

CONSTRAINED VARIATION METHOD IN MOLECULAR QUANTUM MECHANICS.

COMPARISON OF DIFFERENT APPROACHES*

by

D. P. Chong[†] and Yechezkel Rasiel[‡]

University of Wisconsin Theoretical Chemistry Institute

Madison, Wisconsin 53715

ABSTRACT

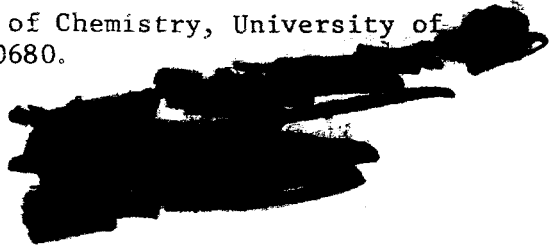
35-1743

Numerical results are presented for the lithium hydride molecule, obtained by different techniques of the constrained variation method. The constrained wave function, the degree to which the constraint is satisfied, and the sacrifice in energy are compared. The constraint operator used was the electronic dipole moment in one case and the total Hellmann-Feynman force in another. In both cases, the parametrization procedure in the direct solution approach (Method IB) was found to be superior.

* This research was supported by the following grant:
National Aeronautics and Space Administration Grant NsG-275-62.

[†] Present Address: Department of Chemistry, University of British Columbia, Vancouver 8, B.C., Canada.

[‡] Permanent Address: Department of Chemistry, University of Illinois, Chicago, Illinois 60680.



INTRODUCTION

The calculation of expectation values of operators other than the energy has been of considerable interest to quantum chemists. Recently Brown¹ formulated a perturbation approach to the constrained variation method which had previously been developed by Whitman and coworkers.^{2,3} To obtain a direct solution for the constrained wave function, Rasiel and Whitman² proposed an iterative scheme, but they did not carry it out much beyond the first iteration. The constrained wave function, however, can be determined accurately without any approximations in a straightforward manner when the basis set is small. The purpose of this paper is to present a comparison of numerical results for lithium hydride obtained by these methods and to discuss their relative merits.

CONSTRAINED VARIATION METHODS

Let \mathcal{h} be the Hamiltonian of the system and let ϕ_1 and ϵ_1 be the approximate ground state wave function and energy as calculated by a free variation. Consider an operator m which does not commute with the Hamiltonian

-
1. W. Byers Brown, University of Wisconsin Theoretical Chemistry Institute Report WIS-TCI-109G (1965).
 2. Y. Rasiel and D. R. Whitman, J. Chem. Phys., 42, 2124 (1965).
 3. D. R. Whitman and R. Carpenter, Bull. Am. Phys. Soc., 9, 231 (1964).

$$[m, h] \neq 0 \quad (1)$$

The constrained variation procedure involves determining a constrained wave function Ψ which minimizes the energy

$$E = \langle \Psi, h \Psi \rangle / \langle \Psi, \Psi \rangle \quad (2)$$

subject to a constraint

$$M = \mu, \quad (3)$$

where

$$M = \langle \Psi, m \Psi \rangle / \langle \Psi, \Psi \rangle \quad (4)$$

and μ is an experimental or theoretical constant within the bounds of M .

It is convenient to define a constraint operator

$$C = m - \mu \quad (5)$$

so that the constraint condition (3) becomes

$$C = M - \mu = 0, \quad (6)$$

where

$$c = \langle \Psi, \mathcal{C}\Psi \rangle / \langle \Psi, \Psi \rangle . \quad (7)$$

The constrained variation principle can be rewritten¹

$$\langle \delta\Psi, (\mathcal{H} - E)\Psi \rangle = 0 , \quad (8)$$

where \mathcal{H} is a (fictitious) constrained Hamiltonian

$$\mathcal{H} = h + \lambda \mathcal{C} \quad (9)$$

and λ is a Lagrangian multiplier, the determination of which will be discussed below.

Linear Variation

It is convenient to use the solutions ϕ_j of the free variation problem as the basis set, so that the overlap matrix \mathcal{S} is a unit matrix and the free Hamiltonian matrix h is diagonal, with $h_{ij} = \delta_{ij} \epsilon_i$. If the constrained wave function is written

$$\Psi = \sum_j a_j \phi_j , \quad (10)$$

where the coefficients a_j are to be determined, then

$$E = \frac{\sum_j a_j^2 \epsilon_j}{\sum_j a_j^2} \quad (11)$$

$$C = \frac{\sum_{i,j} a_i a_j C_{ij}}{\sum_j a_j^2} = 0 \quad (12)$$

where C is the matrix representation of the constraint operator \mathcal{C} in the ϕ basis.

The secular equation and secular determinant are

$$(H - E \mathbf{1}) \tilde{a} = 0 \quad (13)$$

$$D = \det \{ H - E \mathbf{1} \} = 0 \quad (14)$$

where

$$H = h + \lambda C \quad (15)$$

and λ still has to be determined.

Direct Solution

As a first *cycle* in the iterative scheme of Rasiel and Whitman^{2,4} to solve Eqs. (12) and (13) for the a_j and λ , the assumption

$$\Delta E \equiv E - \epsilon_1 \simeq \lambda C_{11} \quad (16)$$

is made in order to obtain approximate expressions for $a_j(\lambda)$ from Eq. (13). These are then substituted into Eq. (12), resulting in a vanishing polynomial in λ . Finally, all terms in λ^3 and higher are neglected and the subsequent quadratic equation in λ is solved for the root which gives a smaller positive ΔE . This solution will be called Method IA.

As Brown¹ pointed out,

$$\partial E / \partial \lambda = \langle \Psi, \mathcal{C} \Psi \rangle / \langle \Psi, \Psi \rangle = C = 0 \quad (17)$$

Therefore, λ and subsequently the a_j can be obtained by parametrization of D . In other words, Eq. (14) is solved with trial values of λ until interpolation and iteration lead to a value of λ at which E is a saddle point (minimum with respect to the coefficients a_j and maximum with respect to λ). However, numerical computations show that the energy E is not very sensitive to λ at the saddle point, so that λ can only be calculated to within ± 0.000005 . *Therefore*, at each trial value *instead* of λ , Eq. (13) is solved and the solutions a_j are substituted into Eq. (12). A few linear interpolations lead to the λ at which $C = 10^{-9}$, much beyond the accuracy of the input matrices. We shall call this Method IB.

4. Y. Rasiel, Ph.D. Thesis, Case Institute of Technology, 1964.

Perturbation Approach

On the other hand, Brown¹ considered the operator \mathcal{C} as a perturbation to \mathcal{h} with λ as the perturbation parameter. Thus, Eq. (17) leads to

$$E^{(1)} = \sum_{n=2}^{\infty} n \lambda^{n-1} E^{(n)}, \quad (18)$$

where $E^{(n)}$ is the familiar n-th order energy in perturbation theory and can be evaluated by standard formulas for expansions in unperturbed eigenfunctions as basis set.⁵ Inversion of the series in Eq. (18) then gives

$$\lambda = \theta - A_3 \theta^2 + [-A_4 + 2A_3^2] \theta^3 + [-A_5 + 5A_4 A_3 - 5A_3^3] \theta^4 + \mathcal{O}(\theta^5), \quad (19)$$

where

$$\theta = -A_1 = -E^{(1)} / (2E^{(2)}) \quad (20)$$

$$A_n = nE^{(n)} / (2E^{(2)}) . \quad (21)$$

-
5. J. O. Hirschfelder, W. Byers Brown, and S. T. Epstein, "Recent Developments in Perturbation Theory" in Advances in Quantum Chemistry, (Academic Press Inc., New York; 1964). The sign in Eq. (II-10) on p. 266 should be positive.

The energy and expectation values of other operators can be calculated in two ways which are closely related. In the first case, which will be called Method IIA, straightforward single and double perturbation procedures^{1,5} are applied. The sacrifice in energy is given¹ by

$$\Delta E = \sum_{n=1}^{\infty} \lambda^n E^{(n)}. \quad (22)$$

To find the expectation value L of some other operator \mathcal{L} , let us consider another fictitious Hamiltonian

$$\mathcal{H} = \mathcal{H} + \gamma \mathcal{L} = \mathcal{h} + \lambda \mathcal{C} + \gamma \mathcal{L}, \quad (23)$$

where both λ and γ are perturbation parameters; and let us assume that the minimum variational energy \mathcal{E} can be expanded as a double power series

$$\mathcal{E} = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \lambda^n \gamma^k \mathcal{E}^{(n,k)}. \quad (24)$$

Then, according to Brown,¹

$$L = \left(\frac{\partial \mathcal{E}}{\partial \gamma} \right)_{\gamma=0} = \sum_{n=0}^{\infty} \lambda^n \mathcal{E}^{(n,1)}. \quad (25)$$

In this work, we have continued the series in Eqs. (19), (22), and (25) up to terms involving θ^4 , $E^{(5)}$, and $\xi^{(3,1)}$ respectively, one more term in each case than Brown did. Owing to the fact that the basis set used in the present work is small, these series still have not converged when truncated.

Finally, in Method IIB, the first, second, and third order wave functions are determined using well-known perturbation theory methods.⁵ A truncated constrained wave function is then found

$$\Psi = \phi_1 + \lambda \psi^{(1)} + \lambda^2 \psi^{(2)} + \lambda^3 \psi^{(3)} \quad (26)$$

and expectation values are calculated with this function.

LiH WAVE FUNCTION

To demonstrate the application of the constrained variation method, Rasiel and Whitman,² and Brown¹ used a variational function for LiH calculated by Robinson.⁶ It is a three-term open-shell valence-bond configuration interaction function with a basis set of six Slater-type atomic orbitals,

$$\phi_1 = \sum_{k=1}^3 c_{1k} \eta_k, \quad (27)$$

6. J. M. Robinson, Ph.D. Thesis, University of Texas, 1957.

where the three configurations η_k have been given explicitly by Rasiel and Whitman.² This function was also chosen for our comparative studies for its simplicity and the availability of all the primitive matrix elements (in the η basis).^{6,2} However, during the course of our investigation, it was found that some of the off-diagonal elements of the transformed free Hamiltonian matrix (in the ϕ basis) are as large as 9×10^{-4} hartree. Therefore, the free variation problem was solved again. The solutions we obtained are

$$\begin{aligned}
 \phi_1 &= 0.1807573 \eta_1 - 0.0026450 \eta_2 + 0.1619637 \eta_3 \\
 \phi_2 &= -0.0737250 \eta_1 + 0.3520028 \eta_2 - 0.0697765 \eta_3 \\
 \phi_3 &= 0.8035686 \eta_1 + 0.3440879 \eta_2 - 0.8818488 \eta_3
 \end{aligned} \tag{28}$$

with

$$\begin{aligned}
 \epsilon_1 &= -9.058791 \text{ hartrees} \\
 \epsilon_2 &= -8.900329 \\
 \epsilon_3 &= -6.618735 .
 \end{aligned} \tag{29}$$

The electronic energy of the ground state is essentially the same as Robinson's - 9.058777 hartrees; but the ϕ_j are slightly different, with the result that the largest off-diagonal element of the transformed free Hamiltonian h is now 3×10^{-7} hartree.

Therefore, all the transformed matrices are computed in the ϕ basis represented by Eq. (28), rather than by Eq. (14) of Reference 2.

CONSTRAINT OPERATOR

Two cases have been considered in this work. In the first case, the electronic dipole moment is constrained to agree with the experimental value obtained by Wharton, Gold, and Klemperer.⁷ Thus,

$$m = e \sum_{i=1}^4 Z_i, \quad (30)$$

where Z_i is measured in a coordinate system in which the Li nucleus is at the origin and the H nucleus is on the positive Z-axis at the internuclear distance of $3.0132976 a_0$; after the nuclear contribution has been subtracted from the experimental dipole moment of 5.882 debyes,⁷ μ is $5.3276 ea_0$; and \mathcal{C} has been defined in Eq. (5) as $m - \mu$.

In the second case we studied, the constraint is that the total molecular Hellmann-Feynman force should vanish. In this case, the constraining procedure ceases to be semi-empirical

7. L. Wharton, L. P. Gold, and W. Klemperer, J. Chem. Phys., 37, 2149 (1962).

since the Hellmann-Feynman theorem⁸⁻¹⁰ requires that the expectation value of this force operator vanishes when averaged over the correct wave function. Thus, μ is zero and

$$m = e^2 \sum_{i=1}^4 \left\{ \frac{3 \cos \theta_i(\text{Li})}{[r_i(\text{Li})]^2} + \frac{\cos \theta_i(\text{H})}{[r_i(\text{H})]^2} \right\}, \quad (31)$$

where the spherical coordinate variables are identical to those of Miller, Gerhauser, and Matsen.¹¹

RESULTS AND DISCUSSION

The results of calculations according to the schemes IA, IB, IIA, and IIB are listed in Table I, where the electronic dipole moment has been constrained, and in Table II, where the total Hellmann-Feynman force is the constraint.

From Tables I and II, one can make the following general remarks about the success of the different constraining techniques. Firstly, the constrained wave function coefficients a_j as well as λ appear to be very similar in both the direct solution and perturbation approaches. However, the degree to

-
8. H. Hellmann, Einführung in die Quantenchemie (Deuticke, Leipzig, 1937), p. 285.
 9. R. P. Feynman, Phys. Rev., 56, 340 (1934).
 10. R. F. W. Bader and G. A. Jones, Can. J. Chem., 41, 255, 586 (1963); J. Chem. Phys., 38, 2791 (1963).
 11. J. Miller, J. M. Gerhauser, and F. A. Matsen, Quantum Chemistry, Integrals and Tables (University of Texas Press, Austin, 1959), p. 11.

which the imposed constraint is satisfied as measured by how closely

$C = \frac{\langle \Psi, \mathcal{C} \Psi \rangle}{\langle \Psi, \Psi \rangle}$ approaches zero, varies widely among the different techniques.

Comparing the two direct solution approaches, one can easily see that Method IB is superior to Method IA in accuracy, both as to adherence to the constraint and minimization of the energy. However, IA is more suitable for desk calculator work and took only half a minute on a CDC G-15 computer. On the other hand, with computer subroutines for matrix diagonalization widely available, IB can be easily programmed for electronic computers and requires about four minutes for each trial value of λ on the CDC G-15. In practice, we used IA to determine an approximate value of λ which was employed as an initial guess in IB.

It should be pointed out that Method IA represents the first cycle in an iterative scheme. One could substitute the value of λ into the terms in λ^3 and higher in the polynomial from Eq. (12) and solve the new quadratic equation for a better value of λ and consequently a better set of coefficients a_j . When convergence occurs, both λ and ΔE could be substituted into the higher order terms of Eq. (13) in order to improve on the assumption expressed by Eq. (16). However, considering the accuracy of Method IB, we feel that it hardly seems worth the effort to pursue further iterations beyond the first cycle as

carried out in IA.

While the direct solution approach is more successful in this case than the perturbation techniques due to the slow convergence of the latter, we believe that the latter should be recommended for functions with **larger** basis sets. As the basis set becomes much larger, Method IA loses its inherent simplicity and becomes unwieldy for desk calculators, while IB may take prohibitive computer time. Moreover, the perturbation approach is expected to converge faster since the perturbation can be better expressed in terms of the enlarged basis set. This expectation is realized in a preliminary study on a system with a large basis set.¹²

Actually, Methods IIA and IIB should be used in conjunction, since IIA furnishes the value of λ for use in IIB, while IIB gives a further check on the convergence of IIA and provides a constrained wave function as well as a direct calculation of the expectation value of the constraint operator.

To provide further comparison of the constraining techniques, the following other physical properties were calculated: the adherence to the virial theorem, $-T/E_{\text{total}}$; the molecular quadrupole moment measured relative to the center of mass, $Q_{(\text{cm})}$; diamagnetic contribution to the susceptibility, χ^{d} ; and diamagnetic contributions to the proton and lithium shielding constants, σ^{d} .

12. D. P. Chong, unpublished results, using as ϕ the 28-term LiH wave functions of J. C. Browne and F. A. Matsen, Phys. Rev., 135A, 1227 (1964), and private communication.

The results are presented in Tables III and IV. The physical constants used have been taken from the latest set recommended by the National Academy of Sciences-National Research Council.¹³

In order to facilitate comparison of using different constraints, C can be made dimensionless by defining a set of reduced quantities

$$\mathcal{C}^* = \mathcal{C} / c_{11} \quad (32)$$

$$C^* = C / c_{11} \quad (33)$$

$$c^* = c / c_{11} \quad (34)$$

$$\lambda^* = \lambda c_{11} \quad (35)$$

Then

$$\mathcal{H} = h + \lambda^* \mathcal{C}^* \quad (36)$$

and C^* takes on the value of unity for the free variation function ϕ_1 and the value of zero for the perfectly constrained function.

Examining the results from this point of view, we find that Method IB gives a λ^* of 0.01206925 for the dipole moment constraint and 0.005749273 for the total force constraint. Thus, the slightly better convergence of the perturbation schemes in the latter case can be better understood.

13. Physics Today, 17, 48 (1964).

ACKNOWLEDGMENTS

We are grateful to Professors S. T. Epstein and
W. Byers Brown for helpful discussions.

TABLE I. Comparison of constrained wave functions. The constraint is the electronic dipole moment.

	Free Variation			Direct solution			Perturbation		
	ϕ_1	Method IA	Method IB	Method IA	Method IB	Method IIA	Method IIA	Method IIB	
C, ea_0	-0.676760	+0.002182	$+2 \times 10^{-9}$	(0)	(0)	(0)	(0)	-0.036396	
Molecular dipole moment, debyes	-4.162	-5.888	-5.882	(-5.882)	(-5.882)	(-5.882)	(-5.882)	-5.789	
$10^3 \Delta E$, hartrees	0	5.113	5.074	4.962	4.962	4.962	4.962	4.453	
$10^3 \lambda$	0	-17.21	-17.833875	-17.191	-17.833875	-17.191	-17.191	(-17.191)	
a_1	1	0.984060	0.984277	0.984277	0.984277	0.984277	0.984277	0.986231	
a_2	0	0.177399	0.176163	0.176163	0.176163	0.176163	0.176163	0.164927	
a_3	0	-0.012449	-0.012871	-0.012871	-0.012871	-0.012871	-0.012871	-0.012158	

TABLE II. Comparison of constrained wave functions. The constraint is the total Hellmann-Feynman force.

	Free variation			Direct solution			Perturbation		
	ϕ_1	Method IA	Method IB	Method IA	Method IB	Method IIA	Method IIA	Method IIB	
$C, e^2 a_0^{-2}$	-0.968705	+0.048369	$+6 \times 10^{-10}$			(0)		+0.055899	
$10^3 \Delta E, \text{ hartrees}$	0	3.554	3.262			3.269		3.600	
$10^3 \lambda$	0	-6.069	-5.935009			-6.2835		(-6.2835)	
a_1	1	0.994790	0.995380					0.994922	
a_2	0	0.097595	0.091648					0.096036	
a_3	0	-0.029479	-0.028614					-0.030110	

TABLE III. Comparison of other physical properties. The constraint is the electronic dipole moment.

Property	Free variation		Direct solution		Perturbation	
	ϕ_1	Method IA	Method IB	Method IIA	Method IIB	
Total force, $e^2 a_0^{-2}$	-0.9687	+0.0367	+0.0407	-0.0960	-0.0378	
$-T/E_{\text{total}}$	0.9319	0.9322	0.9322	0.9320	0.9321	
$Q_{(\text{cm})}$, Buckingham	-5.5486	-7.3396	-7.3447	-7.0842	-7.1770	
$10^6 \chi_{(\text{Li})}^d$, cm^3/mole	-27,281	-25.785	-25.806	-26.025	-25.923	
$10^6 \chi_{(\text{H})}^d$, cm^3/mole	-32.956	-30.195	-30.224	-30.393	-30.381	
$\sigma(\text{Li})^d$, ppm	105.62	105.55	105.54	105.55	105.55	
$\sigma(\text{H})^d$, ppm	33.251	33.909	33.904	33.848	33.863	

TABLE IV. Comparison of other physical properties. The constraint is the total Hellmann-Feynmann force.

Property	Free variation		Direct Solution		Perturbation	
	ϕ_1		Method IA	Method IB	Method IIA	Method IIB
$-T/E_{\text{total}}$	0.9319	0.9313	0.9313	0.9313	0.9312	0.9313
Dipole moment, D	-4.162	-5.378	-5.316	-5.316	-5.396	-5.369
$Q_{(\text{cm})}$, Buckingham	-5.5486	-7.3307	-7.2497	-7.2497	-7.3096	-7.3438
$10^6 \chi_{(\text{Li})}^d$, cm^3/mole	-27.281	-26.956	-27.001	-27.001	-27.022	-26.983
$10^6 \chi_{(\text{H})}^d$, cm^3/mole	-32.956	-31.925	-32.008	-32.008	-31.942	-31.968
$\sigma_{(\text{Li})}^d$, ppm	105.62	105.46	105.47	105.47	105.46	105.46
$\sigma_{(\text{H})}^d$, ppm	33.251	33.618	33.596	33.596	33.613	33.612