Construction of Optimal Basis Sets in Approximate Theories of the Electronic Structure of Molecules

Weighted Löwdin Orthogonalization

Z. B. Maksić *
Organisch-chemisches Institut der Universität Heidelberg

Z. Naturforsch. **36a**, 373–377 (1981); received February 3, 1981

Properties of the weighted Löwdin orthogonalization procedure are thoroughly discussed. It is shown that its flexibility might be very useful in the construction of optimal truncated basis sets in approximate theories of the electronic structure of molecules and solids. The relevance of pairwise nonorthogonal (hybrid) basis functions is pointed out.

Introduction

It is well established by now that the proper choice of the basis set functions in calculations of the electronic structure of molecules and solids is of utmost importance. Computations employing complete basis sets are obviously not feasible. Therefore, the selection of optimal truncated sets has far reaching consequences. Intuition and experience play a great role in this respect. Applications of basis sets designed by using simple but adequate physical concepts and ideas can save a lot of computational efforts extending in the same time the range of practicability of the quantum chemical methods. Excellent illustrative examples are provided by gauge-invariant atomic orbitals (GIAO) [1–3] and electric field variant atomic orbitals (EFV) [4, 5], which proved very useful in studying molecular systems perturbed by the external magnetic and electric fields, respectively. The use of atomic orbitals well adapted to local molecular fields is even more important in approximate theories** of chemical bonding. We have shown in a series of papers that hybrid orbitals exhibiting maximum overlapping provide a fair description of a large number of molecular properties [6–10]. This is not surprising because the local hybrid orbitals intrinsically reflect two most important facets of covalent bonding: (a) pile-up of charge in the bonding regions and (b) spatial arrangement of bonds. Since the hybrids conform the symmetry of local fields they provide a natural starting point for more sophisticated and ambitious calculations being especially suitable for approximate methods. Indeed careful analyses of molecular integrals have shown that various NDDO (neglect of differential orbital orbital) existing schemes are justified to some extent if and only if the basis set of symmetrically orthogonalized hybrid orbitals is used [11–14]. This type of basis set has also a good performance within the framework of separated electron pairs formalism [14–17] offering a description of closed shell molecules which is closest to chemical intuition. It is worth to mention that in all these studies the hybrid orbitals are orthogonalized to inner shells utilizing Schmidt procedure. The inner core orbitals were left intact during the orthogonalization procedure in order to avoid the energy contamination and reflecting chemical inertness of these orbitals. The hybrid orbitals were subsequently symmetrically orthogonalized by using Löwdin method [18] which in turn has a number of remarkable properties. The salient feature of Löwdin orbitals is their localisability and the maximum similarity with initial nonorthogonal (hybrid) orbitals as measured by the least squares' deviations [19]. Extremal likelihood with starting functions makes Löwdin orbitals very convenient for a construction of basis sets in large systems by using localised functions of their fragments. The optimal

---

* On leave of absence from the Theoretical Chemistry Group, "Rudjer Boskovic" Institute, 41001 Zagreb, Yugoslavia and The Faculty of Natural Sciences and Mathematics, University of Zagreb, Marulicev trg 19, 41000 Zagreb, Yugoslavia.

** Semiempirical theories are certainly approximate but it should be pointed out here that there are approximate schemes which do not heavily lean on the empirical data (see Ref. [14]).

Reprint requests to Prof. Dr. Z. B. Maksić, Organisch-chemisches Institut der Universität Heidelberg, Im Neuenheimer Feld 270, D-6900 Heidelberg 1.

0340-4811 / 81 / 0400-0373 $ 01.00/0. — Please order a reprint rather than making your own copy.

localised functions of small molecules, which are building blocks of larger compounds, are transferred and mutually symmetrically orthogonalized in order to ensure minimum of their perturbation at least in the Hilbert space. However, we have to keep in mind two important points. Firstly, Löwdin orbitals are minimaly disturbed on the average, i.e. all participating orbitals are treated in the same time on the same footing. That's how symmetrical orthogonalization earned its name. Secondly, atomic orbitals are not used in molecular calculations only because they are computationally suitable. More importantly, they do have some physical meaning describing an electron in the vicinity of a particular nucleus. In fact, this is the reason behind the success of the relatively simple MO methods based on AOs. We have to racall also that two AOs of different energies form a bonding MO in a way that the latter is composed predominately of the AO possessing lower orbital energy. It follows immediately that symmetrical Löwdin orthogonalization is convenient if AOs with similar atomic energies are involved. However if atoms with, loosely speaking, widely different electronegativities are considered, some sort of discrimination between AOs is desirable. A sort of asymmetrical orthogonalization which minimizes deviations from energetically low lying AOs at the expense of higher AOs is clearly to be preferred. Only in this case we can expect the initial AOs or HAOs are minimally perturbed in energetic sense. This goal is achieved by the weighted Löwdin procedure as shown in the next section.

Weighted Löwdin Orthogonalization

The possibility of weighted least-square orthogonalization was mentioned in the early work of Carlson and Keller [19] but it was, to the best of our knowledge, never discussed in some detail. Hence we give here a derivation. Let's denote the initial set of linearly independent normalized functions by $\varphi_i (i = 1, \ldots, n)$. Withouth much loss of generality we shall suppose that φ_i are real functions and that the set $\{\varphi_i\}$ is finite. The formalism can be easily extended to encompass complex functions and infinite sets. The starting set is nonorthogonal, i.e. $\langle \varphi_i | \varphi_j \rangle = S_{ij} (i \neq j)$ where at least some of S_{ij} are different from zero. Our aim is a construction of orthonormal orbitals $\psi_i (i = 1, \ldots, n)$ which span the same subspace $\{\varphi_i\}$ of the Hilbert space and satisfy the requirement:

$$\sum_{i,j} W_{ji} \langle \psi_i - \varphi_i | \psi_j - \varphi_j \rangle = \min . \qquad (1)$$

where W_{ji} are elements of the symmetric matrix $\tilde{W}$ formed by the predetermined weighting factors. It is also tacitly assumed that $\tilde{W}$ is non-singular but not necessarily positive definite. Each initial function φ_i may be expressed as a linear combination of ψ_i:

$$\varphi_i = \sum_{k=1}^{n} \psi_k C_{ki} .$$

One can easily prove that the following matrix equation holds

$$\tilde{T} = \tilde{C}^t \tilde{C} \qquad (2)$$

where $\tilde{T} = \tilde{I} + \tilde{S}$, $\tilde{I}$ is identity matrix ($\tilde{I}_{ij} = \delta_{ij}$) and t denotes transposed matrix. Substituting

$$\psi_i - \varphi_i = \sum_{k=1}^{n} \psi_k (\tilde{I}_{ki} - \tilde{C}_{ki}) \quad \text{and}$$

$$\psi_j - \varphi_j = \sum_{l=1}^{n} \psi_l (\tilde{I}_{lj} - \tilde{C}_{lj}) \qquad (3)$$

into condition (1) one obtains

$$Tr[\tilde{W}(\tilde{I} + \tilde{T}) - \tilde{W}(\tilde{C} + \tilde{C}^t)] = \min . \qquad (4)$$

The first term in the matrix expression (4) is obviously constant. Consequently

$$Tr[\tilde{W}(\tilde{C} + \tilde{C}^t)] = \max \qquad (5)$$

is equivalent to the previous requirement (4). The coefficients $\tilde{C}$ possessing extremal property (5) are obtained by using an auxiliary matrix $\tilde{A}$

$$\tilde{A} = \tilde{C}\tilde{W}(\tilde{W}\tilde{T}\tilde{W})^{-1/4} - (\tilde{W}\tilde{T}\tilde{W})^{1/4} . \qquad (6)$$

It should be mentioned that $\tilde{W}\tilde{T}\tilde{W}$ is a symmetric matrix because $(\tilde{W}\tilde{T}\tilde{W})^t = (\tilde{W}^t\tilde{T}^t\tilde{W}^t) = \tilde{W}\tilde{T}\tilde{W}$. Since the matrix $\tilde{W}\tilde{T}\tilde{W}$ is also positive definite, the fourth-root is in principle easily found by the diagonalization technique. One can prove by excercising some algebra that

$$\begin{aligned} Tr[\tilde{A}^t\tilde{A}] = Tr[2(\tilde{W}\tilde{T}\tilde{W})^{1/2} \\ - \tilde{W}(\tilde{C} + \tilde{C}^t)] . \end{aligned} \qquad (7)$$

It appears that the requirement (6)

$$\begin{aligned} &Tr[\tilde{W}(\tilde{C} + \tilde{C}^t)] \\ &= Tr[2(\tilde{W}\tilde{T}\tilde{W})^{1/2} - \tilde{A}^t\tilde{A}] = \max \end{aligned}$$

is satisfied if $\tilde{A}$ is a null matrix ($\tilde{A} = \tilde{O}$). Thus

$$\tilde{C}\tilde{W}(\tilde{W}\tilde{T}\tilde{W})^{-1/4} = (\tilde{W}\tilde{T}\tilde{W})^{1/4} \tag{8}$$

implying

$$\tilde{C} = (\tilde{W}\tilde{T}\tilde{W})^{1/2}\tilde{W}^{-1} \quad \text{and}$$
$$\tilde{C}^{-1} = \tilde{W}(\tilde{W}\tilde{T}\tilde{W})^{-1/2}. \tag{9}$$

The general form of the optimal set of orthogonal functions is

$$\psi_i = \sum_{k=1}^{n} \varphi_k [\tilde{W}(\tilde{W}\tilde{T}\tilde{W})^{-1/2}]_{ki}. \tag{10}$$

If $\tilde{W}$ assumes a special case $\tilde{W} = \tilde{I}$ then functions (10) reduce to well known Löwdin symmetrical orbitals. Particularly interesting case arises when $\tilde{W}$ is a diagonal matrix $W_{ij} = w_{ii}\delta_{ij}$. Then the stationary requirement (1) takes a simple form

$$\sum_{ij} w_{ii}\delta_{ij}\langle\psi_i - \varphi_i|\psi_j - \varphi_j\rangle$$
$$= \sum_i w_{ii}\langle\psi_i - \varphi_i|\psi_i - \varphi_i\rangle = \min. \tag{11}$$

In other words, the squares of the deviations from the initial functions are weighted relative to each other yielding additional flexibility compared to symmetrical orthogonalization ($w_{ii} = 1$ for all i). In particular, if we want to simulate Schmidt procedure by excluding a particular function from the orthogonalization process it is attained simply by putting e.g. $w_{kk} = 0$. Orthogonal functions (10) can be now written in a more specific form. Let $\tilde{D}$ be a diagonal matrix $D_{ij} = d_{ii}\delta_{ij}$, where d_{ii} are eigenvalues of the matrix $(\tilde{W}\tilde{T}\tilde{W})_{ij} = w_{ii}w_{jj}T_{ij}$

$$\tilde{O}^t(\tilde{W}\tilde{T}\tilde{W})\tilde{O} = \tilde{D}. \tag{12}$$

The columns of the orthogonal matrix $\tilde{D}$ are the eigenvectors of the matrix $\tilde{W}\tilde{T}\tilde{W}$. Its square root $(\tilde{W}\tilde{T}\tilde{W})^{-1/2}$ is given by

$$(\tilde{W}\tilde{T}\tilde{W})^{-1/2} = \tilde{O}\tilde{D}^{-1/2}\tilde{O}^t, \tag{13}$$

where $(\tilde{D}^{-1/2})_{ij} = d_{ii}^{-1/2}\delta_{ij}$ and the square roots are taken with positive signs. Orthogonal orbitals are then given by

$$\psi_i = \sum_{k}^{n} \varphi_k \left[w_{kk} \sum_{j=1}^{n} O_{kj} O_{ij} d_{jj}^{-1/2}\right]. \tag{14}$$

The most complicated step in the formation of orthogonal orbitals (10, 14) is diagonalization of the matrix $\tilde{W}\tilde{T}\tilde{W}$ but this can be performed by standard computer subroutines. The most striking feature of the functions defined by Eqs. (10) and (14) is that they are asymmetrically orhogonalized but in order to distinguish them from e.g. Schmidt orthogonal functions we shall cell them *weighted Löwdin orbitals*. This name stresses also the important role of weighting factors.

Discussion and Conclusion

The weighted Löwdin orbitals have considerable intrinsic flexibility contained in the free choice of the weighting factors. The diagonal elements w_{ii} are related to orbital energies E_i. The first guess would be $w_{ii} \sim 1/E_i$ in order to ensure minimum disturbance of the low energy orbitals φ_i. It is intuitively clear that the smallest possible perturbation of the most stable AOs must be advantageous. The off-diagonal adjustable weighting factors w_{ij} may be employed to minimize errors occurring in neglect of molecular integrals involving mixed densities $\psi_i\psi_j$. In other words we may try to find out the basis set $\{\psi_i\}$ in which the NDDO approximation holds most accurately. This is of utmost importance because it seems that NDDO approximation is the only way to circumvent the n^4 bottleneck. The use of NDDO scheme in nonorthogonal basis sets is unjustified [11—14]. It should be stressed that the use of minimum perturbation orthogonal orbitals might be favourable in separated electron pairs schemes [15, 16]. Furthermore, the weighting factors procedure is so general that it involves Schmidt-like orthogonalization as a special case. This feature may be useful in theoretical description of hole-states in atoms [20] and molecules [21]. The latter are conveniently studied by the X-ray photoelectron spectroscopy. So far we discussed orthogonalized basis sets but it is likely that pairwise-nonorthogonality might be even more useful [22]. The best argument which supports this statement is the success of the maximum overlap method based on the variable hybridization model [6—10]. Since the overlap integrals represent a good measure of the increased electron density and bond strength we conclude that a pair of overlapping hybris provides the most natural and intuitively appealing description of covalent bonds. The easiest way to reestablish overlapping bond hybrid orbitals is probably achieved by a transformation

$$\chi_i = (\psi_i + a\psi_j)/(1 + a^2)^{1/2} \quad \text{and}$$
$$\chi_j = (\psi_j + a\psi_i)(1 + a^2)^{1/2} \tag{15}$$

the extent of overlapping being $\langle \chi_i | \chi_j \rangle = 2a/(1 + a^2)$. The mixing coefficient a is another disposable parameter which can be obtained by fitting a particular approximate scheme for the evaluation of molecular integrals. The overlapping pair of hybrids (χ_i, χ_j) is apparently convenient for the calculation of very important one-electron integrals β_{ij} (or $h_{ij} = \langle \chi_i | \hat{h} | \chi_j \rangle$) which are traditionally assumed to be proportional to overlap integrals possessing a form $\beta_{ij} = f(h_{ii}, h_{jj}) S_{ij}$. It is well known that β-integrals have dramatic influence on the molecular stabilization [23]. Pairwise-overlapping bonding orbitals should be advantageous for the calculation of the one-center Coulomb repulsion integrals of the type

$$\langle \chi_i \chi_i | \chi_i \chi_i \rangle, \ \langle \chi_i \chi_i | \chi_j \chi_j \rangle,$$

etc., because χ_i orbitals are closer to the initial φ_i orbitals than the ψ_i-ones. This is obvious because the largest perturbation in the orthogonalization process is in fact produced by the nearest-neighbour bonding partner. Consider for example $(\tilde{I} + \tilde{S})^{-1/2}$ expansion of the symmetrically orthogonal orbitals ψ_i

$$\psi_i = \varphi_i - \tfrac{1}{2} \sum_{k}^{(k \neq i)}{}' \varphi_k S_{ki} + \ldots . \tag{16}$$

The series (16) may converge slowly but here we use it only for illustrative purposes. If the series (16) is plug in the expression (15) for χ_i and the normalization factor is neglected one obtains

$$\chi_i \simeq (1 - \tfrac{1}{2} a S_{ij}) \varphi_i + (a - \tfrac{1}{2} S_{ij}) \varphi_j - \tfrac{1}{2} \sum_{k}^{(k \neq i,j)}{}'' \varphi_k (S_{ki} + a S_{kj}) . \tag{17}$$

Taking $a \simeq \frac{1}{2} S_{ij}$ the contribution of the bonding partner φ_j diminish practically to zero, while the coefficient of the initial (hybrid) orbital is smaller roughly only by $\sim (1/4) S_{ij}^2$. Since S_{ij} integral is the largest one by the very definition of the bonding partner and high polarization of hybrid orbital in the particular direction in space, we have removed the largest term in the series (17) leading to appreciable delocalization of the function χ_i. It is noteworthy that the orthogonality properties of the χ_i function is not seriously destroyed because it remains to be orthogonal to the rest of $(n - 2)$ functions. The pairwise-overlapping hybrid orbitals are favourable in constructing approximate integral schemes (vide supra) but they have conceptual advantage since they describe the essential physics of the problem.

To conclude, we feel that weighted Löwdin orbitals with properly deorthogonalized pairs of hybrids have the necessary flexibility for construction of optimal basis sets, which in turn are "conditio sine qua non" in designing theoretically sound and efficient approximate methods applicable to larger molecular systems. Performance of various weighted Löwdin orthogonal (and pairwise non-orthogonal) basis sets in the context of NDDO schemes will be reported in forthcomming paper(s).

Acknowledgement

I would like to thank A. von Humboldt Foundation for a financial support and Prof. R. Gleiter for hospitality.

[1] F. London, J. Phys. Radium **8**, 397 (1937).
[2] H. F. Hameka, Rev. Mod. Phys. **34**, 87 (1962); Mol. Phys. **1**, 203 (1958); J. Chem. Phys. **34**, 366 (1961).
[3] R. Ditchfield, J. Chem. Phys. **56**, 5688 (1972); Chem. Phys. Letters **15**, 203 (1972).
[4] A. J. Sadlej, Chem. Phys. Letters **47**, 50 (1977); Theor. Chim. Acta **47**, 205 (1978); Mol. Phys. **34**, 731 (1977); Mol. Phys. **36**, 1701 (1978); Acta Phys. Pol. **A53**, 297 (1978).
[5] K. Wolinski and A. J. Sadlej, Chem. Phys. Letters **64**, 51 (1979).
[6] Z. B. Maksić and M. Eckert-Maksić, Tetrahedron **25**, 5113 (1969); Lj. Vujisić and Z. B. Maksić, J. Mol. Structure **7**, 431 (1971); Z. B. Maksić, Int. J. Quant. Chem. Symp. **5**, 301 (1971); Z. B. Maksić, Z. Meić and K. Kovačević, Z. Naturforsch. **28A**, 797 (1973).
[7] K. Kovačević and Z. B. Maksić, J. Org. Chem. **39**, 539 (1974); Z. B. Maksić and A. Rubčić, J. Amer. Chem. Soc. **99**, 4233 (1977).
[8] K. Kovačević, M. Eckert-Maksić, and Z. B. Maksić, J. Mol. Structure **21**, 335 (1974); K. Kovačević, M. Eckert-Maksić, and Z. B. Maksić, Croat. Chem. Acta **46**, 249 (1974); Z. B. Maksić and K. Kovačević, Chem. Phys. Letters **73**, 592 (1980).
[9] K. Kovačević, Z. B. Maksić, and A. Moguš, Croat. Chem. Acta **52**, 249 (1979); Theor. Chim. Acta **55**, 127 (1980); J. Mol. Structure (in print).
[10] K. Kovačević, K. Krmpotić, and Z. B. Maksić, Inorg. Chem. **16**, 1421 (1977); M. Eckert-Maksić, K. Kovačević, and Z. B. Maksić, J. Organomet. Chem. **168**, 295 (1979).
[11] D. B. Cook, P. C. Hollis, and R. McWeeny, Mol. Phys. **13**, 553 (1967).
[12] D. B. Cook and P. Palmieri, Mol. Phys. **17**, 271 (1969).
[13] D. B. Cook, Theoret. Chim. Acta **40**, 297 (1975); ibid., **46**, 291 (1977).
[14] D. B. Cook, Structures and Approximations for Electrons in Molecules, Ellis Horwood Limited, Chichester, Sussex, 1978.

[15] M. Klessinger and R. McWeeny, J. Chem. Phys. **42**, 3343 (1965).

[16] M. Klessinger, Chem. Phys. Letters **2**, 562 (1968); J. Chem. Phys. **46**, 3261 (1967); Chem. Phys. Letters **4**, 144 (1969).

[17] A. Tsuchida and K. Ohno, J. Chem. Phys. **39**, 600 (1963).

[18] P. O. Löwdin, J. Chem. Phys. **18**, 365 (1950); Adv. Quant. Chem. **5**, 185 (1970).

[19] B. C. Carlson and J. H. Keller, Phys. Rev. **105**, 102 (1957).

[20] P. S. Bagus, Phys. Rev. **139 A**, 619 (1965).

[21] L. C. Snyder, J. Chem. Phys. **55**, 95 (1971); P. S. Bagus and H. F. Schaeffer, III, ibid., **56**, 224 (1972).

[22] T. Živković and Z. B. Maksić, Croat. Chem. Acta **48**, 301 (1976).

[23] H. Fischer and H. Kollmar, Theor. Chim. Acta **16**, 163 (1970).