Theor Chim Acta (1993) 86: 379–389

Theoretica Chimica Acta

Localization maps by orbital partitioning of the electron density*

Péter R. Surján[1,2], János Pipek[2], and Béla Paizs[1]

[1]Laboratory for Theoretical Chemistry, Eötvös University, H-1518 Budapest 112, P.O. Box 32, Hungary

[2]Quantum Theory Group, Institute of Physics, Technical University, H-1521 Budapest Budafoki út 8, Hungary

Received June 30, 1992/Accepted February 1, 1993

Summary. We define a localization measure for one-determinantal wave-functions based on the partitioning of the total electron density to orbital contributions. The set of occupied orbitals is the more localized the fewer terms are necessary to describe the total density. This measure varies from point to point in space which leads to characteristic localization maps for molecules.

Key words: Localization maps – Orbital partitioning – Electron density

1. Introduction

Localization and localizability of many-electron wavefunctions continues to attract investigators from various fields of solid-state physics [1] and quantum chemistry [2]. Many localization criteria have been proposed [3–7] each of which defines a specific set of localized orbitals. The corresponding localization procedures are based on the fact that the one-determinantal wavefunction is invariant to a unitary transformation of the one-electron orbitals which makes it possible to require the orbitals to obey an extra criterion in addition to the fulfillment of the Hartree–Fock equations (for a recent review, see [8]). The spatial criterion of Foster and Boys [3] maximizes the product $\prod_{i<j} R_{ij}^2$, R_{ij} being the distance between the centroids of MOs i and j. Three other criteria were introduced by Edmiston and Ruedenberg [4]. Their energetic criterion maximizes the self-repulsion of the molecular orbitals (MOs), i.e., the sum $\sum_i (ii|r^{-1}|ii)$. Their second criterion, which maximizes the sum $\sum_i (ii|-r^2|ii)$, resulted in localized MOs similar to those obtained by Foster and Boys [3], and it turned out [5] that it was equivalent to maximizing the sum $\sum_{i<j} R_{ij}^2$. The density-localized orbitals which were introduced by Edmiston and Ruedenberg [4] as their third criterion, and were applied subsequently by von Niessen [6], are defined by maximizing the sum:

$$\sum_{i=1}^{N} \int \left| \int \psi_i^*(r_0)\delta(r-r_0)\psi_i(r_0)\, d^3r_0 \right|^2 d^3r \tag{1}$$

---

* Supported in part by the grant Hungarian Research Fund OTKA 517/1990

A related concept is that of the localization *measure*. This is defined as a number which characterizes the extent of the localization for a given wavefunction. The values of the localization functionals, e.g., can serve for this purpose. One can also define localization indices which measure the extent of localization for a given orbital. If the latter is expanded in an orthogonal basis set of N functions, i.e.:

$$\psi_i(r) = \sum_{\mu=1}^{N} C_{i\mu}\chi_\mu(r), \tag{2}$$

the simplest choice for a localization index is:

$$l_i = \sum_\mu C_{i\mu}^4 \bigg/ \left(\sum_\mu C_{i\mu}^2\right)^2 \tag{3}$$

It is easy to show that l_i fulfills the inequalities:

$$\frac{1}{N} \leqslant l_i \leqslant 1.$$

For maximum localization $l_i = 1$, and for minimum localization $l_i = 1/N$. This index has been used in quantum chemistry [9] and widely applied in the solid state literature [10]. Localization procedures based on its use have been developed by Pipek and Mezey [11] and Perkins [12]. In [11], l_i is defined also for nonorthogonal basis sets applied in *ab initio* theories.

2. Localization index for general distributions

In this section we summarize the mathematics of the localization index used in this study in a somewhat more precise manner. From the intuitive point of view, the localization index of any general distribution (Q_i: $i = 1, 2, \ldots, N$) of some quantities should give the number of indices i where the distribution "essentially" differs from zero. To form such an index, we should define a measure of the 'support' of the distribution in a generalized sense. It is to be emphasized that the number of contributing indices, d, will not be computed "numerically" by counting simply the number of indices for which Q_i exceeds some threshold. Instead, we use a unique expression which takes into account the relative weights of all contributions without any truncation.

The suitable definition has been given in Ref. [13]:

$$d^{-1} = \sum_{i=1}^{N} \left(Q_i \bigg/ \left(\sum_{j=1}^{N} Q_j\right)\right)^2 \tag{4}$$

The properties of the index d satisfy the expectations outlined above as:

$$0 < d \leqslant N \tag{5}$$

and, if the distribution has equal magnitudes on s indices and zero otherwise, it easily follows from Eq. (4) that $d = s$. Moreover, it is also possible to show that for positive distributions ($Q_i \geqslant 0$ for all i) one has:

$$1 \leqslant d \leqslant N$$

Accordingly, a small value for d indicates a larger localization, while large d reflects significant delocalization.

An additional advantage of Eq. (4) is that the *localization index d* is relatively insensitive against small changes of the distribution $\{Q_i\}$ [13]. On the other hand, d is generally noninteger, giving in this way the approximate number of indices where the distribution is 'localized'. By this definition, the concept of *localization* achieves a broader sense that we will apply throughout this paper.

We would like to mention here that the description of distributions (wavefunctions) by a single localization index gives only the most substantial information. A more detailed characterization of the internal structure (shape analysis) is also possible by introducing the *structural entropy* of the distribution $\{Q_i\}$ [14]. This refinement will not be utilized in the present study, however.

3. Localization maps

The total electron density in the Hartree–Fock case is given by:

$$\varrho(r) = \sum_{i=1}^{N} \psi_i^*(r)\psi_i(r) = \sum_{i=1}^{N} \varrho_i(r) \tag{7}$$

One may ask how localized the electron density is at a given point r in terms of molecular orbitals, that is, how many MOs contribute significantly to $\varrho(r)$. In the spirit of the previous section, we identify $Q_i = \varrho_i(r)$, and define:

$$M(r) = \left(\sum_{i=1}^{N} \varrho_i^2(r) \Bigg/ \left(\sum_{j=1}^{N} \varrho_j(r) \right)^2 \right)^{-1} \tag{8}$$

The number of significant MOs may vary from point to point in space indicating that the extent of localization of a given set of MOs, as defined by Eq. (8), varies, too. The resulting *localization map* carries interesting information on the electronic structure represented by a given set of molecular orbitals. It indicates how many molecular orbitals contribute to the total density at a given point of space.

As $M(r)$ is *not* invariant against unitary transformation of orbitals, canonical and localized sets of orbitals, as obtained by usual localization criteria, will result in different localization maps. Similarities and differences of such maps add further interesting pieces of information. Typically, one expects that canonical $M(r)$ values are larger than the localized ones in the regions of chemical bonds, and the former are close to unity only in the inner shell regions where the electron density is dominated by a single orbital. On the other hand, a localized set of MOs will produce $M(r)$ values close to 1 at chemical bonds.

The analysis can, of course, be performed for any set of MOs obtained by unitary transformation. For example, one could also use maximally delocalized orbitals which result from the canonical ones by putting the localization procedure in reverse [15, 16].

The connection between the above localization maps and the density localization procedure is worth mentioning. As a matter of fact, one may ask whether it is possible to define a specific set of MOs so that they make $M(r)$ stationary, i.e. to use $M(r)$ as a localization functional.

It is easy to see that at an arbitrary point r_0 to MOs can be "fully localized" in the sense that it is always possible to find unitary transformed MOs φ_i with the property that only one of them will contribute to the density at r_0. Mathematically this can be expressed by $\varrho(r_0) = |\varphi_1(r_0)|^2$ and $\varphi_i(r_0) = 0$ for $i = 2, 3, \ldots, N$. The proof of the existence of the appropriate unitary matrix U

is trivial, considering that all $N \times N$ matrices transforming the N-vector

$$(\psi_1(r_0), \ldots, \psi_N(r_0))$$

to

$$(\sqrt{\varrho(r_0)}, 0, 0, \ldots, 0)$$

will do the job. Note, however, that the transformation matrix $U(r_0)$, as well as the resulting "fully localized" orbitals, will be different for each point r_0.

In order to define a global criterion, one has to 'average' somehow the r-dependent map $M(r)$. Defining an average by integrating the numerator and the denominator of $M(r)$ separately, i.e.:

$$\overline{M} = \frac{\int \left(\sum_j \varrho_j(r) \right)^2 d^3r}{\int \sum_j \varrho_j^2(r)\, d^3r} \equiv \frac{\int \varrho^2(r)\, d^3r}{\int \sum_j \varrho_j^2(r)\, d^3r},$$

we observe that the numerator of $\overline{M}$ is determined by the electron density, thus it is a unitary invariant constant. Consequently, maximizing $\overline{M}$ is equivalent to minimizing the denominator:

$$\int \sum_j \varrho_j^2(r)\, d^3r \equiv \int \sum_j |\psi_j(r)|^4\, d^3r$$

which is exactly the same as the density localization criterion, as it becomes apparent by carrying out the integration in Eq. (1) over r_0. Accordingly, we can state that the localization map $M(r)$ is a local version of the global density criterion in this sense.

Finally, we have to note that an analysis of a similar problem has been carried out by Ruedenberg et al. [9] in their work on chemical analysis of the FORS (fully optimized reaction space) model.

4. Examples

As an illustration we have computed the localization maps for the water and benzene molecules in 4-21G basis set [17] both for the canonical MOs and for orbitals localized in Boys' procedure [5]. To check the basis dependence of the following results, we repeated the water calculation in 6-311G** basis set [18], as well as in two larger bases which contained also diffuse functions (vide infra). In each case, the nuclear geometry has been optimized at the 4-21G level. The computations were performed by the *ab initio* packages MUNGAUSS [19] and TEXAS [20]. To evaluate the density and the localization indices, we used a separate program GEDASP [21].

4.1. Water

We start the discussion with the water results. Figure 1.a and b show the localization maps for canonical and Boys' localized MOs, respectively, obtained in a simple split-shell basis set. The oxygen atom is placed into the origin, while the hydrogens lie in the plane of the figure at the coordinates (±1.473, 1.069 in

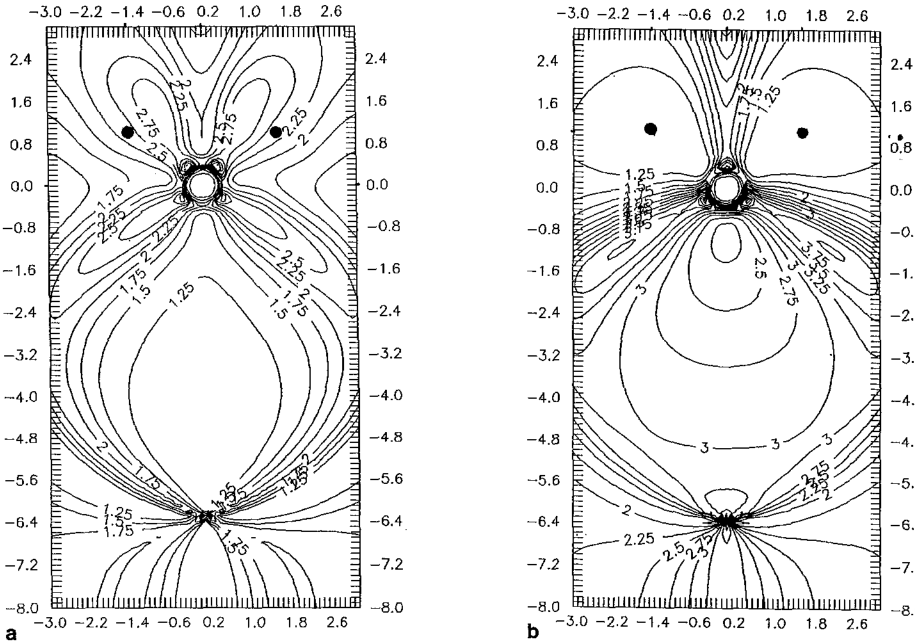

Fig. 1a,b. Localization maps for the water in the molecular plane. Units of distances on the x and y axes are given in Bohrs (a.u.). Heavy dots indicate the position of the hydrogen atoms while the oxygen is situated in the origin. Basis set: 4-21G (**a**) canonical orbitals (**b**) Boy's localized orbitals

a.u.) The curves of the figure follow the lines where $M(r)$ is constant ('isolocalization contours'). The maps are normalized so that the value of $M(r)$ gives the number of contributing doubly occupied orbitals.

At the origin, the $M(0)$ value is practically 1. This clearly indicates that around the oxygen nucleus, the electron density is dominated by the single inner shell $1s$ orbital, so the localization is perfect. This is a common feature of the canonical and localized representations. Moving away from the nucleus, the $M(r)$ values increase sharply and one can see a characteristic localization map for the valence shell which is quite different for the canonical and localized orbitals. In the OH bond region, one has $M^{\mathrm{can}}(r) \cong 2.8$ while $M^{\mathrm{loc}}(r) \cong 1.0$. This is easy to understand since in the canonical case the orbitals $2a_1$, $1b_1$, and $3a_1$ dominate, while a single MO is responsible for the bond if the orbitals are localized. Between the two hydrogens, at the midpoint, one should consider the canonical MOs $2a_1$ and $3a_1$ (the contribution of $1b_1$ vanishes due to symmetry), while the two OH bonding MOs have an equal contribution in the localized case. Accordingly, $M^{\mathrm{can}}(r) \cong M^{\mathrm{loc}}(r) \cong 2$ in this region.

The canonical $M(r)$ values never exceed 3. This is because, from the 5 occupied MOs, the $1s$ orbital sharply decreases due to the increased exponent of the dominating $1s$ AOs, and does not contribute to the valence region. The π lone pair has no effect in the molecular plane either. On the other hand, there are domains where $M^{\mathrm{loc}}(r)$ is nearly 4, indicating that we are at a roughly equal distance from the centroids of the 4 valence orbitals (two lone pairs and two OH bonds).

The lone pair region is also interesting. On the symmetry axis we find that $M^{\mathrm{can}}(r) \cong 1$ while $M^{\mathrm{loc}}(r) \cong 2$. The latter value comes from the two "rabbit ear" lone pairs, characteristic to the Boys' procedure

Passing the valence region one finds a critical point on the molecular symmetry axis at approximately 3 Å from the oxygen. In this region the isolocalization curves cumulate, and one can see a structure similar to the O atom. This feature is present both on the localized and canonical maps. To illustrate the similarity of this critical point and an atomic region, we show Fig. 2 in which the three-dimensional diagrams of the $M(r)$ surfaces are compared.

The origin of this critical point on the localization maps merits further discussion. For very large distances, one expects $M(r)$ to tend to a constant value in any direction, since for large r the orientation of the MOs is not important. In the critical region the $\varrho(r)$ values themselves are already very small, thus we performed careful numerical checks to exclude the possibility of round-off errors.

The calculations have been repeated in the (6-311G**) basis set, which is of triple-ζ quality in the valence shell and is augmented with polarization functions. The results for $M(r)$ in this basis are not reported here as they are quite similar to those in Figs. 1 and 2. This indicates the similar character of the 4-21G and 6-3111G** basis sets: both consist of a small number of primitives and are designed to describe the valence shell region. Indeed, the electron density is more-or-less well described within the valence region already at the 4-21G level,

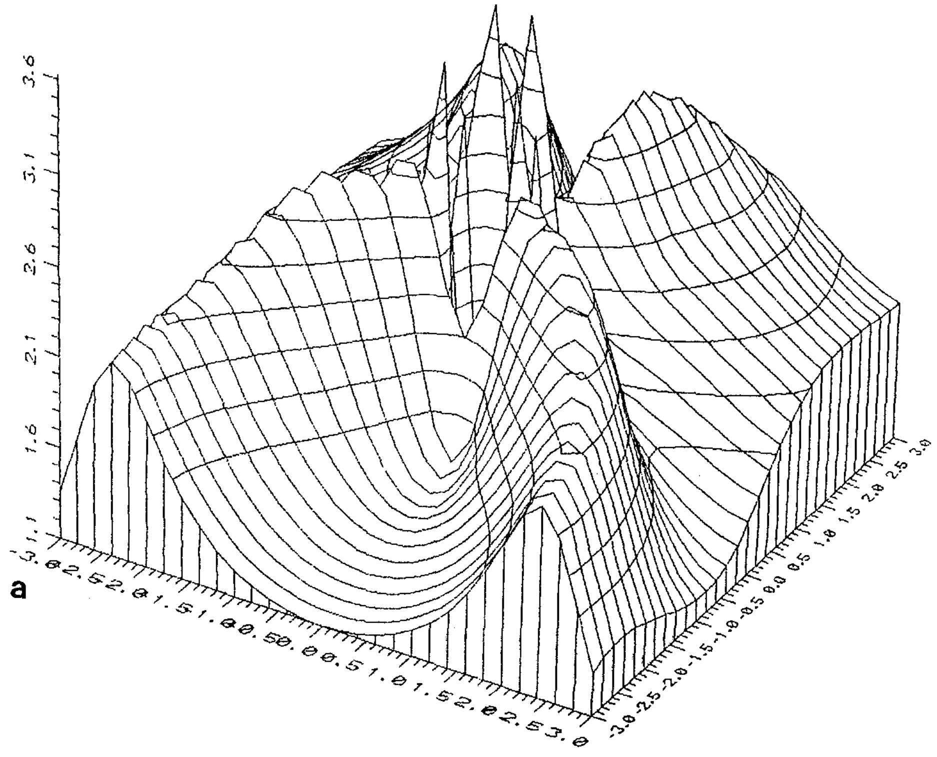

Fig. 2a

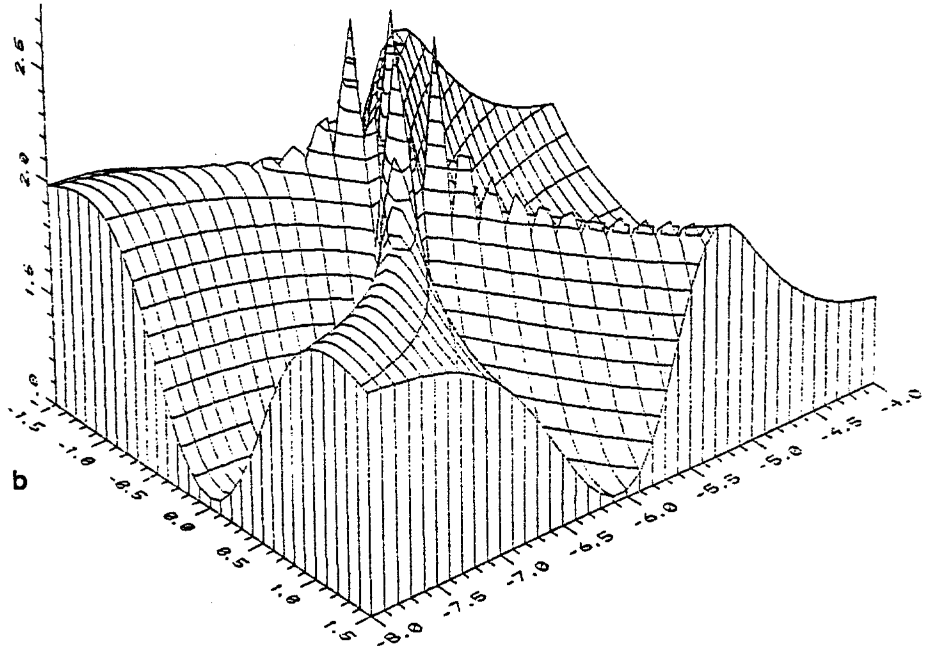

Fig. 2b

Fig. 2. Three-dimensional plots of $M(r)$ surfaces for water at around the oxygen (**a**) and at a point 3 Å from O in the lone pair direction on the molecular symmetry axis (**b**). Basis set: 4-21G

while the diffuse functions, which are expected to be important in the far-lying regions, are still lacking in 6-311G**.

To see the role of diffuse functions, we performed calculations in two larger bases, for which the $M(r)$ maps are presented in Figs. 3 and 4. The two bases can be specified as follows (for details, see [22]):

Basis set A): [5s/3p] for oxygen, contracted by the scheme [53111/412]
[4s] for hydrogen, contracted according to the scheme [5311]

Basis set B): The same as A) for oxygen;
[5s] for hydrogen, contracted according to the scheme [43111]

That is, the two based differ only in the contraction scheme for H. It is important that the smallest s-type exponent is 0.13 for O, and 0.028 for H; the smallest p-type exponent on the oxygen is 0.072.

The effect of the diffuse functions is remarkable. The localization maps $M(r)$ for water in basis sets A) and B) are plotted in Figs. 3 and 4, respectively. The maps in the vicinity of atoms are very similar to those in Fig. 1. However, the results differ significantly in the far-lying regions; though the remote critical point is still present at nearly the same position, its structure is much less similar to the atomic region. Moreover, the two plots in bases A) and B) differ also from each other at large distances from the oxygen. This indicates the basis-set sensitivity of the structure of the density in the outer regions. The specific structure of the remote critical point seen in Figs. 1 and 2 should be considered to result from the artifact of using small basis sets. Accordingly, the present

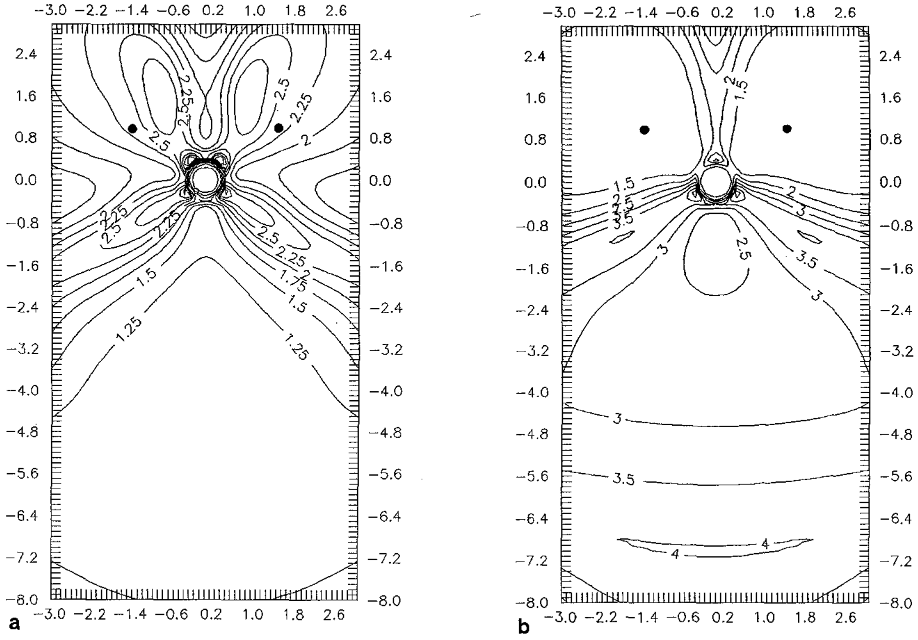

Fig. 3. The same as Fig. 1, but in basis set A) (see text)

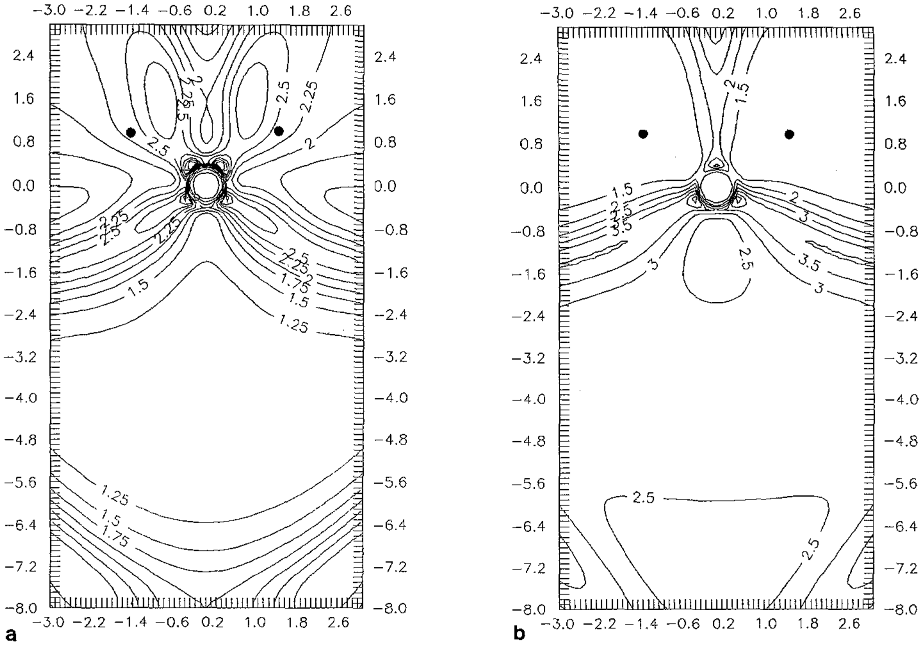

Fig. 4. The same as Fig. 1, but in basis set B) (see text)

localization maps can be applied to judge the quality of basis sets in far-lying regions where other (e.g. energetic) criteria are not applicable.

4.2. *Benzene*

As a second illustration we show Fig. 5 in which the localization maps for benzene are depicted in the molecular plane. Most features are similar to the water case. It is interesting, however, that the $M(r)$ maps break the 6-fold rotational symmetry of the molecule. This is a consequence of the fact that only complex orbitals can span the irreducible representation of benzene. Using real MOs, one mixes the degenerate MOs spanning the two-dimensional representation. This mixing does not affect the density which is a unitary invariant but it affects $M(r)$.

Regardless of this symmetry-breaking feature, Fig. 5 offers an interesting insight into the electronic structure of benzene. First of all, the large values of $M(r)$ in the canonical case (Fig. 5a) indicate that 5–8 molecular orbitals contribute to the σ density all over the valence region. (Note that we see merely

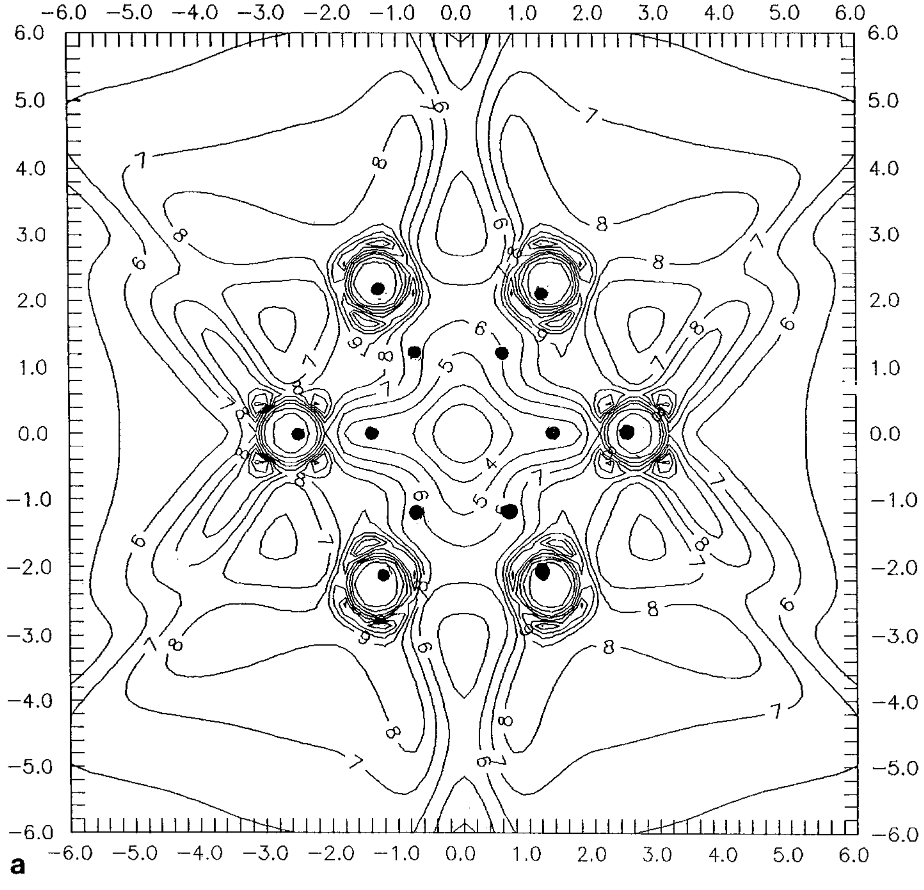

Fig. 5a

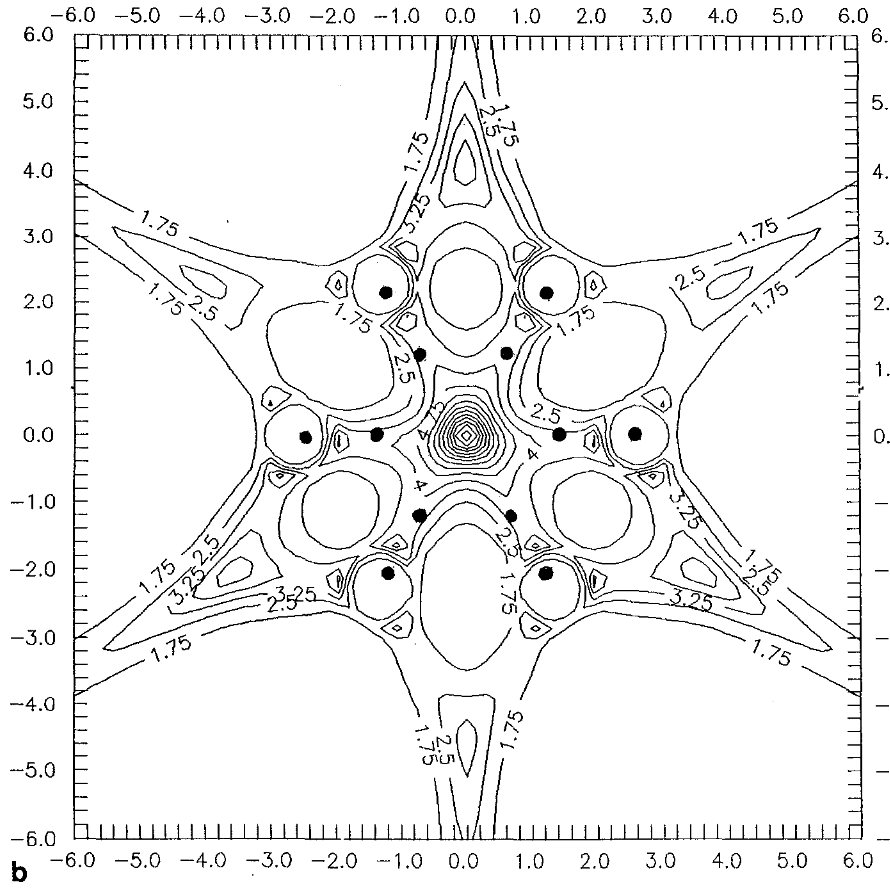

Fig. 5b

Fig. 5. Localization maps for benzene using real canonical (**a**) and Boys' localized (**b**) orbitals in 4-21G basis set. *Heavy dots* indicate the position of the atoms

the σ orbital contributions, as the antisymmetric π MOs do not affect the density in the molecular plane.) Localizing the orbitals by Boys' procedure, the map changes significantly, but the $M(r)$ values remain close to 2 in the region of the CC bonds. This is because the σ and π MOs are mixed now. In the center of the molecule, the M^{loc} values are much larger (indeed, they have a cusp at $x = y = 0$) indicating that localizing spatially the MOs somewhere may result in significant delocalization elsewhere.

References

1. Brezini A, Zekri N (1992) Phys Status Solidi B 169:253
2. Chalvet O, Daudel R, Diner S, Malrieu JP (eds) (1976) Localization and delocalization in quantum chemistry, Vols I–II. Reidel, Dordrecht
3. Foster JM, Boys SF (1960) Rev Mod Phys 32:300

4. Edmiston C, Ruedenberg K (1963) Rev Mod Phys 35:457; (1965) J Chem Phys 43:S97; Ruedenberg K (1965) in: Sinanoglu O (ed) Modern quantum chemistry. Academic, NY, p 85
5. Boys SF (1966) in: Löwdin PO (ed) Quantum theory of atoms, molecules and the solid state. Academic, NY
6. Niessen W Von (1970) J Chem Phys 56:4290; (1972) Theor Chim Acta 27:9
7. Magnasco V, Perico A (1967) J Chem Phys 47:971; (1968) ibid 48:800
8. Surján PR (1990) in: Maksic ZB (ed) Theoretical models of chemical bonding, Part 2, Springer, Berlin
9. Ruedenberg K, Schmidt MW, Gilbert MM (1982) Chem Phys 71:51
10. Bell RJ, Dean P (1970) Discuss Faraday Soc 50:55
11. Pipek J, Mezey PG (1989) J Chem Phys 90:4916
12. Perkins PG, Stewart JPS (1982) J Chem Soc Faraday II 78:285
13. Pipek J (1989) Int J Quant Chem 36:487
14. Pipek J, Varga I (1992) Phys Rev A 46:3148
15. Pipek J, Mezey PG (1988) Int J Quant Chem S22:1
16. Mezey PG (1990) in: Lipkowitz KB, Boyd DB (eds) Reviews in comput chemistry. VCH Publ, NY, p 265
17. Pulay P, Fogarasi G, Dang F, Boggs JE (1979) J Am Chem Soc 101:2550
18. Krishnan R, Binkley JS, Seeger R, Pople JA (1980) J Chem Phys 72:650
19. Program MUNGAUSS: Poirier RA, Peterson MR, Dept Chemistry, Memorial Univ, St. John's, Nfld, Canada
20. Program TEXAS: Pulay P (1979) Theoret Chim Acta 50:299
21. Program GEDASP: Pipek J (1989) Dept Chemistry, Univ of Saskatchewan Saskatoon, Canada
22. Poirier R, Kari R, Csizmadia IG (1985) Handbook of Gaussian basis sets. Elsevier, Amsterdam