0.218	0.123	—	—
Cope rearrangement of 1,5-hexadiene	0.500	0.500	0.206	0.206	0.081	0.056
Benzene → benzvalene	0.611	0.611	0.334	0.321	0.161	0.134

^a The values 0.794 and 0.737 given in ref.¹⁰ are loaded with a numerical error.

more significant than the decrease itself in the absolute values of the respective indices is the fact that the transition to the similarity indices of higher orders not only removes the insufficient differentiation of some reactions at the level of r_{AB} index, but also the respective difference between allowed and forbidden processes are increased in accordance with expectation. Thus another generally expected result is confirmed, viz. that increasing information content of similarity indices of higher orders results also in increasing resolvability of relatively subtle differences in electronic reorganizations of allowed and forbidden processes.

Beside this undoubtedly important conclusion it can be seen that the data of Table I allow a number of other interesting conclusions. First of all we must mention a certain regularity which can be read from the respective values concerning reactions which cannot be differentiated on the basis of the original similarity index r_{AB} . It turns out that all such cases can generally be included among the reactions belonging to the so-called multibond processes¹⁶ (according to a recent classification by Dewar). This result is very useful, since the fact that the differentiation between an allowed and forbidden mechanism of these reactions necessitates – at least – the second-order similarity index indicates a certain delicacy of the reactions. This delicacy follows from the fact that the differences in electronic reorganization between allowed and forbidden reactions are so fine in this case that their respecting requires to involve the electron correlation reflected – at least partially – in the g_{AB} index. Particularly significant in this respect is the case of the Cope rearrangement: Table I shows that even the g_{AB} index is insufficient for differentiation between the allowed and forbidden mechanism, the differentiation being attained at the level of the q_{AB} index only. This result indicates that an adequate description of reaction in this case obviously needs inclusion of correlations of higher orders.

All these results indicate that the position of the so-called multibond reactions really is rather special, and that also their exclusion from the whole class of the other pericyclic reactions has deep physical grounds, regardless of whether or not detailed quantum chemical studies will confirm the mechanistic implications suggested by Dewar¹⁶.

APPENDIX

In a general case it would be possible to define the similarity index of the m -th order for structures A and B as follows:

$$f_{AB}^{(m)} = \frac{\int \varrho_A^{(m)}(1, 2, \dots, m) \varrho_B^{(m)}(1, 2, \dots, m) d\tau_1 d\tau_2 \dots d\tau_m}{[\int \varrho_A^{(m)}(1, \dots, m) d\tau_1 \dots d\tau_m]^{1/2} [\int \varrho_B^{(m)}(1, \dots, m) d\tau_1 \dots d\tau_m]^{1/2}},$$

where $\varrho_A^{(m)}(1, 2, \dots, m)$ and $\varrho_B^{(m)}(1, 2, \dots, m)$ are the spinless reduced density matrices of the m -th order. This equation can be rewritten for any m in a way similar to the

relations (5)–(7), which results from the fact that the discussed density matrices $\varrho_A^{(m)}$ and $\varrho_B^{(m)}$ can always be expressed as functions of the density matrices $\varrho_A^{(1)}$ and $\varrho_B^{(1)}$, respectively. For this purpose it is advantageous to use the standard notation of a reduced density matrix of the m -th order, $\Gamma^{(m)}$, of the wave function Ψ :

$$\Psi = |\phi_1 \phi_2 \dots \phi_N|$$

in the form

$$\Gamma^{(m)} = \begin{vmatrix} \Gamma^{(1)}(1, 1) & \Gamma^{(1)}(1, 2) & \dots & \Gamma^{(1)}(1, m) \\ \Gamma^{(1)}(2, 1) & & & \vdots \\ \vdots & & & \vdots \\ \Gamma^{(1)}(m, 1) & \dots & \dots & \Gamma^{(1)}(m, m) \end{vmatrix}.$$

In the case of closed shell wave functions (2), however, the simple replacement of $\Gamma^{(1)}(i, j)$ matrices by $\varrho^{(1)}(i, j)$ matrices in the determinant does not represent a direct identification of the $\Gamma^{(m)}$ matrix with $\varrho^{(m)}$ matrix. In order to obtain it, one must develop the whole determinant and simultaneously carry out replacement of each chain:

$$\begin{aligned} & \varrho^{(1)}(i_1, i_2) \varrho^{(1)}(i_2, i_3) \dots \varrho^{(1)}(i_{h-1}, i_h) \varrho^{(1)}(i_h, i_1) \rightarrow \\ & \rightarrow \frac{1}{2^{h-1}} \varrho^{(1)}(i_1, i_2) \varrho^{(1)}(i_2, i_3) \dots \varrho^{(1)}(i_{h-1}, i_h) \varrho^{(1)}(i_h, i_1), \end{aligned}$$

where $2 \leq h \leq m$, $1 \leq i_j \leq m$, and $i_j \neq i_k$ for $j \neq k$. In this way it is easily possible to obtain the spinless reduced density matrices of structures A and B of any order (see e.g. relation (3) for $m = 3$).

Since, however, the deriving of the final relation for similarity index of general order does not concern the problems of this present paper, we will not deal with it here. In conclusion it must be noted that the denotations still used for the individual similarity indices correspond to the following identities: $r_{AB} \equiv f_{AB}^{(1)}$, $g_{AB} \equiv f_{AB}^{(2)}$, and $q_{AB} \equiv f_{AB}^{(3)}$.

REFERENCES

1. Thornber C. W.: Chem. Soc. Rev. 8, 563 (1979).
2. Amooore J. E.: *Molecular Basis of Odor*. Thomas, Springfield 1970.
3. Amooore J. E., Palmieri G., Wanke E., Blum M. S.: Science 165, 1266 (1969).
4. Carbo R., Leyda L., Arnau M.: Int. J. Quantum Chem. 17, 1185 (1980).
5. Polansky O. E., Derflinger G.: Int. J. Quantum Chem. 1, 379 (1969).
6. Hodgkin E. E., Richards W. G.: J. Chem. Soc., Chem. Commun. 1986, 1342.
7. Hodgkin E. E., Richards W. G.: Int. J. Quantum Chem., Quant. Biol. Symp. 14, 105 (1987).
8. Bowen-Jenkins P. E., Cooper D. L., Richards W. G.: J. Phys. Chem. 89, 2196 (1985).
9. Cooper D. L., Allan N. L.: J. Comput. Aided Design, in press.

10. Ponec R.: *Collect. Czech. Chem. Commun.* 52, 555 (1987).
11. Ponec R.: *Collect. Czech. Chem. Commun.* 49, 455 (1984).
12. Ponec R.: *Z. Phys. Chem. (Leipzig)* 268, 1180 (1987).
13. Ponec R.: *Z. Phys. Chem. (Leipzig)* 270, 365 (1989).
14. Ponec R., Strnad M.: *Collect. Czech. Chem. Commun.*, in press.
15. Ponec R., Strnad M.: *Collect. Czech. Chem. Commun.* 55, 896 (1990).
16. Dewar M. J. S.: *J. Am. Chem. Soc.* 106, 209 (1984).

Translated by J. Panchartek.