PARTITIONING OF THE ORBITAL OVERLAP MATRIX AND THE LOCALIZATION CRITERIA*

J. CIOSLOWSKI

Department of Chemistry and Supercomputer Computations Research Institute, Florida State University, Tallahassee, FL 32306-3006, USA

Abstract

A new molecular orbital localization procedure is proposed. The approach is based on partitioning of the overlap matrix into atomic contributions in accordance with Bader's topological theory of atoms in molecules. The new procedure has several advantages over other schemes. It preserves the σ/π-separability in planar systems and allows for a straightforward interpretation of the localized orbitals in terms of their localization indices and atomic occupancies. The new procedure is tested on the H_2O, LiF, N_2, CO, $BH_3 \cdot CO$ and Li_2 molecules.

1. Introduction

A single-determinantal many-electron wave function is invariant with respect to any unitary transformation among its one-electron functions (spin-orbitals) [1]. The canonical molecular orbitals (CMOs) that diagonalize the Fock operator reflect the symmetry of the irreducible representations of the molecular point group and therefore are delocalized over the entire molecular framework. The classical Lewis–Linnett picture [2] of localized bonds, lone pairs, etc. can be retrieved by carrying out a localization transformation among CMOs [3–5].

From the mathematical point of view, localization of CMOs is equivalent to finding a unitary transformation matrix C that maximizes the localization sum

$$L = \sum_i \sum_{klmn} C_{ik}^* C_{il} C_{im}^* C_{in} T_{klmn} .$$ (1)

The matrix C that relates LMOs to CMOs is usually computed through a series of 2×2 rotations [6] or by a second-order maximization procedure [7]. Due to its unitary character, the localization transformation leaves the many-electron wave function unaltered. The tensor T_{klmn} defines the localization procedure. The most widely used choices for T_{klmn} are

$$T_{klmn}^{ER} = \langle k_1 m_2 | r_{12}^{-1} | l_1 n_2 \rangle, \tag{2a}$$

(the Edmiston–Ruedenberg (ER) localization [3]);

$$T_{klmn}^{FB} = -\langle k | r | l \rangle \cdot \langle m | r | n \rangle, \tag{2b}$$

(the Foster–Boys (FB) localization [4]); and

$$T_{klmn}^{N} = \langle km | ln \rangle, \tag{2c}$$

(the von Niessen (N) localization [5]).

For a molecule with M electrons, the ER and N localization methods require $\approx M^5$ mathematical operations, while for the FB localization the computational effort is proportional to M^3. In the case of (locally) planar molecules, the LMOs obtained from any of the above procedures often have a form of so-called τ or banana orbitals rather than the σ and π ones [8]. One should point out that, although there are no definite physical reasons for preferring the σ/π LMOs over the τ ones, one obtains a much clearer picture of bonding using the former LMOs. The LiF molecule can serve as a typical example: the FB localization yields the following LMOs: two of the inner-shell type, one σ lone-pair on the F atom and three equivalent LMOs of the τ type [9]. This is in contrast to the set of two inner-shell LMOs together with one s and three p lone-pair LMOs on the F atom anticipated from the classical Li^+F^- picture. Another drawback of the above localization procedures is that they do not provide any simple indices that could be used for judging the extent of localization/ delocalization and/or calculating atomic electron populations. Usually, the character of LMOs has to be examined from their plots.

Recently, some new approaches to localization of molecular orbitals have been put forward. The natural localized orbitals of Weinhold et al. [10] take advantage of the natural population analysis [11]. The LMOs of Pipek and Mezey [12] are derived using Mulliken population analysis. Both of these methods rely on reference to atomic orbitals as prescribed by the basis set used in a particular calculation. Therefore, so obtained LMOs are not true molecular properties, since they cannot be derived from the wave function itself. This introduces a certain degree of arbitrariness and precludes the use for one-center basis sets, or basis sets with bond polarization functions or ghost atoms.

In the present paper, we propose a new localization procedure that is free of these disadvantages.

2.   The atomic overlap matrices

Let us assume that we were able to define *an atomic overlap matrix* (AOM) $\langle i | j \rangle_A$ with the following summation property:

$$\sum_A \langle i|j \rangle_A = \langle i|j \rangle = \delta_{ij} .$$ (3)

In eq. (3), the summation extends over N atoms A of the molecule under consideration. Definition of AOMs constitutes a partitioning of the overlap matrix. The quantity $\langle i|i \rangle_A$ defines the atomic population of the ith orbital on the atom A (the atomic occupancy). The quantity [12]

$$L_i = \sum_A \langle i|i \rangle_A^2 ,$$ (4)

is *the localization index* of the ith orbital. L_i^{-1} is equal to the effective number of atoms spanned by the ith orbital. If for all atoms A the eigenvalues of the respective AOM lie between 0 and 1, it follows from the Cauchy inequality that L_i^{-1} is bounded by the following inequalities:

$$1 \le L_i^{-1} \le N.$$ (5)

Taking the above observation into account, we propose a new localization scheme with the tensor

$$T_{klmn} = \sum_A \langle k|l \rangle_A \langle m|n \rangle_A .$$ (6)

In order to have a practical localization method that uses eq. (6), one has to specify $\langle i|j \rangle_A$. Several definitions are possible. In the approach by Pipek and Mezey [12], one uses Mulliken population analysis to determine AOMs. As explained in the introduction, this results in arbitrary partitioning of AOMs. Another potential possibility would be the use of recently introduced GAPT population analysis [13, 14]. However, we have found that the approach often yields negative electron populations for the localized orbitals.

The quantum theory of atoms in molecules as proposed by Bader [15] provides an elegant way of defining AOMs. The definition uses no references to the basis set used in the actual calculations. Hence, the resulting AOMs are true molecular properties. The relevant equation reads [16]

$$\langle i|j \rangle_A = \int_{\text{over basin of } A} \psi_i^*(r) \psi_j(r) \, dr.$$ (7)

The atomic basins are uniquely specified by gradient paths in the electron density. This yields "atoms" that satisfy several quantum-mechanical theorems, the virial theorem being one example. For more details, the interested reader should consult refs. [15, 16]. In table 1, we show the AOMs for the H_2O molecule.

Table 1

Atomic overlap matrices for the H_2O molecule (6-311G** basis set)

H_1				
0.000				
0.000	0.040			
0.000	0.064	0.109		
0.000	− 0.040	− 0.065	0.044	
0.000	0.000	0.000	0.000	0.006
O				
1.000				
0.000	0.920			
0.000	0.000	0.781		
0.000	0.079	0.000	0.911	
0.000	0.000	0.000	0.000	0.988
H_2				
0.000				
0.000	0.040			
0.000	− 0.064	0.109		
0.000	− 0.040	0.065	0.044	
0.000	0.000	0.000	0.000	0.006

Since in planar molecules $\langle \sigma \mid \pi \rangle_A = 0$ for all atoms A, one can block-localize the orbitals without affecting the localization sum L. This means that there is no σ/π mixing and no possible formation of banana bonds.

3.   Computational details

The CMOs were obtained using the GAUSSIAN-82 package [17]. All molecules had their geometries optimized within a given basis set. The Bader AOMs were computed using EXTREME/PROAIM system of programs [16]. In the following tables, all CMOs are listed in the order of increasing orbital energies.

4.   Results

The localized orbitals were computed for the H_2O, LiF, N_2, CO, $BH_3 \cdot$ CO and Li_2 molecules. Values of L_i^{-1} close to one indicate core or lone-pair orbitals. Those between one and two are characteristic for LMOs describing localized bonds and those larger than two correspond to delocalized bonds. In the H_2O molecule, the CMOs are already fairly localized, as indicated by a small change in the localization sum L when going from CMOs to LMOs (table 2). The first LMO is a core orbital

Table 2

Results for the H_2O molecule (6-311G** basis set)

(a) Canonical orbitals ($L = 4.296$)				
CMO i	Localization number L_i^{-1}	Atomic occupancies		
		H_1	O	H_2
1	1.000	0.000	1.000	0.000
2	1.177	0.040	0.920	0.040
3	1.575	0.109	0.782	0.109
4	1.197	0.044	0.912	0.044
5	1.024	0.006	0.988	0.006
Total		0.199	4.602	0.199

(b Localized orbitals ($L = 4.346$)				
LMO i	Localization number L_i^{-1}	Atomic occupancies		
		H_1	O	H_2
1	1.000	0.000	1.000	0.000
2	1.009	0.002	0.996	0.002
3	1.450	0.004	0.809	0.187
4	1.450	0.187	0.809	0.004
5	1.024	0.006	0.988	0.006
Total		0.199	4.602	0.199

Table 3

Results for the LiF molecule (6-31+G* basis set)

(a) Canonical orbitals ($L = 5.907$)			
CMO i	Localization number L_i^{-1}	Atomic occupancies	
		Li	F
1	1.000	0.000	1.000
2	1.022	0.989	0.011
3	1.010	0.005	0.995
4	1.043	0.021	0.979
5	1.010	0.005	0.995
6	1.010	0.005	0.995
Total		1.025	4.975

(b Localized orbitals ($L = 5.922$)			
LMO i	Localization number L_i^{-1}	Atomic occupancies	
		Li	F
1	1.000	0.000	1.000
2	1.015	0.993	0.007
3	1.000	0.000	1.000
4	1.046	0.023	0.977
5	1.010	0.005	0.995
6	1.010	0.005	0.995
Total		1.025	4.975

on the oxygen atom. The second and fifth LMOs are the σ and π lone pairs on the oxygen atom. Finally, the third and fourth LMOs describe two equivalent O–H bond orbitals. This interpretation of localized orbitals can be carried out entirely by considering their localization indices and the relevant atomic occupancies. As expected, the total atomic occupancies are invariant to the localization transformation.

Results for the LiF molecules are also clear-cut (table 3). In this case, the localization increases the sum L only very slightly. The set of LMOs comprises F and Li core orbitals ($i = 1$ and 2, respectively) together with one s and three p lone-pair orbitals on the F atom ($i = 3, 4, 5$ and 6). This is in accordance with the ionic Li^+F^- bond.

In the N_2 molecule, all CMOs are completely delocalized due to the molecular symmetry requirements (table 4). The localization transformation yields two core orbitals ($i = 1$ and 2), one σ bonding orbital ($i = 3$), two lone-pair orbitals

Table 4

Results for the N_2 molecule (6-31G* basis set)

(a) Canonical orbitals ($L = 3.500$)			
CMO i	Localization number L_i^{-1}	Atomic occupancies	
		N_1	N_2
1	2.000	0.500	0.500
2	2.000	0.500	0.500
3	2.000	0.500	0.500
4	2.000	0.500	0.500
5	2.000	0.500	0.500
6	2.000	0.500	0.500
7	2.000	0.500	0.500
Total		3.500	3.500

(b Localized orbitals ($L = 5.481$)			
LMO i	Localization number L_i^{-1}	Atomic occupancies	
		N_1	N_2
1	1.000	0.000	1.000
2	1.000	1.000	0.000
3	2.000	0.500	1.500
4	1.010	0.995	0.005
5	1.010	0.005	0.995
6	2.010	0.500	0.500
7	2.000	0.500	0.500
Total		3.500	3.500

($i = 4$ and 5) and two equivalent bonding π orbitals ($i = 6$ and 7). This is accompanied by a significant increase in L.

The situation in the isoelectronic CO molecules is quite similar (table 5). However, due to its heteronuclear character, the resulting LMOs are less delocalized, as reflected by smaller values of the corresponding localization indices. There are

Table 5

Results for the CO molecule (6-31+G* basis set)

(a) Canonical orbitals ($L = 5.948$)			
CMO i	Localization number L_i^{-1}	Atomic occupancies C	O
1	1.000	0.000	1.000
2	1.006	0.997	0.003
3	1.152	0.071	0.929
4	1.273	0.122	0.878
5	1.290	0.129	0.871
6	1.290	0.129	0.871
7	1.335	0.853	0.147
Total		2.301	4.699

(b Localized orbitals ($L = 6.249$)			
LMO i	Localization number L_i^{-1}	Atomic occupancies C	O
1	1.000	0.000	1.000
2	1.005	0.998	0.002
3	1.228	0.104	0.896
4	1.001	0.000	1.000
5	1.290	0.129	0.871
6	1.290	0.129	0.871
7	1.124	0.941	0.059
Total		2.301	4.699

two core orbitals ($i = 1$ and 2), one σ bonding orbital ($i = 3$), two equivalent π bonding orbitals ($i = 5$ and 6) together with lone pairs on the O ($i = 4$) and the C ($i = 7$) atoms. The C lone pair is slightly delocalized.

CMOs of the $BH_3 \cdot$ CO molecular complex are highly delocalized (table 6). The ninth orbital is spread across five atoms (three hydrogens, the B and the C atoms). The localization transformation brings about a substantial increase in L. There are eleven LMOs: The three first ones ($i = 1, 2$ and 3) are core orbitals of the O, C and B atoms. The fourth LMO is a σ bonding orbital of the C–O bond. There is an increased participation of the C atom in bonding compared to the isolated CO molecule. The fifth LMO is basically a lone-pair orbital on the C atom which is delocalized to some extent. The sixth LMO is clearly a lone-pair orbital on the O

Table 6

Results for the $BH_3 \cdot CO$ molecule (6-31G* basis set)

(a) Canonical orbitals ($L = 5.948$)

CMO i	Localization number L_i^{-1}	Atomic occupancies					
		B	H_1	H_2	H_3	C	O
1	1.000	0.000	0.000	0.000	0.000	0.000	1.000
2	1.006	0.000	0.000	0.000	0.000	0.997	0.003
3	1.008	0.996	0.001	0.001	0.001	0.001	0.000
4	1.163	0.000	0.000	0.000	0.000	0.076	0.924
5	2.789	0.054	0.036	0.036	0.036	0.406	0.432
6	2.885	0.068	0.057	0.057	0.057	0.235	0.527
7	1.375	0.003	0.007	0.002	0.002	0.146	0.840
8	1.375	0.003	0.001	0.005	0.005	0.146	0.840
9	5.147	0.131	0.190	0.190	0.190	0.260	0.040
10	3.233	0.128	0.031	0.380	0.380	0.048	0.034
11	3.212	0.124	0.499	0.148	0.148	0.048	0.034
Total		1.506	0.822	0.818	0.818	2.362	4.674

(b) Localized orbitals ($L = 8.843$)

LMO i	Localization number L_i^{-1}	Atomic occupancies					
		B	H_1	H_2	H_3	C	O
1	1.000	0.000	0.000	0.000	0.000	0.000	1.000
2	1.005	0.000	0.000	0.000	0.000	0.998	0.002
3	1.007	0.997	0.001	0.001	0.001	0.001	0.000
4	1.259	0.001	0.001	0.001	0.001	0.114	0.884
5	1.376	0.066	0.016	0.016	0.016	0.849	0.038
6	1.000	0.000	0.000	0.000	0.000	0.000	1.000
7	1.298	0.001	0.001	0.000	0.001	0.130	0.868
8	1.298	0.001	0.001	0.001	0.000	0.130	0.868
9	1.673	0.146	0.023	0.022	0.757	0.048	0.004
10	1.675	0.148	0.023	0.756	0.021	0.048	0.004
11	1.672	0.148	0.757	0.022	0.021	0.048	0.005
Total		1.506	0.822	0.818	0.818	2.362	4.674

atom. The next two LMOs ($i = 7$ and 8) are equivalent π orbitals of the C—O bond. Finally, the last three LMOs ($i = 9, 10$ and 11) are equivalent σ orbitals of the B—H bonds.

Two points regarding the above results should be emphasized. First, there is no LMO that would describe the B—C bonding. Second, there are some small differences in the localization indices and atomic occupancies for the equivalent LMOs. These are caused by a limited numerical accuracy associated with the discrete integration used in the Bader partitioning.

The interpretation of the localized orbitals resulting from the Bader definition of AOMs may run into difficulties in some rare cases where there is a discrepancy between the number of atomic basins and nuclei [18,19]. The Li_2 molecule is a classical example of such a case [17]. To localize the orbitals of the Li_2 molecule, one has to modify eqs. (3), (4) and (6) by including the empty basin (denoted by $(\cdot)$ in table 7) in the summation over atoms A. The localization yields two core orbitals ($i = 1$ and 2), and one orbital delocalized over two Li atoms and the empty basin ($i = 3$).

Table 7

Results for the Li_2 molecule (6-311G* basis set)

(a) Canonical orbitals ($L = 1.405$)				
CMO i	Localization number L_i^{-1}	Atomic occupancies		
		Li_1	Li_2	$(\cdot)$
1	2.000	0.500	0.500	0.000
2	2.000	0.500	0.500	0.000
3	2.469	0.224	0.224	0.552
Total		1.224	1.224	0.552

(b) Localized orbitals ($L = 2.405$)				
LMO i	Localization number L_i^{-1}	Atomic occupancies		
		Li_1	Li_2	$(\cdot)$
1	1.000	1.000	0.000	0.000
2	1.000	0.000	1.000	0.000
3	2.468	0.224	0.224	0.552
Total		1.224	1.224	0.522

5. Discussion

Our localization procedure has several advantages over other schemes: First, like the FB localization it is computationally an M^3 procedure. Second, it yields the localization indices and atomic occupancies for both CMOs and LMOs that provide useful tools for interpretation of the results. In fact, the character of the localized orbitals can be inferred directly from these quantities without the need of plotting LMOs. The present scheme extends the applicability of the topological theory of atoms in molecules. Finally, the σ/π-separation is built into our approach. The LMOs do not break symmetry imposed by the nuclear framework of molecules.

In the light of the above results, one should wish that a more efficient numerical integration technique be incorporated into the Bader scheme, yielding a better accuracy of the computed AOMs. At the present time, this accuracy is limited to approximately 3 digits.

We devote the final comment to the relation between the present approach and the localization procedure first proposed by Perkins and Stewart [20] and corrected by Paniagua et al. [21]. If one defines the atomic overlap matrix as

$$\langle i|j\rangle_A = \sum_{k \in A} C_{ki}^* C_{kj} , \qquad (8)$$

where C is the matrix of MO coefficients, one retrieves the Perkins–Stewart–Paniagua scheme. Our localization index becomes equivalent to the "bond centre number" [20]. One should remember, however, that the definition (8) is relevant only for the semi-empirical methods based on the ZDO approximation.

References

[1] V. Fock, Z. Phys. 61(1930)126;
     P. Surján, in: *Theoretical Models of Chemical Bonding*, Part 2, ed. Z.B. Maksić (Springer, 1990), p. 205.
[2] J.W. Linnett, J. Amer. Chem. Soc. 83(1961)2643.
[3] C. Edmiston and K. Ruedenberg, J. Chem. Phys. 43(1965)S97.
[4] J.M. Foster and S.F. Boys, Rev. Mod. Phys. 32(1960)300.
[5] W. von Niessen, J. Chem. Phys. 56(1972)4290.
[6] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. 35(1963)457.
[7] J.M. Leonard and W.L. Luken, Theor. Chim. Acta 62(1982)107.
[8] R. Hart, J. Chem. Phys. 43(1965)S115.
[9] W. von Niessen, Theor. Chim. Acta 29(1973)29.
[10] A.E. Reed and F. Weinhold, J. Chem. Phys. 83(1985)1736.
[11] A.E. Reed, R.B. Weinstock and F. Weinhold, J. Chem. Phys. 83(1985)735.
[12] J. Pipek and P.G. Mezey, J. Chem. Phys. 90(1989)4916.
[13] J. Cioslowski, Phys. Rev. Lett. 62(1989)1469.
[14] J. Cioslowski, J. Amer. Chem. Soc. 111(1989)8333;
     J. Cioslowski, P.J. Hay and J.P. Ritchie, J. Phys. Chem. 94(1990)148.
[15] For a review on quantum theory of atoms in molecules, see:
     R.F.W. Bader and T.T. Nguyen-Dang, Adv. Quant. Chem. 14(1981)63;
     R.F.W. Bader, T.T. Nguyen-Dang and Y. Tal, Rep. Progr. Phys. 44(1981)893.
[16] F.W. Biegler-König, R.F.W. Bader and T.H. Tang, J. Comput. Chem. 3(1982)317.
[17] GAUSSIAN-82, J.S. Binkley, M.J. Frisch, D.J. DeFrees, K. Raghavachari, R.A. Whitehead, H.B. Schlegel, E.M. Fluder and J.A. Pople, Department of Chemistry, Carnegie-Mellon University, Pittsburgh, PA. The CRAY X-MP48 implementation by Dr. R. Martin (LANL).
[18] C. Gatti, P. Fantucci and G. Pacchioni, Theor. Chim. Acta 72(1987)433;
     W.L. Cao, C. Gatti, P.J. MacDougall and R.F.W. Bader, Chem. Phys. Lett. 141(1987)380.
[19] R.F.W. Bader, Y. Tal, S.G. Anderson and T.T. Nguyen-Dang, Israel J. Chem. 19(1980)8.
[20] P.G. Perkins and J.J.P. Stewart, J. Chem. Soc. Faraday Trans. 2 79(1983)37.
[21] J.C. Paniagua, A. Moyano and L.M. Tel, Int. J. Quant. Chem. 26(1984)383.