Local-Energy Method in Electronic Energy Calculations

ARTHUR A. FROST, REID E. KELLOGG, AND EARL C. CURTIS

Department of Chemistry, Northwestern University, Evanston, Illinois

INTRODUCTION

ELECTRONIC energies of molecules are calculated nonempirically almost exclusively by the Ritz variation method. As long as the approximate wave function is expressed as linear combinations and products of atomic orbitals the resulting fragmentary integrals involve no more than two electrons and four centers. Several groups are successfully evaluating such integrals with the aid of high-speed digital computers. However, even with the integrals available it is still an enourmous problem to carry out the necessary configuration interaction, or what is equivalent, the mixing of valence bond covalent and ionic structures, in order to attain the desired accuracy.

An alternative conceptual scheme for constructing a suitable wave function involves the concept of correlated molecular orbitals in which the interelectronic distances are included in the wave function by way of a correlation factor. In the helium atom this was first done by Hylleraas¹ and recently extended by Weiss and Roothaan² with excellent success. For the hydrogen molecule this idea has been shown to have conceptual value and also to provide a surprisingly good energy.3-5 The extension of this method to many-electron molecules or atoms leads to integrals of great difficulty.

It can be concluded that any method that can provide sufficiently accurate solutions of the Schroedinger equation without such time-consuming or difficult integrations should be most acceptable.

CRITERION OF CONSTANT ENERGY

Dividing the general Schroedinger equation $H\psi =$ $E\psi$ by ψ yields

$$H\psi/\psi = E.$$
 (1)

 $H\psi$ and ψ are both functions of the coordinates of the electrons. Therefore, for each point in configuration space a "local" energy E may be evaluated. Bartlett⁶ has pointed out that for the exact solution E is a constant whereas an approximate wave function ϕ

¹ E. Hylleraas, Z. Physik **65**, 209 (1930). ² A. W. Weiss and C. C. J. Roothaan, Bull. Am. Phys. Soc. Ser. II; **4**, 173 (1959); see also C. C. J. Roothaan and A.W. Weiss, Revs. Modern Phys. **32**, 194 (1960), this issue.

will lead to a local energy ϵ calculated from $H\phi/\phi$ as a varying function of position. It is then clear that the constancy of ϵ can, in principle, be used as a criterion of the excellence of a wave function

$$H\phi/\phi = \epsilon.$$
 (2)

The use of this criterion has the great advantage over the variation method in that integration is not required. The Hamiltonian being a differential operator. it is always possible to evaluate $H\phi$ no matter how complicated the wave function. A least squares method of approximation was outlined and tested some years ago on the hydrogen molecule-ion.7 In the absence of modern computing facilities it was considered impractical to extend the method at that time. A. von Mohrenstein⁸ modified the scheme and applied it to the helium atom and the hydrogen molecule, again using only desk computers.

More recently Bartlett⁹ used $H\phi/\phi$ as a test of a numerically calculated wave function for helium. This application differs from that of Frost and of von Mohrenstein in that Bartlett does not make use of the criterion of constant energy to improve the original function. Furthermore, his function is not an analytical expression.

In this paper the original least squares method is improved and tested more adequately using a mediumspeed digital computer. A subsequent paper will describe attempts to get constancy of local energy by analytical procedures. In both cases the hydrogen molecule-ion is used as a test case.

LEAST SQUARES METHOD

The general method introduced in 1942 is first summarized with improved notation.

Let ϕ be a variation function with one or more parameters to be determined. H is the Hamiltonian operator for the given system. Pick a suitable set of representative points in configuration space \mathbf{r}_p where p =1, 2, $\cdots N$. For each point, identified by p, calculate $\epsilon_p = H\phi(\mathbf{r}_p)/\phi(\mathbf{r}_p)$. Now evaluate the mean energy $\bar{\epsilon}$ and the mean square energy $\langle \epsilon^2 \rangle_{av}$ using an appropriate weighting factor g_p :

$$\bar{\epsilon} = \sum_{p} g_{p} \epsilon_{p} / \sum_{p} g_{p}, \qquad (3)$$

$$\langle \epsilon^2 \rangle_{\rm av} = \sum_p g_p \epsilon_p^2 / \sum_p g_p.$$
 (4)

^a A. A. Frost, J. Braunstein, and W. Schwemer, J. Am. Chem. Soc. **70**, 3292 (1948).

⁴A. A. Frost and J. Braunstein, J. Chem. Phys. 19, 1133 (1951).

⁵ W. Kolos and C. C. J. Roothaan, Bull. Amer. Phys. Soc. Ser. II, 4, 174 (1959). ⁶ J. H. Bartlett, Phys. Rev. 51, 661 (1937).

⁷ A. A. Frost, J. Chem. Phys. 10, 240 (1942).

⁸ A. von Mohrenstein, Z. Physik 134, 488 (1953).
⁹ J. H. Bartlett, Phys. Rev. 98, 1067 (1955).

The variance V or mean square deviation of ϵ from the mean is

$$V = \langle \epsilon^2 \rangle_{\rm av} - (\bar{\epsilon})^2. \tag{5}$$

The method consists, then, of minimizing V for all possible variations in the parameters of the wave function ϕ .

APPLICATION TO LINEAR VARIATION FUNCTIONS

Let

$$\phi = \sum_{i=1}^{n} c_i \phi_i$$

where the ϕ_i are not necessarily orthonormal. It is convenient in simplifying the algebra to take the weighting factor as

$$g_p = w_p \phi(\mathbf{r}_p)^2, \tag{6}$$

where w_p is typically the volume element associated with a given point.

Upon substitution in (2)-(4) and introducing matrix components defined as

$$S_{ij} = \sum_{p} w_{p} \phi_{i}(\mathbf{r}_{p}) \phi_{j}(\mathbf{r}_{p}), \qquad (7)$$

$$H_{ij} = \frac{1}{2} \sum_{p} w_{p} \left[\phi_{i}(\mathbf{r}_{p}) H \phi_{j}(\mathbf{r}_{j}) + \phi_{j}(\mathbf{r}_{p}) H \phi_{i}(\mathbf{r}_{p}) \right], \quad (8)$$

$$G_{ij} = \sum_{p} w_{p} H \phi_{i}(\mathbf{r}_{p}) H \phi_{j}(\mathbf{r}_{p}), \qquad (9)$$

there results

$$\bar{\epsilon} = \sum_{i,j} c_i c_j H_{ij} / \sum_{i,j} c_i c_j S_{ij}, \qquad (10)$$

$$\langle \epsilon^2 \rangle_{\rm av} = \sum_{i,j} c_i c_j G_{ij} / \sum_{i,j} c_i c_j S_{ij}.$$
(11)

The matrix components H_{ij} , Eq. (8), are defined in symmetrical form for convenience. This is permissable because their only use is by way of double sums as in (10) and succeeding equations. Putting (10) and (11) in (5), the variance expression becomes

$$V = \frac{(\sum_{ij} c_i c_j G_{ij}) (\sum_{i,j} c_i c_j S_{ij}) - (\sum_{i,j} c_i c_j H_{ij})^2}{(\sum_{ij} c_i c_j S_{ij})^2}.$$
 (12)

The conditions that V be stationary, i.e., that

$$\partial V/\partial c_k = 0$$
 for all k , (13)

requires solution of a set of n simultaneous equations which can be put in the form

$$\sum_{i} c_{i} [G_{ik} - 2\bar{\epsilon}H_{ik} + (\bar{\epsilon}^{2} - V) S_{ik}] = 0,$$

$$k = 1, 2, \cdots n. \quad (14)$$

These variation equations are analogous to the usual set for a typical eigenvalue problem, but in this case there are two parameters $\bar{\epsilon}$ and V and the one comes in quadratically. Equations (14) are a set of nonlinear

equations, which is expected considering the fourthdegree terms in the numerator and denominator of (12).

METHODS FOR SOLVING THE VARIATION EQUATIONS

Two methods have been used for the approximate solution of Eq. (14) that express the condition that V be stationary.

The *first method* is a *relaxation* method as follows:

1. Assume a starting set of coefficients c_i .

2. Calculate the corresponding ϵ and V using (10), (11), and (5).

3. Solve the *n*th equation [(14)] in a form that will give Δc_n , the change in c_n .

4. Repeat 2.

5. Solve the (n-1)-th equation [(14)] for Δc_{n-1} .

6. Continue through all n equations [(14)] in this way.

7. Repeat the cycle of relaxations of Eq. (14) as many times as are required to make ΔV as small as desired.

The *second method* is an *iteration* method which was devised later and which in the calculations so far performed is faster by a factor of about ten in machine calculations. This method consists of:

1. Substitution of approximate values of $\bar{\epsilon}$ and V in Eqs. (14).

2. Omit the first equation and set c_1 equal to unity.

3. Solve the remaining n-1 simultaneous inhomogeneous equations for the n-1 coefficients c_i .

4. Substitute these c_i into (10), (11), and (5) to obtain new values of $\bar{\epsilon}$ and V.

5. Repeat 1, 2, and 3 with the new $\bar{\epsilon}$ and V and continue the iteration until V converges to a minimum.

6. Alternatively one could begin at step 4 by guessing an approximate set of c_i .

TEST OF THE METHOD ON THE HYDROGEN MOLECULE-ION

The ground state wave function of the hydrogen molecule-ion was expressed by the nine-term function

$$\phi = \sum_{i=1}^{z} c_i \phi_i, \qquad (15)$$

where

$$\phi_i = \exp(-z\mu)\mu^m \nu^n, \tag{16}$$

 μ and ν being elliptical coordinates, and $i=1, 2, \dots 9$ corresponding to the *m*, *n* values 0, 0; 0, 2; 0, 4; 1,0; 1, 2; 1,4; 2, 0; 2, 2; 2, 4. The first two terms constitute the function used so successfully by James¹⁰ in his variation method calculation.

In these test calculations two values of z have been used; 1.35, the best value found by James; and 1.48501, the value required to give the correct limiting energy

¹⁰ H. M. James, J. Chem. Phys. 3, 7 (1935).

 $H\phi_i/\phi_i$ is easily obtained as only the part involving the Laplacian operator need be calculated. Since, in atomic units,

$$H \equiv -\frac{1}{2} \nabla^2 - (r_a)^{-1} - (r_b)^{-1}, \qquad (17)$$

and in elliptical coordinates,

$$H = -\frac{1}{2} \nabla^2 - \left[\frac{4\mu}{R(\mu^2 - \nu^2)} \right], \tag{18}$$

there results

$$H\phi_i/\phi_i = -\left[2/R^2(\mu^2 - \nu^2)\right]\left[P_m(\mu) + Q_n(\nu)\right], \quad (19)$$

where

$$P_{m}(\mu) = [1 - (\mu^{2})^{-1}][(z^{2}\mu - 2mz)\mu + m(m-1)] + 2[(R-z)\mu + m] \quad (20)$$

and

$$Q_n(\nu) = n\{(n-1)[(\nu^2)^{-1} - 1] - 2\}.$$
(21)

The choice of points for evaluation of the local energy and the subsequent minimization of its variance appears to be a matter for experiment. Mohrenstein⁸ devised a method of choosing points of equal weight by spacing them farther apart as the square of the wave function diminishes. This method is probably more economical of points but has the disadvantage that a separate summation over points must be carried out for every new set of coefficients c_i that is tried. On the contrary in the present method, as outlined in the last section, the summations over the points are included in the mattrix component formulas (7)-(9) and are carried out once and for all regardless of the variation of the c_i .

For the present, consider that μ , ν space, out to an arbitrary maximum μ , is cut into equal areas according to successive equal intervals of μ and ν . Points are then chosen at the center of each area and are weighted according to Eq. (6) with

$$w_{p} = (R^{3}/8) (\mu^{2} - \nu^{2}) \Delta \mu \Delta \nu,$$
 (22)

which is equivalent to the volume element in elliptical coordinates. The angle φ does not appear inasmuch as the ground state is σ and therefore cylindrically symmetric, but the increasing weight of points further re-

moved from the axis due to integration over a circle of increasing radius is automatically taken care of by the $(\mu^2 - \nu^2)$ term.

A set of 32 points was first chosen with initial value, interval, and final value of each coordinate as follows:

$$\mu = 1.125(0.25)2.875$$

$\nu = 0.125(0.25)0.875.$

These points correspond to the area $\mu = 1.0$ to 3.0 and $\nu = 0.0$ to 1.0. Because of the symmetry of the wave function with respect to the midplane it is necessary to take only positive ν . That μ covers the range only up to 3.0 is reasonable inasmuch as the principal factor in the weighting function $\exp(-2z\mu)$ drops to a fraction $\exp(-4z)$, or about 4.5×10^{-3} of its maximum value. At the midplane, $\exp(-2z\mu)(\mu^2 - \nu^2)$ drops to a fraction of 4×10^{-2} .

An additional set of eight points defined as

$$\mu = 3.25(0.5)4.75$$

$$\nu = 0.25(0.5)0.75$$

was added to the original 32 to make a 40-point calculation. These points have double intervals in μ and ν over what was used in the original 32. The $\Delta\mu\Delta\nu$ of the weighting factor (22) automatically weights these points quadruply, nevertheless, the contribution of these points to the more important matrix components is only a few percent.

The calculations were carried out on an IBM 650 computer at the Northwestern Computing Center. After computing the matrix components, the mean energy $\bar{\epsilon}$ and the coefficients c_i were determined by solving Eq. (14) by relaxation. First the 2×2 system was solved using a starting vector $c_1 = c_2 = 1$. The system is successively expanded by one more term with several cycles of relaxation before adding another until finally the nineterm system is obtained. For an *n*-term system a cycle of n successive relaxations are carried out determining successive new values of c_n , c_{n-1} , $\cdots c_1$. The cycle is repeated until the variance changes no more than $1 \times$ 10^{-6} atomic units in a cycle. In the nine-term stage the relaxation cycle was continued to a maximum change in V of 2×10^{-7} . Table I shows the result for the 40-point calculation.

TABLE I. Hydrogen molecule-ion calculation: 40 points. Mean energy $\bar{\epsilon}$, variance V, standard deviation $(V)^{\frac{1}{2}}$, and coefficients c_i ; atomic units used throughout; R=2.00; z=1.35; $c_1=1.00$.

Terms	—ē	V×10 ⁶	$(V)^{\frac{1}{2}} \times 10^{3}$	C2	C3	C4	C5	C ₆	C7	C ₈	C 9
1	1.07839	55050.0	234.6			•••		•••			
2	1.10201	662.5	25.7	0.4369					•••	• • • · · · · · · · ·	• • • • · · ·
3	1.10387	365.5	19.1	0.4022	0.0469					•••	• • •
4	1.10379	341.9	18.5	0.4060	0.0501	0.00878	• • • •	• • • •		· · · · · ·	••••
5	1.10368	336.8	18.4	0.4020	0.0474	0.00785	0.00552		•••		•••
6	1.10377	335.2	18.3	0.4013	0.0458	0.00731	0.00517	0.00250			
7	1.10317	39.6	6.3	0.4202	0.0559	0.0819	0.00851	-0.00319	-0.0329		•••
8	1.10309	32.7	5.7	0.4223	0.0516	0.0868	0.0121	0.000272	-0,0249	-0.00309	• • •
9	1.10304	30.1	5.5	0.4220	0.0473	0.0896	0.0142	0.00531	-0.0258	-0.00390	-0.00125

No. of points	z	—ē	V×106	<i>c</i> 1	<i>C</i> 2	C3	C4	C5	C 6	C7	C ₈	Cy
40 32 40 32	$1.35 \\ 1.35 \\ 1.485 \\ 1.485 \\ 1.485$	1.10294 1.10193 1.10276 1.10233	25.3 7.1 5.8 2.1	$1.0 \\ 1.0 \\ 1.0 \\ 1.0 \\ 1.0$	$\begin{array}{c} 0.4170 \\ 0.4013 \\ 0.4119 \\ 0.4029 \end{array}$	0.02215 0.04931 0.03372 0.04895	0.08192 0.09896 0.25874 0.26837	0.01486 0.04450 0.09411 0.11120	0.0429 0.00409 0.03499 0.0130	$\begin{array}{r} -0.0237 \\ -0.0301 \\ -0.0216 \\ -0.0249 \end{array}$	$\begin{array}{r} -0.00293 \\ -0.01356 \\ -0.00482 \\ -0.01087 \end{array}$	$\begin{array}{r} -0.01489 \\ -0.00121 \\ -0.00907 \\ -0.00118 \end{array}$

TABLE II. Hydrogen molecule-ion calculation using iteration method. Exponential parameter z, mean energy $\bar{\epsilon}$, variance V, coefficients c_i ; atomic units used throughout; R=2.00.

In the usual system of a given number of terms the procedure required only three or four cycles of relaxation. The seven-term system was the exception and needed 19 cycles to reduce ΔV to 1×10^{-6} . Perhaps the reason for this is that the seventh term is the first introduction into the wave function of the second degree in μ . It is noticed that V drops by a surprisingly large amount at this stage.

The mean energy is already very good, to within 0.0006 a.u. of the exact energy -1.1026, with the twoterm-function. Addition of more terms does not improve the mean energy but does improve the variance. That the energy is not better may be due to poor distribution of points either in the neighborhood of a nucleus where the local energy is likely to have a singularity or in the region far removed from the nuclei. That the mean energy goes below the accepted value is quite possible inasmuch as the variation theorem does not apply to this method.

Table II shows the results obtained by the interaction method for 32-point and 40-point calculations using both z=1.35 and z=1.48501. The latter value of z is obtained by letting μ go to infinity in Eqs. (19) and (20), thus getting $\epsilon = -z^2/2$ and setting ϵ equal to the known correct value -1.10263 a.u. This value of zalso occurs in the accurate solution as summarized by Bates et al.¹¹ The results show that with the nine-term function the energy is rather insensitive to the value of z although the latter value is slightly better. Probably in future calculations on other systems it will be possible to predetermine the corresponding quantity or quantities.

Comparison with Table I for the 40-point z=1.35 calculation shows that the iteration method gives an improved result in terms of V and $\bar{\epsilon}$ as compared with the relaxation method. The coefficients c_i shown in Table II are likely to be more correct.

For both z's the 32-point calculation results in a smaller V than does the 40-point calculation. This is not unexpected inasmuch as a calculation with as few as nine points should result in zero V since only nine coefficients are to be determined. On the other hand, increasing the member of points and keeping them well distributed should lead asymptotically to a fixed V determined only by the particular set of approximate functions used in the linear combination. Possibly 40 points already is a large set from this standpoint.

It may have been noticed that no use has so far been made of the fact that the H_2^+ problem is separable in elliptical coordinates. If this is the case the coefficients c_i should be expressible as $c_i = a_j b_k$ where a_j and b_k are coefficients of powers of μ and of ν , respectively. A solution of this form was obtained by first assuming the coefficients a_i , solving for b_k by relaxation, substituting the latter and solving for a new set of a_i by relaxation, repeating until V was minimized. This technique gave a quicker and slightly better result than did the relaxation method applied to the nine-term function, but it was not as good as the second iteration method. For example, for z=1.35 and 40 points, $V \times 10^6$ for the nine-term relaxation method was 30.1, while using separability, a value of 26.5 was obtained and with the iteration method 25.3. Separability in the present methof of calculation presumably depends on the choice of points used in the calculation.

The accuracy of the mean energy can be judged to some extent by the standard deviation which in the 40-point z=1.485 calculation has diminished to less than 3×10^{-3} . It would be desirable to get this down to less than 1×10^{-3} . Unfortunately, at this point the variance is such a small difference between two large numbers, $\langle \epsilon^2 \rangle_{av}$ and $(\bar{\epsilon})^2$, that only one significant figure remains even though each term is obtained to eight digits.

Nevertheless, the mean energy $\bar{\epsilon}$ in the z=1.485 calculations is within 3×10^{-4} of the correct value. It would be very acceptable for chemical purposes if molecular energies in general could be calculated to

TABLE III. Comparison of sums and integrals.

	S_{11}	H ₁₁	ē1					
$\overline{\begin{array}{c} 32 \text{-point sums} \\ 1 < \mu < 3 \end{array}}$	0.0407762	-0.0440059	-1.07920					
Integrals up to $\mu = 3$	0.0405746	-0.0438616	-1.08101					
$\begin{array}{c} 40 \text{-point sums} \\ 1 < \mu < 5 \end{array}$	0.0419659	-0.0452555	-1.07839					
Integrals up to $\mu = 5$	0.0415178	-0.0448507	-1.08027					
Integrals to $\mu = \infty$	0.0418605	-0.0451955	-1.07967					

¹¹ D. R. Bates, K. Ledsham, and A. L. Stewart, Phil. Trans. Roy. Soc. (London) A246, 215 (1953).

.

this accuracy since 3×10^{-4} a.u. of energy is only 8×10^{-3} ev or about 0.2 kcal per mole.

RELATION TO NUMERICAL INTEGRATION

The choice of points and weighting factors is reminiscent of a numerical integration. Although the present method does not depend upon the summations being approximations to integrals, the later possible use of such functions to calculate expectation values of other quantum mechanical quantities would have to depend upon integration. Therefore it is of interest to see how closely a given matrix component defined by a limited summation approaches to the corresponding integral.

As an example of an integral

$$S_{11} = (R^3/8) \int_1^\infty \int_0^1 \exp(-2z\mu) (\mu^2 - \nu^2) d\mu d\nu,$$

Table III shows the corresponding sums and integrals for both S_{11} and H_{11} and the ratio $\bar{\epsilon}_1$. The sums agree with the corresponding integrals to within 1.0%. The energy values are much better and agree to within 0.2%.

ACKNOWLEDGMENTS

One of us (A. A. F.) had helpful discussions with J. O. Hirschfelder, M. P. Barnett, and C. C. J. Roothaan. This research was supported by the National Science Foundation and the Office of Ordnance Research.

REVIEWS OF MODERN PHYSICS

VOLUME 32, NUMBER 2

APRIL, 1960

Relativistic Corrections in Many-Electron Systems*

ANDERS FRÖMAN

Quantum Chemistry Group, Uppsala University, Uppsala, Sweden

I. INTRODUCTION

IN almost every application of quantum mechanical methods to the methods to the description of many-electron systems, it has been assumed that a nonrelativistic treatment should give results accurate enough. This assumption is very seldom checked, and in the few cases where an approximate relativistic calculation has been carried out, the results indicate that for many systems the preceding assumption is not justified. In the present paper the interest is focused on the shifts in the total energies caused by relativistic effects, and because almost nothing is known so far about these shifts in molecular systems, results are given only for atomic systems. The limited material presented here does not imply that the theory of relativity has been completely forgotten in quantum chemistry. In fact, a great deal of work has been performed in studying the effect of spin-orbit interaction in atomic and molecular spectra and relativistic effects have also been considered in the theories of the Auger spectra¹ and of the cohesive energies of alkali metals,² to mention a few examples. But because the concept of the total energy of a system plays such a dominating role in the formulation of quantum mechanics, in particular in constructing the wave equation determining the wave function, the influence of the theory of relativity on the total energy is of special importance. This is even more obvious from the fact that the only possible way at present to calculate an energy difference, e.g., dissociation energies and spectra, is to calculate the total energies for the two states involved and then take the difference.

As an introductory and illustrative example of cases where relativistic effects are of importance, we may consider a hydrogenlike system with nuclear charge Z. In the ground state of this system the electron has, according to the Schrödinger theory, a root-meansquare velocity of the order of $2.2 \times Z \times 10^8$ cm/sec, which means that even for fairly small Z the velocity of the electron is comparable to the velocity of light. By using classical arguments for Z=92, one should expect a relativistic correction of about 50% of the nonrelativistic energy. In the relativistic quantum mechanics as developed by Dirac, the correction predicted in this case is around 15%. The structure of the Dirac theory is in many respects different from the nonrelativistic Schrödinger theory. The main differences are that the wave function for one electron has four components which are determined from first-order equations and that the spin of the electron from the beginning is built into the theory and not added as an extra postulate as in the Schrödinger theory. This inclusion of spin leads to relativistic terms which have no counterpart in classical relativistic mechanics. Apart from some small

^{*} Work supported in part by the King Gustaf VI Adolf's 70-Years Fund for Swedish Culture, Knut and Alice Wallenberg's Foundation, the Swedish Natural Science Research Council, and in part by the Aeronautical Research Laboratory, Wright Air Development Center of the Air Research and Development Command, U. S. Air Force through its European Office under a contract with Uppsala University.

with Uppsala University. ¹W. N. Asaad, Proc. Roy. Soc. (London) **A249**, 555 (1959). ²J. Callaway, R. D. Woods, and V. Sirounian, Phys. Rev. **107**, 934 (1957).