Theor Chim Acta (1987) 72: 13-26

Theoretica Chimica Acta
© Springer-Verlag 1987

Localized virtual and occupied molecular orbitals

Resolution of the difficulties with use of local orbital population as the criterion for localization

Michel Rajzmann[1], **Bruno Brenier**[2] **and Keith F. Purcell**[3]

[1]Laboratoire de Chimie Inorganique Moleculaire, Faculté des Sciences de Saint-Jerome,
Rue Henri Poincaré, F-13397 Marseille Cédex 4, France
[2] Compagnie Internationale de Service en Informatique, Applications Scientifiques et Techniques (CFAO), Les Bureaux Parc du Griffon, 590 Route de la SEDS, F-13127 Vitrolles, France
[3] Department of Chemistry Kansas State University, Manhattan, KS 66502, USA

(Received February 6/Accepted March 8, 1987)

This report describes the generation of localized from canonical molecular orbitals such that the method (1) be consistently applicable to occupied bond, lone pair and unoccupied orbitals and (2) permit symmetry related orbitals in molecules of two-fold or higher symmetry. Minimization of populations completely external to each local pair region effectively meets these criteria. Finally, conventional strategy for finding the global extremization point is costly in time and memory to implement; a much more efficient numerical search procedure for the global extremum is described. Results for ethylene, butadiene and benzene are presented to clarify the difficulties and their resolution.

Key words: Localized orbitals—Virtual orbitals—Overlap population

1. Introduction

Molecular orbital functions for a given molecule exist in an infinity of forms; for a closed shell molecule, the infinity of Slater determinants which can be written are related simply by phase factors [1]. In practice, only two types are useful: the canonical type (that which arises directly from diagonalization of the Fock matrix under the Mulliken-Roothan procedure) and the localized type (those which arise when the amplitudes of the individual waves are maximized in certain molecular regions and minimized elsewhere).

The canonical type has been found to be useful for interpretation of reaction mechanisms [2], of molecular spectroscopic properties [3] and of certain aspects of molecular structure [4]. The localized forms have been found useful for interpretations of reaction mechanisms [5] and of classical (localized) concepts of chemical bonding (the electron density between two atomic centers, bond polarity [6], ring strain [7], ligand donor properties in conventional acid base [8] and transition metal complexes [9] and (when the intramolecular stereochemical environment of an interatomic link is fixed) justification of the findings that the magnitudes of certain physical properties of the link can be understood in terms of additive local effects [10]). From a theoretical point of view, an advantage of a localized orbital basis is found in the study of configuration interaction, where useful results can be obtained from a configuration interaction matrix of reduced size [11, 12].

One obtains the localized molecular orbitals from the canonical orbitals by carrying out a specific unitary transformation of the latter such that the total wave function is invariant. Existing definitions of the transformation are based on localization criteria which may be categorized in either of two groups. The intrinsic criteria most widely used are electron/electron repulsion (Edminston and Rudenberg [13]), interelectron distance (Boys [14]) and interbond electron density (von Niessen [15]). The extrinsic criteria are represented by Magnasco and Perico's use of local orbital populations (atomic or overlap) to generate the transformation matrix [16].

While there have been many studies of localized occupied orbitals, there are few literature reports of studies of localized virtual orbitals [8, 12, 17]. Similarly, few studies have included molecules with lone pair electrons. In this article we address the problem of defining a single extrinsic localization function which can be applied to both virtual and occupied orbitals in molecules with both bond and lone pair orbitals. We are led to a new localization function which consistently avoids the difficulties encountered with the conventional overlap population criterion for localization when it is applied to molecules possessing lone pair electrons and to virtual orbitals. Implementation of this function requires a general, numerical search procedure for its extremization, which turns out to be very efficient and implicitly yields extrema satisfying the second-order condition.

This approach to orbital localization can be used with all overlap-normalized orbital functions, including the zero-differential-overlap methods when the canonical orbitals have been transformed to their overlap-normalized equivalents [18]. It is very easy to apply since no new integrals need be evaluated; the only data required are the canonical LCAO coefficients and the overlap integrals. We use here the extended Hückel theory [19] canonical orbital functions because they are the least expensive to obtain of the overlap-normalized methods.

2. Results and discussion

2.1. Localization functions derived from local orbital populations

Following Magnasco and Perico [16], a local total-population for molecular orbital ϕ_i is defined as P_i

$$P_i = 2 \sum_{\mu\in\Gamma_i} \sum_{\nu\in\Gamma_i} C_{i\mu}C_{i\nu}S_{\mu\nu} = 2 \sum_{\mu\in\Gamma_i} C_{i\mu}^2 + 2 \sum_{\mu\in\Gamma_i} \sum_{\substack{\nu\in\Gamma_i \\ \mu\neq\nu}} C_{i\mu}C_{i\nu}S_{\mu\nu} = A_{\Gamma_i} + B_{\Gamma_i}, \tag{1}$$

where ao's μ, ν constitute the set Γ_i which defines the bond or lone pair region of interest. The set of N localized orbitals gives as the localization function P

$$P = \sum_{i=1}^{N} P_i \tag{2}$$

which is to be extremized. P_i is called the local total-population because it represents the one- and two-center electron density contribution for molecular orbital ϕ_i to the bond or lone pair defined by the local set of atomic orbitals Γ_i. The $c_{i\mu}$ are ao coefficients for the set Γ_i and the $S_{\mu\nu}$ are the ao overlap integrals. For LCAO mo's normalized without overlap, as in the zero-differential-overlap methods, only the one-center term A_{Γ_i} survives. P_i in this case is always >0 and <2, so that localization of all bond, lone and anti-bonding orbitals is uniformly achieved by maximization of P in every case.

For LCAO mo's normalized with overlap integrals, $B_{\Gamma_i} \neq 0$. This has interesting consequences for P_i and P which can best be understood by considering the partitioning of the global density function for molecular orbital ϕ_i:

$$\begin{aligned} 2 = 2\int \phi_i^2 d\tau &= 2 \sum_{\mu\in\Gamma_i} C_{i\mu}^2 + 2 \sum_{\mu\in\Gamma_i} \sum_{\substack{\nu\in\Gamma_i \\ \nu\neq\mu}} C_{i\mu}C_{i\nu}S_{\mu\nu} \\ &\quad + 2\left[\left\{ \sum_{\mu\in\Gamma_i} \sum_{\nu\notin\Gamma_i} + \sum_{\mu\notin\Gamma_i} \sum_{\nu\notin\Gamma_i} \right\} C_{i\mu}C_{i\nu}S_{\mu\nu}\right] \\ &= A_{\Gamma_i} + B_{\Gamma_i} + O_i \\ &= P_i + O_i, \end{aligned} \tag{3}$$

where only A_{Γ_i} and B_{Γ_i} are purely local terms; O_i contains both partly- and fully-remote terms.

Au début, one has a choice of four terms to extremize. Of these, only A_{Γ_i} is necessarily >0; B_{Γ_i} will be positive (negative) for local bonding (antibonding); O_i can be positive (net remote bonding) or negative (net remote antibonding) without restriction and its sign will vary from mo to mo.

Depending on the sign of the non-local terms (O_i), P_i can in principle be $>$ or <2 at any point during the localization process. Since, in P, one encounters both $P_i > 2$ and $P_i < 2$, it is impossible to define a single localization procedure by the customary method of successive two-dimensional

$$\cot 2\theta = \frac{\sum_{\mu\in\Gamma_i} \sum_{\nu\in\Gamma_i} [C_{i\mu}C_{i\nu} + C_{j\mu}C_{j\nu}]S_{\mu\nu} + \sum_{\mu\in\Gamma_j} \sum_{\nu\in\Gamma_j} [C_{j\mu}C_{j\nu} + C_{i\mu}C_{i\nu}]S_{\mu\nu}}{\sum_{\mu\in\Gamma_i} \sum_{\nu\in\Gamma_i} [C_{i\mu}C_{j\nu} + C_{j\mu}C_{i\nu}]S_{\mu\nu} - \sum_{\mu\in\Gamma_j} \sum_{\nu\in\Gamma_j} [C_{i\mu}C_{j\nu} + C_{j\mu}C_{i\nu}]S_{\mu\nu}}$$

rotations by the angle θ_{ij} in the orbital vector space to make each $P_i \rightarrow 2$, from either direction.

In practice, we have found that *maximization* of P for *occupied* orbitals by the successive orbital-pair rotation technique leads to successful convergence of the localization process. Even though it cannot be proven that each P_i must, in general, be <2, the successes of P maximization depend on the fact that, as the localization proceeds, eventually all $(A_{\Gamma_i}+B_{\Gamma_i})$ *and* O_i become >0. In practice, difficulties have been encountered, however, with some $P_i>2$ and other $P_i<2$ in the localization of *virtual* orbitals (see Table 1). In one attempt to avoid the attendant complications for both occupied and virtual mo's, we adopted a numeric method [20] (see the appendix) for searching for a minimized P', using the following definition of P'_i for both occupied and virtual mo's.

$$P'_i=|2-P_i|=|O_i| \quad \text{and} \quad P'=\sum_i P'_i=\sum_i |O_i|. \tag{4}$$

This method can lead to local extrema on the localization surface characterized by sets of localized orbitals with local total-populations near 2 for each P_i, however, some of the resulting orbitals are not well localized and orbitals expected to be chemically equivalent are not found to be so (see Table 2). The problem appears to derive from the fact that O_i is not explicitly treated by the localization requirement, so that O_i can approach zero (*from either direction*) by near cancellation of positive and negative components (these components are of two types: $O_i^{l,r}$; local/remote, where $\mu \in \Gamma_i$, $\nu \notin \Gamma_i$, and O_i^r; purely remove, where μ, $\nu \notin \Gamma_i$). When these components are large in magnitude the "localized" orbitals are in fact still "delocalized".

In an attempt to eliminate this delocalization problem, we studied the use of P''_i, defined as

$$P''_i=|2-P_i|+\sum_{\mu\notin\Gamma_i}|c_{i\mu}|=|O_i|+\sum_{\mu\notin\Gamma_i}|C_{i\mu}| \tag{5}$$

for minimization. This solved the delocalization problem for the virtual orbitals, but there appears to remain difficulty with generation of non-equivalent occupied orbitals which are in fact symmetry related (see Table 3).

It appears that there is no single, reliable, local total-population function on which to base localization and which is simultaneously applicable to the occupied and virtual mo's. Because of these problems, we abandoned the local total-population concept and studied the use of local overlap populations (two-center terms B_{Γ_i}, only) for the definition of an extremization function.

Table 1. P_i values, Eq. (1), for the virtual canonical molecular orbitals of ethylene

mo number[a]	1 CC_σ	2 CH	3 CH	4 CH	5 CH	6 $CC\pi$
P_i	2.41	1.46	0.0	3.22	0.0	0.0

[a] The mo's are listed in order of decreasing canonical eigenvalue

Table 2. Occupied and virtual localized molecular orbitals of ethylene calculated using the function P'_i (Eq (4))

		CC	CH	CH	CH	CH	CC_π
C_1	s	−0.38663	0.41961	0.02411	−0.10873	−0.24505	0.00000
C	p_z	0.00000	0.00000	0.00000	0.00000	0.00000	0.62725
C	p_x	0.17169	0.35453	0.03579	0.00626	0.33694	0.00000
C	p_y	−0.34593	−0.18251	0.12840	−0.01955	0.35504	0.00000
C_2	s	−0.37784	0.05845	−0.37861	0.28451	0.16595	0.00000
C	p_z	0.00000	0.00000	0.00000	0.00000	0.00000	0.62725
C	p_x	−0.01551	0.02708	0.29593	0.41737	−0.08523	0.00000
C	p_y	0.35019	0.00740	−0.32366	0.21021	−0.15576	0.00000
H_3	s	0.03995	−0.01887	−0.01243	0.54347	−0.10340	0.00000
H_4	s	0.09741	−0.00092	−0.51024	−0.16662	0.10276	0.00000
H_5	s	−0.08424	−0.09457	−0.02539	−0.01322	−0.53972	0.00000
H_6	s	0.21113	0.50467	−0.01792	−0.07802	0.04985	0.00000
P_i		2.00000	2.00000	2.00000	2.00000	1.94745	2.00000

		CC_π*	CH*	CH*	CH*	CH*	CC*
C_1	s	0.00000	0.46433	0.33388	−1.01902	0.37935	1.10673
C	p_z	−0.82807	0.00000	0.00000	0.00000	0.00000	0.00000
C	p_x	0.00000	−0.79124	0.95711	−0.22958	0.00000	−0.16816
C	p_y	0.00000	−0.34692	−0.05576	0.18106	−0.72819	1.06779
C_2	s	0.00000	−0.36737	−0.25422	0.77645	1.15147	−0.77838
C	p_z	0.82807	0.00000	0.00000	0.00000	0.00000	0.00000
C	p_x	0.00000	−0.40289	−1.03888	−0.61229	0.00000	−0.08132
C	p_y	0.00000	−0.28196	−0.00239	0.01855	0.29743	1.28777
H_3	s	0.00000	0.72245	0.77957	0.06718	−0.72093	−0.16671
H_4	s	0.00000	−0.05958	−0.51488	−0.88694	−0.72093	−0.32660
H_5	s	0.00000	−1.04075	0.37926	0.36786	−0.54273	−0.15761
H_6	s	0.00000	−0.21781	−0.77547	0.85237	−0.54273	0.10884
P_i		2.00000	2.00000	2.00000	2.00000	2.00000	2.13819

2.2. Localization functions derived from overlap populations

A localization function which ignores the >0 contribution of A_{Γ_i} (i.e., the one center contribution) to P_i and focuses on $B_{\Gamma_i} > 0$ for bonding orbitals and $B_{\Gamma_i} < 0$ for the virtual orbitals is

$$P_i = 2 \sum_{\mu\in\Gamma_i} \sum_{\substack{\nu\in\Gamma_i \\ \nu\neq\mu}} C_{i\mu} C_{i\nu} S_{\mu\nu} = B_{\Gamma_i}. \tag{6}$$

Again, maximizing P_i for bonding mo's and minimizing P_i for anti-bonding mo's works well; however, this definition of P_i is not applicable when the molecule contains occupied (lone pairs) or vacant one-center orbitals, because $S_{\mu\nu} = 0$ for all μ and ν confined to one center. To deal with this situation we tried a two step approach: in the first pass at localization, orbital rotations among all occupied mo's produce localized bond functions and among all empty mo's produce anti-bond functions; in the second pass, the resulting delocalized, non-equivalent

Table 3. Occupied and virtual localized molecular orbitals of ethylene calculated using the function P'' (Eq (5))

		CC	CH	CH	CH	CH	CC_π
C_1	s	−0.29374	0.00000	0.00168	0.47068	0.30031	0.00000
C	p_z	0.00000	0.00000	0.00000	0.00000	0.00000	0.62725
C	p_x	−0.01895	−0.00064	−0.01374	−0.28753	0.43219	0.00000
C	p_y	−0.46101	−0.02666	0.00000	−0.22962	−0.17301	0.00000
C_2	s	−0.29869	0.45497	0.31903	0.00342	0.00000	0.00000
C	p_z	0.00000	0.00000	0.00000	0.00000	0.00000	0.62725
C	p_x	0.01352	−0.29226	0.42937	0.00755	−0.00101	0.00000
C	p_y	0.45861	0.24726	0.15387	0.00306	0.02740	0.00000
H_3	s	0.14740	0.00000	0.53285	0.00000	−0.04999	0.00000
H_4	s	0.20094	0.49791	−0.20129	−0.06486	0.02988	0.00000
H_5	s	0.15448	−0.04770	0.02742	0.49133	−0.19965	0.00000
H_6	s	0.11975	0.00404	−0.07362	0.00000	0.53701	0.00000
P_i		2.09785	2.02633	2.00016	2.01201	2.00233	2.00000
		CC_π*	CH*	CH*	CH*	CH*	CC*
C_1	s	0.00000	0.00000	0.78986	−0.79826	0.00000	1.21373
C	p_z	−0.82807	0.00000	0.00000	0.00000	0.00000	0.00000
C	p_x	0.00000	0.14928	0.89214	0.88535	0.14518	0.00171
C	p_y	0.00000	−0.00280	−0.52367	0.52300	0.01812	1.13075
C_2	s	0.00000	0.79117	0.00000	0.00000	−0.79923	−1.21224
C	p_z	0.82807	0.00000	0.00000	0.00000	0.00000	0.00000
C	p_x	0.00000	−0.88841	−0.15181	−0.14019	−0.88944	0.00659
C	p_y	0.00000	0.52727	0.00552	−0.01182	−0.51735	1.13174
H_3	s	0.00000	−0.02722	0.01079	−0.00750	1.29567	0.04514
H_4	s	0.00000	−1.29539	0.00286	−0.00001	0.02499	0.05522
H_5	s	0.00000	0.01078	−0.02232	1.29563	−0.00116	−0.04936
H_6	s	0.00000	0.00566	−1.29535	0.02976	−0.00002	−0.05346
P_i		2.00000	2.04783	2.04907	2.04489	2.04648	2.00935

lone pair functions were rotated among themselves using the definition

$$P_i = \sum_{\mu \in \Gamma_i} C_{i\mu}^2 = A_{\Gamma_i}. \qquad (7)$$

The extremization of $P = \sum B_{\Gamma_i}$ in the first step produces results meeting the symmetry requirement in the absence of lone pairs (Table 4) but suggesting incomplete localization (i.e., the CC sigma bond orbital of ethylene). However, when lone pairs are present the unnatural, mutual exclusion of lone and bond orbital sets during localization leaves the lone pair orbitals to bear the majority of the intrinsic orbital delocalization for the complete set of occupied mo's. When the molecule possesses only twofold symmetry, the situation is further exacerbated by the σ/π separability within the bonding and lone pair sets. These findings are apparent from the study of HOCl (Table 5) and of p-difluorobenzene (not tabulated here). This severely limits the classes of molecules to which the Magnasco/Perico overlap population concept can be usefully applied. Furthermore, the need for dual localization functions is unsatisfying. All these difficulties have been circumvented by the technique of the next section.

Table 4. Occupied and virtual localized molecular orbitals of ethylene calculated with valence overlap population, P_i (Eq 6)

		CC	CH	CH	CH	CH	CC_π
C_1	s	−0.38273	0.35235	0.35235	−0.04016	−0.04016	0.00000
C	p_z	0.00000	0.00000	0.00000	0.00000	0.00000	0.62725
C	p_x	0.00000	0.36739	−0.36739	−0.00584	0.00584	0.00000
C	p_y	−0.39424	−0.25660	−0.25660	−0.06629	−0.06629	0.00000
C_2	s	−0.38273	−0.04016	−0.04016	0.35235	0.35235	0.00000
C	p_z	0.00000	0.00000	0.00000	0.00000	0.00000	0.62725
C	p_x	0.00000	−0.00584	0.00584	0.36739	−0.36739	0.00000
C	p_y	0.39424	0.06629	0.06629	0.25660	0.25660	0.00000
H_3	s	0.08646	−0.04709	0.02928	0.53825	−0.08885	0.00000
H_4	s	0.08646	0.02928	−0.04709	−0.08885	0.53825	0.00000
H_5	s	0.08646	−0.08885	0.53825	0.02928	−0.04709	0.00000
H_6	s	0.08646	0.53825	−0.08885	−0.04709	0.02928	0.00000
P_i		2.18900	2.08363	2.08363	2.08363	2.08363	2.00000
		CC_π*	CH*	CH*	CH*	CH*	CC*
C_1	s	0.00000	−0.81405	−0.01941	−0.81405	−0.01941	1.18664
C	p_z	−0.82807	0.00000	0.00000	0.00000	0.00000	0.00000
C	p_x	0.00000	0.88885	0.14662	−0.88885	−0.14662	0.00000
C	p_y	0.00000	0.50440	−0.02799	0.50440	−0.02799	1.14729
C_2	s	0.00000	0.01941	0.81405	0.01941	0.81405	−1.18664
C	p_z	0.82807	0.00000	0.00000	0.00000	0.00000	0.00000
C	p_x	0.00000	−0.14662	−0.88885	0.14662	0.88885	0.00000
C	p_y	0.00000	−0.02799	0.50440	−0.02799	0.50440	1.14729
H_3	s	0.00000	−0.00494	−0.02671	−0.00605	−1.29618	0.02922
H_4	s	0.00000	−0.00605	−1.29618	−0.00494	−0.02671	0.02922
H_5	s	0.00000	1.29618	0.00605	0.02671	0.00494	−0.02922
H_6	s	0.00000	0.02671	0.00494	1.29618	0.00605	−0.02922
P_i		2.00000	2.04617	2.04617	2.04617	2.04617	2.00316

2.3. A localization function derived from remote populations

Returning to the total mo density expression (3), we have seen that minimization of the magnitudes of the terms remote to the bonding, anti-bonding or lone pair region in question (i.e., $|O_i| = |2 - P_i|$), does not necessarily lead to localization of orbitals at the bond or atomic region (Table 2). This is because O_i may approach zero with large, but nearly cancelling, values for its local/remote ($O_i^{l,r}$) and purely remote (O_i^r) components. Thus, we were led to investigate the possibility of minimizing the purely remote terms only. Accordingly, we define the purely remote density function

$$P_i^r = 2 \sum_{\mu\notin\Gamma_i} \sum_{\nu\notin\Gamma_i} C_{i\mu} C_{i\nu} S_{\mu\nu} = O_i^r \tag{8}$$

and minimize the function P defined as

$$P = \sum_{i=1}^{N} P_i^r. \tag{9}$$

Table 5. Occupied and virtual molecular orbitals of HOCl calculated with overlap population, P_i (Eqs. (6, 7))

		OCl	OH	O(lp_1)	O(lp_2)	Cl(lp_1)	Cl(lp_2)
C_1	s	−0.36366	0.08958	0.29715	0.0	0.77849	−0.32649
	p_z	−0.00000	0.00000	−0.00000	0.68991	−0.00000	−0.00000
	p_x	−0.43884	0.12248	−0.25497	0.0	−0.35465	0.05168
	p_y	0.00449	0.05521	0.50570	0.0	0.48667	0.71300
O	s	−0.38520	−0.49195	0.38982	0.0	−0.15483	0.41584
	p_z	0.00000	−0.00000	0.00000	0.65179	0.00000	0.00000
	p_x	0.47960	−0.23304	0.36274	0.0	0.10776	0.20173
	p_y	0.03674	−0.50836	−0.50056	0.0	0.05158	−0.49913
H	s	0.08248	−0.36835	−0.01441	0.0	0.03522	−0.02954
P_i		2.03856	2.02314	1.06822	0.84266	1.93734	1.23527

		Cl(lp_3)	OCl*	OH*
C_1	s	−0.00000	0.40727	−0.03481
	p_z	−0.72559	0.0	0.0
	p_x	0.00000	0.86420	0.02560
	p_y	−0.00000	−0.00114	−0.02663
O	s	−0.00000	−0.53171	−0.69449
	p_z	0.71160	0.0	0.0
	p_x	−0.00000	0.84364	−0.21561
	p_y	0.00000	0.06790	−0.68236
H	s	0.00000	−0.03124	1.18761
P_i		1.05296	2.18319	1.99508

Tables 6–9 present the results for the occupied and virtual orbitals of ethylene, butadiene, benzene and HOCl. The results are extremely satisfying in that the orbitals are well localized, and chemical equivalence, when required by symmetry, is obtained. Furthermore, rotations among all occupied orbitals (bond and lone) are treated, avoiding the difficulties with the two-pass technique and the separate function definitions.

It should be noted that in Tables 6–9 the symmetry separation of σ and π mo sets under 2-fold symmetry was preserved by forbidding interset rotations. Note further that for butadiene and benzene the final localized π mo's are of the two center type, corresponding in the case of benzene to a Kekulé resonance structure. The alternate Kekulé structure is equally valid, of course, and is obtained by specification of alternate carbon atom pairs as the bonded pairs [21].

The use of definition (8) for P_i^r involves more terms than does the Magnasco/Perico definition and so requires more computation time than the latter. Of course, this is acceptable to avoid arbitrarily localized orbitals.

2.4. Global extrema

Finally, we would like to point out that the orbital pair rotation technique is not guaranteed to produce globally localized orbitals. To test for global extremization

Table 6. Occupied and virtual localized molecular orbitals of ethylene calculated with remote overlap population, P_i (Eq (8))

		CC	CH	CH	CH	CH	CC_π
C_1	s	0.38810	0.04351	0.34900	0.34900	−0.04351	0.00000
C	p_z	0.00000	0.00000	0.00000	0.00000	0.00000	0.62725
C	p_x	0.00000	−0.00584	0.36739	−0.36739	−0.00584	0.00000
C	p_y	0.38856	0.06969	−0.26000	−0.26000	−0.06969	0.00000
C_2	s	0.38810	−0.34900	−0.04351	−0.04351	0.34900	0.00000
C	p_z	0.00000	0.00000	0.00000	0.00000	0.00000	0.62725
C	p_x	0.00000	0.36739	−0.00584	0.00584	0.36739	0.00000
C	p_y	−0.38856	−0.26000	0.06969	0.06969	0.26000	0.00000
H_3	s	−0.08269	0.08811	−0.04636	0.03001	0.53898	0.00000
H_4	s	−0.08269	−0.53898	0.03001	−0.04636	−0.08811	0.00000
H_5	s	−0.08269	0.04636	0.08811	0.53898	0.03001	0.00000
H_6	s	−0.08269	−0.03001	0.53898	0.08811	−0.04636	0.00000
P_i		2.18898	2.08342	2.08342	2.08342	2.08342	2.00000
		CC_π*	CH*	CH*	CH*	CH*	CC*
C_1	s	0.00000	−0.01453	−0.80917	0.80917	0.01453	1.19344
C	p_z	−0.82807	0.00000	0.00000	0.00000	0.00000	0.00000
C	p_x	0.00000	0.14662	0.88885	0.88885	0.14662	0.00000
C	p_y	0.00000	−0.02330	0.50910	−0.50910	0.02330	1.14334
C_2	s	0.00000	0.80917	0.01453	−0.01453	−0.80917	−1.19344
C	p_z	0.82807	0.00000	0.00000	0.00000	0.00000	0.00000
C	p_x	0.00000	−0.88885	−0.14662	−0.14662	−0.88885	0.00000
C	p_y	0.00000	0.50910	−0.02330	0.02330	−0.50910	1.14334
H_3	s	0.00000	−0.02658	−0.00481	0.00592	1.29605	0.03469
H_4	s	0.00000	−1.29605	−0.00592	0.00481	0.02658	0.03469
H_5	s	0.00000	0.00592	1.29605	−0.02658	−0.00481	−0.03469
H_6	s	0.00000	0.00481	0.02658	−1.29605	−0.00592	−0.03469
P_i		2.00000	2.04622	2.04622	2.04622	2.04622	2.00441

it is necessary to subsequently construct, and diagonalize, a second-order matrix (at least) of the localization function [22]. From the eigenvalues one can tell if a true extremum has been found. As described in the appendix, the numeric search procedure we have used incorporates a second-order search for the global extremum and does not require an analytical localization function for differentiation to second order.

Table 7. Localized π molecular orbitals of butadiene calculated with remote overlap population, P_i^r (Eq (8))

	Occupied		Virtual	
C_1	−0.145	0.632	−0.113	−0.805
C_2	0.085	0.596	−0.198	0.853
C_3	0.596	0.085	0.853	−0.198
C_4	0.632	−0.145	−0.805	−0.113

Table 8. Localized π molecular orbitals of benzene calculated with remote overlap population, P_i^r (Eq (8))

	Occupied			Virtual		
C_1	0.561	−0.185	0.187	0.162	0.316	−0.792
C_2	0.561	0.187	−0.185	0.316	0.162	0.792
C_3	0.187	0.561	−0.185	−0.792	−0.162	−0.315
C_4	−0.185	0.561	0.187	0.792	−0.316	−0.162
C_5	−0.185	0.187	0.561	−0.316	0.792	0.162
C_6	0.187	−0.185	0.561	−0.162	−0.792	0.315

2.5. Localized hybrid orbital characteristics

Having obtained a satisfactory procedure for generating the localized orbitals, we examined the results for several of the molecules for their chemical interpretability, choosing fractional *s* character of the atomic hybrids for this purpose. The results are displayed in Table 10.

The first column of figures gives the fractional *s* character for the designated atom (O and Cl) in its hybrid contribution to the localized bond or lone pair orbital (for bond orbitals, the molecular orbital is a linear combination of normalized atomic hybrid orbitals on each of the two atoms forming the bond).

Table 9. Occupied and virtual molecular orbitals of HOCl calculated with remote overlap population, P_i^r (Eq (8))

		OCl	OH	O(lp_1)	O(lp_2)	Cl(lp_1)	Cl(lp_2)
Cl	s	0.05630	0.04744	0.02524	−0.00000	0.68379	0.68379
	p_z	0.00000	−0.00000	0.00000	−0.07754	−0.00000	−0.00000
	p_x	0.58319	0.05013	0.00794	−0.00000	−0.16873	−0.17286
	p_y	0.00585	0.03662	−0.02251	−0.00000	−0.70721	0.70852
O	s	0.01858	−0.41787	0.72789	0.00000	−0.00969	−0.01475
	p_z	−0.00000	−0.00000	−0.00000	0.96451	0.00000	−0.00000
	p_x	−0.61263	−0.10817	−0.28397	0.00000	−0.01588	−0.01730
	p_y	0.05346	−0.61359	0.57793	0.00000	0.01791	−0.03161
H	s	−0.12895	−0.35795	−0.00048	0.00000	−0.00189	−0.00682
P_i		2.01539	2.00016	1.88892	1.86056	1.99238	1.99889

		Cl(lp_3)	OCl*	OH*
C_1	s	−0.00000	0.33800	−0.04539
	p_z	−1.00052	0.0	0.0
	p_x	−0.00000	0.89440	0.02062
	p_y	−0.00000	0.00201	−0.02735
O	s	0.00000	0.48364	−0.67310
	p_z	0.03042	0.0	0.0
	p_x	0.00000	0.82379	−0.23190
	p_y	−0.00000	0.03568	−0.68415
H	s	−0.00000	−0.00859	1.19243
P_i		2.00209	2.13585	2.00415

Table 10. Hybrid *s* orbital percentages and polarization

		Remote		Overlap	
		% *s*	$\lvert sp \rvert$	% *s*	$\lvert sp \rvert$
$HONH_2$	ON	32.1	0.304	58.3	0.474
	OH	59.9	0.480	67.7	0.498
	O(*lp*)	50.2	0.448	41.4	0.374
HOCl	OCl	29.9	0.286	63.4	0.459
	OH	54.3	0.456	65.7	0.496
	O(*lp*)	53.7	0.453	42.5	0.385
HOOH	OO	22.0	0.214	58.4	0.474
	OH	59.8	0.479	69.4	0.500
	O(*lp*)	52.9	0.449	43.5	0.392
HOF	OF	10.3	0.103	78.9	0.485
	OH	56.5	0.466	69.2	0.500
	O(*lp*)	55.0	0.460	44.3	0.397
$ClCH_3$	ClC	38.5	0.355	60.8	0.483
	Cl(*lp*)	53.3	0.440	47.2	0.416
$ClNH_2$	ClN	20.2	0.198	58.4	0.474
	Cl(*lp*)	54.6	0.457	50.4	0.436
ClOH	ClO	9.2	0.092	64.0	0.492
	Cl(*lp*)	55.5	0.462	51.8	0.443

The results are chemically useful in that a comparison of hybrids directed to two different substituents at an atom reveals a higher concentration of *s* character in that atom's hybrid directed toward its less electronegative substituent. This is expected from Bent's arguments [23]. Also in accordance with Bent's principle, variation of one substituent at an atom (e.g., the series HOX, $X = NH_2$, Cl, OH, F) causes the fractional *s* character in the hybrid directed to the varying substituent to decrease as the substituent electronegativity increases. A specially significant finding is that the loss of *s* character in a bond hybrid appears to cause an increase in the fractional *s* character of the lone pair hybrids in preference to another bond hybrid. This is a pleasing result because there are sound energetic reasons [24] for maximization of *s* electron density in molecules, and lone pair orbitals inherently satisfy this physical principle (contrast bond orbitals, for which there must be antibonding counterparts containing *s* character). Clearly, however, there is no quantitative adherence of the results to the "isovalency" presumption of Bent's full arguments, nor should there be because that presumption ignores the fact that *s* character is necessarily directed into virtual orbitals and not confined to the occupied set.

The third column of figures gives the results for hybrid *s* character when localization is derived from the overlap population criterion. These results are characterized by high *s* characters for the bond hybrids, they exhibit low *s* character for lone pair electrons and they do not follow the sound energetic basis for Bent's principle of *s* distribution among bond hybrids. The reason for the high concentration of *s* character in the bond hybrids derives from the localization criterion itself, that is, the localized bond orbitals are dictated to have maximum interference of bonding atomic hybrids and this is achieved by s characters in the vicinity of 50%. To emphasize this point the fourth column of figures gives the magnitude of the product of s and p coefficients for the atomic hybrids; this is a quantitative measure of the extent to which a hybrid orbital is polarized [25]. The greater a hybrid's polarization, the greater the overlap population. Comparison of column four values with those in column two for the remote localization criterion are striking.

As a consequence of these observations, we conclude that the remote localization criterion is of greater value than the overlap population criterion with regard to the interpretation of chemical effects. Combined with its advantages in consistency and ease of application, its rigor to second order, and its nonprejudicial foundation, we feel the case is exceptionally strong for using the remote criterion for localization studies.

3. Conclusions

The localization concept of Magnasco and Perico has been generalized and extended to provide easy access to lone and virtual orbital localization. This is done by defining the localization function in terms of overlap populations completely remote to the two- and one-center molecular regions defining each localized mo. Other definitions were considered, but each led to problems with consistency and/or arbitrariness of application and with chemically useful results. Furthermore, being forced to adopt a numerical search procedure has made possible automatic searching to second-order for optimal extremization of the localization function.

That the numerical minimization of purely remote density resolves the problem of the consistent definition of a localization function for LCAO mo's will be of great value for configuration interaction studies, because of the considerable reduction in size of the CI matrix needed and the attendant reduction in cost of such calculations to yield useful chemical results for large molecules.

Finally, and with respect to a reviewer comment concerning the limits of applicability of our technique to expanded atomic orbital bases, the remote population extremization procedure can be applied only to those expansion schemes which preserve the atomic centering of the expansion functions. Other schemes, such as the one-center expansion method, cannot be treated as we have in this work.

Acknowledgement. We wish to thank the Compagnie Internationale de Service en Informatique, and in particular the personnel of the Marseille office, for their assistance in making these calculations. We also wish to thank Professor L. Pujol for his constructive criticism of the manuscript.

Appendix A: Numeric extremization of the localization functions on the interval $(0, \pi)$

The standard localization procedure entails successive two-dimensional rotations in orbital space with the object of extremizing an analytical localization function for a pair of orbitals; the rotation angle θ spans the interval 0, π. Such localization procedures utilize a function which is differentiable over the interval 0, π and which leads to an analytical solution for the extremizing value of θ. Some of the functions with which we have had to deal in this work do not readily lend themselves to this analytical simplification. In such cases one must resort to a numeric solution.

We have chosen for our application an iterative method based on quadratic (i.e., second-order) interpolation of three function values to search for the function minimum. Given a starting value of the function, the direction of the minimum from that point is determined by calculation of the function value at two neighboring points on the extremization curve. Moving to the point of lower function value, the process is repeated until either the minimum is located or an interval boundary is encountered. Once the region of the minimum is enclosed, the quadratic search for the minimum begins, using the enclosing values of θ and an intermediate point; the location of the minimum is estimated and two new enclosing and an intermediate points are taken to repeat the quadratic search for the function minimum. The search is completed when the convergence criteria are satisfied.

Appendix B: Convergence criteria

Three parameters are controlled by the user to define the quality of the convergence desired. The first parameter, ε, defines the tolerance of the value of θ at the minimum, that is, when two successive iterations give values of θ within ε. The second control of the iterative procedure is NFMAX, which limits the number of search attempts. Because the localization functions with which we have had to work can present local minima, it is necessary to divide the 0, π interval for θ into NINTER subintervals of equal range in θ, for each of which a local minimum is searched. The global minimum is then simply given by the lowest of these local minima. In those special cases where the function is independent of θ, within the computer arithmetic precision, the initial value of θ is taken.

Appendix C: Accuracy of the results

The arithmetic accuracy of the results is determined by the use of single- or double-precision arithmetic. The results presented in this article were all obtained with double-precision arithmetic (64 bit floating point words), although single-precision arithmetic would have sufficed. For values of NINTER between 10 and 100 there is no effect on the specification of the global minimum for any of the molecules considered in this work. When NFMAX is large enough, selection of ε over the range 10^{-6} to 10^{-15} produces results essentially indistinguishable from each other. For $\varepsilon = 10^{-8}$, twenty iterations suffice, in general, to locate the global minimum. Regarding computation times on the KSU NAS 6630 system, one to two seconds are required for $\varepsilon = 10^{-8}$ and NFMAX = 100, to be compared with ten to twenty seconds with $\varepsilon = 10^{-15}$ and NFMAX = 1000. For most of the results presented here, $\varepsilon = 10^{-8}$, NFMAX = 100 and NINTER = 10.

References and notes

1. Millie P, Levy B, Berthier G (1975) Localisation and relocalisation in orbital theories. In: Chalvet O, Daudel R, Diner S, Malrieu JP (eds) Localization and delocalization in quantum chemistry. Reidel, Dorcrecht Boston, p 59
2. Anh NT (1970) Les regles de Woodward Hoffman. Edisciences, McGraw-Hill, Paris
3. Dixon DA, Kleier DA, Lipscomb WN (1978) J Am Chem Soc 100: 5681
4. Gimarc BM (1974) Acc. Chem Res 7:384
5. Trindle C, Sinanoglu O (1969) J Am Chem Soc 91:4054
6. Sunberg KR, Bicerano J, Lipscomb WN (1979) J Chem Phys 71:1515
7. Kleier DA, Halgren TA, Hall, Jr. JH, Lipscomb WN (1974) J Chem Phys 61:3305
8. Purcell KF, Dolf TGM (1972) J Am Chem Soc 94:2693
9. Marynick DS, Kirkpatrick CM (1983) J Phys Chem 87:3273
10. Daudey JP, Malrieu JP, Rojas O (1975) Empirical and theoretical partitions of the molecular energy into local contributions: additive systematic and conformation problems. In: Chalvet O,

Daudel R, Diner S, Malrieu JP (eds) Localisation and delocalisation in quantum chemistry. Reidel, Dordrecht Boston, p. 155
11. Caldwell D, Eyring H (1978) Localized orbitals in spectroscopy. In: Lowdin PO (ed) Advances in quantum chemistry, vol. 11. Academic Press, New York San Francisco London, p 93
12. Cizeck J, Forner W, Ladik J (1983) Theor Chim Acta 64:107
13. Edmiston C, Rudenberg K (1983) Rev Mod Phys 35:457
14. Boys SF (1960) Rev Mod Phys 32:296; Foster JM, Boys SF (1960) Rev Mod Phys 32:300
15. Niessen WV (1972) J Chem Phys 56:4290
16. Magnasco V, Perico A (1967) J Chem Phys 47:971
17. Bohm MC (1981) Theor Chim Acta 59:609. Note: it is not clear that a correct localization function for virtual mo's under a non-zero-differential-overlap Fock operator was defined in this work
18. Offenhay, PO'D (1970) Atomic and molecular orbital theory. McGraw-Hill, New York, Appendix 2
19. Hoffman R (1963) J Chem Phys 39:1397
20. Vignes J (1980) Algorithmes numeriques, analyse et mise en oeuvre. Editions Technip, Publications de l'Institut Francais du Petrole
21. Newton MD, Switkes E (1971) J Chem Phys 54:3179
22. Purcell KF, Dolph TGM, Shultz MJ program 198.1. QCPE, Indiana University, Bloomington, Ind 47401; Switkes E, Lipscomb WN, Newton MD (1970) J Am Chem Soc 92:3847
23. Bent HA (1961) Chem Rev 61:275
24. Hall MB (1978) J Am Chem Soc 100:6333
25. Purcell KF, Kotz JC (1977) Inorganic chemistry. Saunders, Philadelphia Exercise 3, p 107