Variationally Optimized Numerical Orbitals for Molecular Calculations

James D. Talman
*Department of Applied Mathematics and Centre for Chemical Physics, University of Western Ontario,
London, Ontario, Canada N6A 5B7*

(Received 5 January 1999)

The problem of variational optimization of the atomic orbitals used in molecular calculations is investigated. It is shown that the variational principle leads to an equation similar to the radial Schrödinger equation but containing an inhomogeneous term. As an example, the equations are solved for the minimum basis set orbitals for the methane molecule. The results show a substantial improvement over those of a previous calculation optimizing in a minimum basis of Slater orbitals.

PACS numbers: 31.15.Pf, 31.15.Ne

Practically all contemporary calculations of molecular electronic properties are based on discrete basis approximations using Gaussian orbitals (GTOs) as proposed originally by Boys [1]; such calculations have come to be commonplace in modern chemistry using computer packages such as GAUSSIAN 94 [2]. GTOs are used, of course, since they render the problem of computing multicenter integrals, integrals of products of angular momentum orbitals centered at different points, almost trivial. On the other hand, GTOs are not very well suited for an accurate description of electronic properties in molecules since they behave analytically at the nuclei, whereas the true electronic wave functions exhibit a cusp behavior. In addition, at large distances, GTOs decrease more rapidly than the exponential decrease expected for bound state functions. In momentum space, GTOs also decrease in Gaussian fashion at large $|\mathbf{k}|$ whereas for an orbital of angular momentum l they should decrease like $|\mathbf{k}|^{-4-l}$.

These problems with GTOs are overcome to some extent by using superpositions of several GTOs to mimic the correct physical behavior and there is a very large amount of literature concerned with the construction of these "contracted" orbitals. However, the use of these superpositions defeats, to some extent, the advantage obtained from the simplicity of using these orbitals. At the same time, problems arise from the use of large numbers of orbitals. Therefore it may be useful to construct orbitals that are more physically realistic.

To a large extent, atomic calculations are based on numerical orbitals, i.e., single-particle functions that are defined on a numerical mesh. Such orbitals have the advantage of complete flexibility [3]. The use of such orbitals is, of course, feasible because of the spherical symmetry of the atomic problem, which is lost in the molecular situation. Atomic orbitals are usually computed variationally using the Hartree-Fock or multiconfiguration Hartree-Fock methods giving results of high accuracy. Since the energies in GTO calculations depend on the orbital parameters, i.e., the ζ parameters, in a very complicated way, it is not in general feasible to optimize these calculations on the parameters, and it is usual practice to use sets of standardized orbitals. A recent paper by Lee and Head-Gordon [4] reviews and discusses the possibility of making partial optimizations.

In the past, I have proposed several methods for computing the multicenter integrals required for molecular calculations for arbitrary, numerical orbital basis functions [5,6]. These methods employ an efficient method [7] for computing spherical Hankel transforms required for the transformation of angular momentum wave functions from ordinary space to momentum space and vice versa. Techniques have been described for computing overlap integrals, nuclear attraction integrals, and two-electron integrals, and the feasibility of applying these methods has been demonstrated in a MO-LCAO calculation for the methane molecule [8]. However, these methods are of limited utility unless the numerical orbitals can be constructed to be physically realistic.

In this Letter the problem of variational optimization of the radial factors in the atomic orbitals that form the basis of an MO-LCAO calculation is discussed. The principal complication in going beyond the radial Hartree-Fock equations for atoms is the nonorthogonality of orbitals on different centers; the main result is to provide a method for dealing with this problem. The equations to be solved resemble, but are somewhat different from, the radial Hartree-Fock equations. As an example, the equations will be solved for the electronic ground state of the methane molecule in the minimum basis.

In an MO-LCAO calculation, a Slater determinant is formed from molecular orbitals ϕ_i that are linear combinations of atomic orbitals.

$$\phi_i(\mathbf{r}) = \sum_{k\le i} d_{ki}\chi_k(\mathbf{r}), \tag{1}$$

where $\chi_k(\mathbf{r})$ is of the form $f_{lm}(\mathbf{r}-\mathbf{a}_k)$ where $f_{lm}(\mathbf{r}) = \chi_k(r)Y_{l_k m_k}(\hat{r})$ and is centered at $\mathbf{a}_k$, usually one of the nuclear coordinates. The $\phi_j(\mathbf{r})$ are constructed to be orthonormal using the Gram-Schmidt procedure. This is implemented by forming the overlap matrix Ω with elements

$$\omega_{ij} = \int \chi_i(\mathbf{r})\chi_j(\mathbf{r})\,d\mathbf{r} \tag{2}$$

and making the Cholesky factorization $\Omega = R^T R$, where R is upper triangular. Then, if $D = R^{-1}$, $D^T \Omega D = I$ and the coefficients d_{ki} are the elements of D.

The Hamiltonian for the problem is

$$\mathcal{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_{i,n}\frac{Z_n}{|\mathbf{r}_i - \mathbf{R}_n|} + \sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{3}$$

where Z_n and $\mathbf{R}_n$ are the charge and position of nucleus n. Orthonormal Hartree-Fock orbitals ψ_i are formed as linear combinations of the ϕ_j:

$$\psi_i(\mathbf{r}) = \sum_j u_{ji}\phi_j(\mathbf{r}) = \sum_{kj} d_{kj}u_{ji}\chi_k(\mathbf{r}) = \sum_k c_{ki}\chi_k(\mathbf{r}). \tag{4}$$

The total energy is then given, in terms of matrix elements with respect to the χ_k, by

$$\begin{aligned} E &= \sum_{kk'} w_{kk'}\langle k|H_{\mathrm{sp}}|k'\rangle + \frac{1}{2}\sum_{klk'l'}\left[w_{kk'}w_{ll'} - \frac{1}{2}w_{lk'}w_{kl'}\right]\langle kl|v|k'l'\rangle \\ &= \sum_{kk'} w_{kk'}\langle k|H_{\mathrm{sp}}|k'\rangle + \frac{1}{2}\sum_{klk'l'} w_{kk'}w_{ll'}\left[\langle kl|v|k'l'\rangle - \frac{1}{2}\langle lk|v|k'l'\rangle\right], \end{aligned} \tag{5}$$

where H_{sp} is the sum of the kinetic energy operator and the nuclear attraction potentials. The $w_{kk'}$ are elements of the density matrix W:

$$w_{kk'} = 2\sum{}' c_{ki}c_{k'i}, \tag{6}$$

where the prime on the sum indicates that it is on the occupied single-particle states. (The simplest situation in which there are an even number n of electrons, and the lowest $n/2$ states are doubly occupied, is being considered.) Minimizing E with respect to the coefficients c_{ki} leads to the matrix Hartree-Fock equation

$$\hat{H}_0\mathbf{v}^i = \lambda_i\Omega\mathbf{v}^i, \tag{7}$$

where $\hat{H}_0$ is the matrix with elements

$$\hat{h}^0_{kk'} = \langle k|H_{\mathrm{sp}}|k'\rangle + \sum_{ll'} w_{ll'}\left[\langle kl|v|k'l'\rangle - \frac{1}{2}\langle lk|v|k'l'\rangle\right]. \tag{8}$$

The generalized eigenvalue problem (7) can be replaced by the usual equation

$$\hat{H}\mathbf{u}^i = \lambda_i\mathbf{u}^i, \tag{9}$$

where $\hat{H} = D^T\hat{H}_0 D$, $\mathbf{v}^i = D\mathbf{u}^i$. This equation is solved iteratively with Eq. (6) to obtain the self-consistent density matrix $w_{kk'}$. The density matrix is then given by

$$W = 2D\sum{}' \mathbf{u}^i\mathbf{u}^{i^T} D^T = DPD^T, \tag{10}$$

where $P = 2\sum{}' \mathbf{u}^i\mathbf{u}^{i^T}$.

The variational problem for the optimized orbitals is

$$\frac{\delta E}{\delta\chi_m(\mathbf{r})} = 0. \tag{11}$$

E depends on $\chi_m(\mathbf{r})$ directly through the Hamiltonian matrix elements and indirectly through the elements of the density matrix. The direct terms are

$$\left[\frac{\delta E}{\delta\chi_m(\mathbf{r})}\right]_1 = 2\sum_{k'} w_{mk'}H_{\mathrm{sp}}\chi_{k'}(\mathbf{r}) + 2\sum_{k'}\sum_{ll'}\left\{\left[w_{mk'}w_{ll'} - \frac{1}{2}w_{lk'}w_{ml'}\right]v_{ll'}(\mathbf{r})\right\}\chi_{k'}(\mathbf{r}), \tag{12}$$

where

$$v_{ll'}(\mathbf{r}) = \int |\mathbf{r} - \mathbf{r}'|^{-1}\chi_l(\mathbf{r}')\chi_{l'}(\mathbf{r}')\,d\mathbf{r}' \tag{13}$$

is the potential produced by the product of the orbitals $\chi_l(\mathbf{r})$ and $\chi_{l'}(\mathbf{r})$.

The indirect terms are

$$\left[\frac{\delta E}{\delta\chi_m(\mathbf{r})}\right]_2 = \sum_{kk'}\hat{h}^0_{kk'}\frac{\delta w_{kk'}}{\delta\chi_m(\mathbf{r})} = \mathrm{Tr}\left[\hat{H}_0\frac{\delta W}{\delta\chi_m(\mathbf{r})}\right]. \tag{14}$$

From $W = DPD^T$,

$$\delta W = \delta DPD^T + D\delta PD^T + DP\delta D^T. \tag{15}$$

Since $D^T\Omega D = I$,

$$\delta D^T\Omega D + D^T\delta\Omega D + D^T\Omega\delta D = 0, \tag{16}$$

$$\delta D^T R^T + R\delta D = -D^T\delta\Omega D. \tag{17}$$

The first term on the left hand side is lower triangular, and the second term is upper triangular. The equation for δD can be solved by writing

$$D^T\delta\Omega D = \delta X = \delta X_U + \delta X_L \tag{18}$$

with $\delta X_L = \delta X_U^T$. It follows that

$$\delta D^T = -\delta X_L D^T \tag{19}$$

and similarly $\delta D = -D\delta X_U$.

It can be shown that the term $D\delta PD^T$ in Eq. (15) does not contribute:

$$\mathrm{Tr}[\hat{H}_0 D\delta P D^T] = \mathrm{Tr}[\hat{H}\delta P] = 2\sum_i{}' \mathrm{Tr}[\hat{H}\mathbf{u}_i \delta\mathbf{u}_i^T + \delta\mathbf{u}_i\mathbf{u}_i^T\hat{H}] = 4\sum_i{}' \lambda_i \mathrm{Tr}[\delta\mathbf{u}_i\mathbf{u}_i^T] = 4\sum_i{}' \lambda_i(\mathbf{u}_i, \delta\mathbf{u}_i) = 0\,. \quad (20)$$

(Since $\mathbf{u}_i$ is normalized, $\delta\mathbf{u}_i$ is orthogonal to $\mathbf{u}_i$.)

Therefore

$$\delta W = -D\delta X_U P D^T - DP\delta X_L D^T. \quad (21)$$

Now

$$[\delta\Omega/\delta\chi_m(\mathbf{r})]_{ij} = \sum_n [\delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}]\chi_n(\mathbf{r}), \quad (22)$$

$$[D\delta\Omega/\delta\chi_m(\mathbf{r})D^T]_{ij} = \sum_n [d_{im}d_{jn} + d_{in}d_{jm}]\chi_n(\mathbf{r})\,. \quad (23)$$

If Y_{mn} denotes the upper triangular part of the matrix with elements $d_{mi}d_{nj} + d_{ni}d_{mj}$,

$$\frac{\delta W}{\delta\chi_m(\mathbf{r})} = -\sum_n [DY_{mn}PD^T + DPY_{mn}^T D^T]\chi_n(\mathbf{r}), \quad (24)$$

$$\mathrm{Tr}\left[\hat{H}_0 \frac{\delta W}{\delta\chi_m(\mathbf{r})}\right] = -\sum_n \mathrm{Tr}[\hat{H}(Y_{mn}P + PY_{mn}^T)]\chi_n(\mathbf{r})\,. \quad (25)$$

The variational equation for the atomic orbitals is therefore

$$\sum_n \left\{H_{mn} - \frac{1}{2}\mathrm{Tr}[\hat{H}(Y_{mn}P + PY_{mn}^T)]\right\}\chi_n(\mathbf{r}) = 0\,, \quad (26)$$

where the operators H_{mn} are defined implicitly in Eq. (12). The radial factor χ_m satisfies the projection of this equation onto the appropriate angular momentum subspace centered at $\mathbf{a}_m$.

The resulting set of coupled equations can be solved self-consistently by viewing the equation arising from the variation of χ_m as an equation for χ_m with the other orbitals fixed. This assumes that χ_m remains normalized between iterations. The term in Eq. (26) with $m = n$ is then separated to give a differential equation for χ_m; the remaining terms then contribute an inhomogeneous term. The resulting Schrödinger-like radial equation is

$$-\frac{1}{2r}\frac{d^2}{dr^2}r\chi_m(r) + \frac{l_m(l_m+1)}{2r^2}\chi_m(r) + [V_d(r) - E_m]\chi_m(r) + Z_m(r) = 0\,, \quad (27)$$

where $Z_m(r)$ arises from the inhomogeneous terms. The constant E_m, which arises from the term with $m = n$, plays the role of a Lagrange multiplier and is chosen to maintain the normalization. A very effective iterative scheme to adjust E_m that converges in two or three iterations has been used. The equations are then solved iteratively to obtain a self-consistent solution. In the calculations, all the terms proportional to χ_m have been included in the form $V_d(r)\chi_m(r)$ for the purpose of the self-consistent solution.

The equation for χ_m is generated by expanding all the terms about the center $\mathbf{a}_m$. An angular momentum wave function $f_{lm}(\mathbf{r} - \mathbf{R})$ can be expanded as (regarding $\mathbf{a}_m$ as the origin)

$$f_{lm}(\mathbf{r} - \mathbf{R}) = \sum_{LM} F_{LM}(r,R)Y_{LM}(\hat{r})\,. \quad (28)$$

The functions $F_{LM}(r,R)$ are computed numerically using a Fourier transform technique. The procedure involves spherical Hankel transforms [8]. These numerical transforms are carried out very efficiently by using a logarithmic mesh for both the radial spatial and momentum meshes [7]. The expansion of the nuclear attraction potentials uses the Laplace expansion. The product of two such expansions of the form of Eq. (28) can be reexpressed in the same form using the usual expansion for the product of spherical harmonics.

To demonstrate the feasibility of the variational procedure, I have made calculations for the methane molecule in the minimum basis of $1s$, $2s$, $2p_{x,y,z}$ orbitals centered on the C nucleus, and $1s$ orbitals centered on the protons. Similar calculations have been made previously by Pitzer [9] using Slater orbitals, varying the nonlinear parameter ζ in the four different orbital forms; these provide a valuable comparison.

The calculations require infinite summations as implied in Eq. (28). These are truncated at various values. In the nuclear attraction three-center integrals, which are singly infinite, the summation is truncated at a value $L_{\max}^{(N)}$, and a Padé approximation method is used to accelerate the convergence.

The two-electron integral sums are doubly infinite if all four centers are different, and are singly infinite if both factors in a product are on the same center. The expansion of the product of two orbitals is truncated at a value $\Lambda_{\max}^{(1)}$ if the sum is singly infinite, and at $\Lambda_{\max}^{(2)}$ if the sum is doubly infinite. If the orbitals in a product are on different centers, they are expanded about the midpoint of the two centers in the total energy calculation and about the appropriate orbital center in the optimization calculation. $L_{\max}$ denotes the maximum used in the orbital expansions for the two-electron integrals. (The total energy calculation and the optimization are made independently.) The time requirement is largely determined by the calculation of the four-center, two-electron integrals with the orbitals on four different centers as governed by the parameter $\Lambda_{\max}^{(2)}$.

To obtain self-consistency, input and output values of $V_d(r)$ and $Z_m(r)$ are combined with weights of 0.4 and 0.6, respectively, at each iteration. Since the energy does not always decrease monotonically, the minimum over about ten iterations is determined.

Table I gives results of the optimization calculation with $L_{\max}^{(N)} = 12$, $L_{\max} = 14$, $\Lambda_{\max}^{(1)} = 10$, and $\Lambda_{\max}^{(2)} = 2, 4, 6$. The energy calculations are made with the same limits and $\Lambda_{\max}^{(2)} = 6$. The results are given for the values of

TABLE I. Total energies in hartrees of the methane molecule at three values of r_0, the C-H bond length. (a) Ref. [9]. (b) Value obtained with the orbitals of Ref. [9] and $\Lambda_{\max}^{(2)} = 6$. (c) Optimized value with $\Lambda_{\max}^{(2)} = 2$. (d) Optimized value with $\Lambda_{\max}^{(2)} = 4$. (e) Optimized value with $\Lambda_{\max}^{(2)} = 6$.

Calculation/r_0	2.0	2.05	2.1
(a)	−40.125 68	−40.128 22	−40.126 98
(b)	−40.125 22	−40.127 79	−40.126 52
(c)	−40.180 02	−40.181 06	−40.178 36
(d)	−40.186 75	−40.187 90	−40.185 38
(e)	−40.187 28	−40.188 29	−40.185 62

r_0, the C-H bond length, used by Pitzer. Table I also shows the results obtained by Pitzer together with the result of applying the present program to the orbitals found by Pitzer; the latter results are in agreement to better than a mhartree. The number of iterations to obtain a minimum energy increases from 2 or 3 for $\Lambda_{\max}^{(2)} = 2$ to 7 or 8 for $\Lambda_{\max}^{(2)} = 6$.

It is seen that the orbital optimization makes a substantial improvement, of the order of 60 mhartrees, even on the optimized Slater orbitals of Pitzer. A very large calculation using Slater orbitals with fixed ζ parameters with 14 s and 18 p orbitals on the C nucleus and two s orbitals on each proton has been made by Woznick [10] at $r_0 = 2.0665$. The result of −40.181 is in close agreement with these results. A calculation at the same value of r_0 by Krauss [11] using GTOs with 9 s orbitals, 12 p orbitals on the C nucleus, and 3 s orbitals on each proton has given a somewhat higher energy of −40.1668. Hashimoto and Osamura [12] have made a complete optimization of the orbital parameters in a GTO calculation using 9 s orbitals and 15 p orbitals on the C nucleus and 4 s orbitals on each proton. Their result of −40.189 01 at $r_0 = 2.067$ is slightly below the present results.

The accuracy of these calculations is limited by the finite difference approximation, the truncation of the angular momentum sums, and the degree of self-consistency. The calculations have been made on a 256 point mesh, and the finite difference approximation appears to be a negligible source of error. The error due to lack of self-consistency is also negligible. The error arising from the angular momentum truncations seems to be most sensitive to the values of $L_{\max}$ and $\Lambda_{\max}^{(1)}$ and is substantially less than a mhartree.

This study illustrates the feasibility of numerically optimizing the orbitals used in molecular calculations. The use of such optimized orbitals can substantially reduce the size of basis sets required and thereby reduce or eliminate some of the problems associated with using very large bases. At the very least, the optimized orbitals can provide valuable benchmarks for comparison with or construction of contracted Gaussian orbitals.

Although the development here has been based on the single-configuration MO-LCAO approach, the same methods will be applicable to other methods for molecular electronic structure calculations.

This work has been supported by the Natural Sciences and Engineering Research Council of Canada.

[1] S. F. Boys, Proc. R. Soc. London A **200**, 542 (1950).

[2] GAUSSIAN 94, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Forestman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople (Gaussian, Inc., Pittsburgh, PA, 1995).

[3] C. Froese Fischer, T. Brage, and P. Jönsson, *Computational Atomic Structure: An MCHF Approach* (IOP, Philadelphia, 1997).

[4] M. S. Lee and M. Head-Gordon, J. Chem. Phys. **107**, 9085 (1997).

[5] J. D. Talman, J. Chem. Phys. **80**, 2000 (1984).

[6] J. D. Talman, J. Chem. Phys. **84**, 6879 (1986).

[7] J. D. Talman, J. Comput. Phys. **29**, 35 (1978); Comput. Phys. Commun. **30**, 93 (1983).

[8] J. D. Talman, Int. J. Quantum Chem. Quantum Chem. Symp. **27**, 321 (1993).

[9] R. M. Pitzer, J. Chem. Phys. **46**, 4871 (1967).

[10] B. J. Woznick, J. Chem. Phys. **40**, 2860 (1964).

[11] M. Krauss, J. Chem. Phys. **38**, 564 (1963).

[12] K. Hashimoto and Y. Osamura, J. Chem. Phys **95**, 15 (1991).