Mo/ller–Plesset convergence issues in computational quantum chemistry

Frank H. Stillinger

Citation: The Journal of Chemical Physics **112**, 9711 (2000); doi: 10.1063/1.481608
View online: http://dx.doi.org/10.1063/1.481608
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/112/22?ver=pdfcov
Published by the AIP Publishing

Møller–Plesset convergence issues in computational quantum chemistry

Frank H. Stillinger
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

(Received 27 December 1999; accepted 15 March 2000)

The Hartree–Fock self-consistent-field approximation has provided an invaluable conceptual framework and a standard computational procedure for atomic and molecular quantum theory. Its shortcomings are significant however, and require remediation. Møller–Plesset perturbation theory offers a popular correction strategy: it formally expands eigenfunctions and eigenvalues as power series in a coupling parameter λ that switches the Hamiltonian continuously between the Hartree–Fock form ($\lambda=0$) and the electron-correlating ''physical'' Hamiltonian ($\lambda=1$). Recent high-order Møller–Plesset numerical expansions indicate that the series can either converge or diverge at $\lambda=1$ depending on the chemical system under study. The present paper suggests at least for atoms that series convergence is controlled by the position of a singularity on the negative real λ axis that arises from a collective all-electron dissociation phenomenon. Nonlinear variational calculations for the two-electron-atom ground state illustrate this proposition, and show that series convergence depends strongly on oxidation state (least favorable for anions, better for neutrals, better yet for cations). © *2000 American Institute of Physics.* [S0021-9606(00)30222-7]

I. BACKGROUND

On account of its conceptual simplicity, computational convenience, and adequate accuracy, the Hartree–Fock self-consistent-field approximation has exerted a dominating influence on the evolution of atomic and molecular electronic structure studies. Indeed its single-particle terminology (e.g., orbitals, bands, sigma and pi electrons, Fermi surfaces, holes, etc.) has permanently entered the general scientific vocabulary. Not surprisingly, the Hartree–Fock approximation often serves as the starting point for more ambitious computational procedures that attempt to come to grips with electron correlation phenomena by appending systematic corrections to that starting point. One of these procedures was initiated by Møller and Plesset in 1934,[1] and forms the subject of this paper.

As explained in the following Sec. II, the Møller–Plesset formalism casts the electronic structure problem into the format of Rayleigh–Schrödinger perturbation theory.[2] The self-consistent-field Hartree–Fock Hamiltonian serves as the unperturbed problem, and the electronic wave function and energy are then developed in power series in a perturbation parameter λ whose increase from 0 to 1 continuously transforms and connects the independent-particle description to the fully correlated electronic structure problem.

Modern advances in computing power have enabled numerical studies to carry out Møller–Plesset expansions to high order, at least for some atomic and molecular systems of modest size. Although basis set adequacy always remains a significant concern, one can safely assume that behavior patterns exhibited by published Møller–Plesset series for energy eigenvalues are at least qualitatively correct. These patterns appear to fall into two categories: (a) convergent (examples are BH, CH_2) and (b) divergent with even–odd sign alternation in the high-order series coefficients (observed for Ne, HF, H_2O).[3] In view of the fact that the physical state of interest lies at $\lambda=1$, these cases (a) and (b), respectively, correspond to λ series with radii of convergence greater than, and less than, unity. Specifically, pattern (b) is characteristic of an energy singularity closest to the origin that lies on the negative real λ axis at less than unit distance from the origin. The purpose of the present study is to argue that this singularity stems from a multielectron autoionization phenomenon.

Beside presenting the general formalism required for this analysis, Sec. II also offers an intuitive description of how the eigenvalue problem is expected to evolve along the real λ axis. Section III bolsters this view with results from a set of nonlinear variational calculations for ^1S ground states of two-electron atoms; in particular these results show a two-electron autoionization whose position on the negative λ axis depends strongly on nuclear charge Z. The final Sec. IV discusses the potentially most useful next-stage calculations designed to elucidate Møller–Plesset convergence issues, and how they might be used to increase the productivity of computational quantum chemistry.

II. FORMALISM

In the interest of maximum clarity, the following will concentrate on the case of $2n$ electrons with equal numbers of up and down spins, i.e., a spin singlet state. The corresponding self-consistent-field Hamiltonian, to be denoted by $\mathbf{H}(0)$, consists strictly of a sum of $2n$ identical operators for the $2n$ electrons:

$$\mathbf{H}(0)=\sum_{j=1}^{2n}\left[-(1/2)\boldsymbol{\nabla}_j^2-\sum_{k=1}^{N}Z_k/|\mathbf{r}_j-\mathbf{R}_k|+\mathbf{V}_j^{(\mathrm{scf})}\right]. \tag{2.1}$$

Here k indexes the N nuclei, with charges Z_k and positions $\mathbf{R}_k$. $\mathbf{V}_j^{\mathrm{scf}}$ is the self-consistent-field operator for electron j, including both Coulomb and exchange portions. The ''physi-

cal'' Hamiltonian, $\mathbf{H}(1)$, replaces the sum of self-consistent-field operators with the pairwise sum of all $2n^2-n$ electron–electron Coulomb repulsions. More generally, let $\mathbf{H}(\lambda)$ be the Hermitian operator that linearly interpolates between these two cases:

$$\mathbf{H}(\lambda)=\sum_{j=1}^{2n}\left[-(1/2)\boldsymbol{\nabla}_j^2-\sum_{k=1}^{N}Z_k/|\mathbf{r}_j-\mathbf{R}_k| +(1-\lambda)\mathbf{V}_j^{(\mathrm{scf})}+\lambda\sum_{j<1}^{2n}1/|\mathbf{r}_j-\mathbf{r}_l|\right]. \tag{2.2}$$

For each electronic state of interest, two λ-dependent energy functions need to be distinguished. The first is the corresponding eigenvalue of $\mathbf{H}(\lambda)$, to be denoted by $E(\lambda)$, associated with normalized wave function $\psi(1,...,2n,\lambda)$:

$$\mathbf{H}(\lambda)\psi(1,...,2n,\lambda)=E(\lambda)\psi(1,..,2n,\lambda). \tag{2.3}$$

The second, $W(\lambda)$, is the expectation value of the physical Hamiltonian $\mathbf{H}(1)$ in the state described by $\psi(1,...,2n,\lambda)$:

$$W(\lambda)=\langle\psi(1,...,2n,\lambda)|\mathbf{H}(1)|\psi(1,...,2n,\lambda)\rangle. \tag{2.4}$$

It is this latter quantity whose power series represents the Møller–Plesset expansion:

$$W(\lambda)=W_0+W_1\lambda+W_2\lambda^2+\cdots. \tag{2.5}$$

In general $E(\lambda)$ and $W(\lambda)$ are not equal. The obvious exception occurs at the physical value of the coupling constant,

$$E(1)=W(1). \tag{2.6}$$

The Rayleigh–Ritz variational principle[4] requires for the ground electronic state (or indeed for the lowest-energy state of any given symmetry) that $W(\lambda)$ must pass through its absolute minimum at $\lambda=1$, an attribute not shared by $E(\lambda)$.

It is traditional to express the wave function at $\lambda=0$ as a $2n\times 2n$ Slater determinant whose elements are space–spin orbitals. However, this is not necessary in view of the fact that Hamiltonian $\mathbf{H}(\lambda)$ is spin independent for all λ. Instead, we can confine attention to any one spin–space component, say that for electrons 1,...,n with spins down and $n+1$,...,$2n$ with spins up, and consider just the position-space dependence of $\psi(\lambda=0)$. Let $\varphi_1(\mathbf{r})\cdots\varphi_n(\mathbf{r})$ be an appropriate orthonormal set of position-space orbitals. Then we can set

$$\psi(\mathbf{r}_1\cdots\mathbf{r}_{2n},\lambda=0)=D(\mathbf{r}_1\cdots\mathbf{r}_n)D(\mathbf{r}_{n+1}\cdots\mathbf{r}_{2n}), \tag{2.7}$$

where D is an $n\times n$ determinant

$$D(\mathbf{r}_1\cdots\mathbf{r}_n)=(n!)^{-1/2}\det[\varphi_i(\mathbf{r}_j)]. \tag{2.8}$$

Self-consistency requires that the orbitals φ_i which compose eigenfunction $\psi(\lambda=0)$ both determine the operators $\mathbf{V}_j^{(\mathrm{scf})}$ and also minimize $W(0)$. Each of the $\mathbf{V}_j^{(\mathrm{scf})}$ resolves into Coulomb (c) and exchange (e) portions,

$$\mathbf{V}_j^{(\mathrm{scf})}=\mathbf{V}_j^{(c)}-\mathbf{V}_f^{(e)}. \tag{2.9}$$

The first of these is just an $\mathbf{r}$-space-function multiplier,

$$\mathbf{V}_j^{(c)}=2\sum_{k=1}^{n}\int d\mathbf{s}\,\varphi_k^2(\mathbf{s})/|\mathbf{r}_j-\mathbf{s}|, \tag{2.10}$$

while the second is an integral operator with the property

$$\mathbf{V}_j^{(e)}\cdot f(\mathbf{r}_j)=\sum_{k=1}^{n}\varphi_k(\mathbf{r}_j)\int d\mathbf{s}\,\varphi_k(\mathbf{s})f(\mathbf{s})/|\mathbf{r}_j-\mathbf{s}|. \tag{2.11}$$

The effect of the self-consistent-field essentially is to provide a static negative charge cloud that is spatially distributed according to the extension of the orbitals comprised in $\psi(\lambda=0)$. The exchange operators reduce the magnitude of the corresponding repulsion somewhat, but only to a partial extent. As λ increases from 0 to 1 in $\mathbf{H}(\lambda)$, Eq. (2.2), this diffuse repulsion continuously switches off while being replaced by explicit electron-pair repulsions. Formally extending λ to even larger positive values greater than 1 causes the self-consistent field to convert to a diffuse attraction surrounding the nuclei, while electron pairs become even more repulsive. For very large positive λ we can expect the $2n$ electrons to concentrate around configurations that represent a compromise between these competing strong attractive and repulsive interactions.

As λ moves from the origin along the negative real axis, interaction roles are reversed in comparison with the $\lambda>1$ regime. The self-consistent field becomes ever more repulsive, overcoming the direct nuclear attractions, thereby diminishing the capacity of the electrons to remain bound in the neighborhood of the nuclei. But now the explicit electron–electron pair interactions have become attractive. As a result the $2n$ electrons in isolation from nuclei have the capacity to form their own bound state whose energy would have the form ($\lambda<0$)

$$-A(n,n)\lambda^2, \tag{2.12}$$

where A is a suitable positive constant. This last expression (2.12) locates an autoionization threshold at which the $2n$ electrons spontaneously leave the neighborhood of the nuclei, together as a bound composite particle, tunneling through a repulsive barrier due to the magnified self-consistent field. The negative λ value at which this occurs can be identified by equating eigenvalue $E(\lambda)$ to quadratic expression (2.12). Just as an analogous autoionization threshold for the Z^{-1} expansion of the two-electron-atom ground state creates a wave function and energy singularity in that context,[5,6] so too can we expect the same for $\psi(\lambda)$, $E(\lambda)$, and $W(\lambda)$. Hence we propose that this collective autoionization phenomenon determines the Møller–Plesset convergence radius.

III. SIMPLE ILLUSTRATIVE EXAMPLE

In order to provide support for the qualitative ideas expressed at the end of Sec. II, we now set up and carry out a simple nonlinear variational calculation. While it is desirable eventually to use a more sophisticated and precise calculation, the following example will suffice for present purposes. Specifically we consider the general two-electron atom (nuclear charge Z) in its singlet ground state, for which the spatial wave function is symmetric under electron interchange. Our task is to estimate $E(\lambda)$ and $W(\lambda)$, and for this purpose we introduce the following two-parameter correlated variational wave function (the nucleus is at the origin):

$$\psi_\nu(\mathbf{r}_1,\mathbf{r}_2,\lambda)=C(\alpha,\beta)\exp[-\alpha(r_1+r_2)-\beta r_{12}]. \quad (3.1)$$

The normalizing constant has the value

$$C(\alpha,\beta)=\left[\frac{8\alpha^3(\alpha+\beta)^5}{\pi^2(8\alpha^2+5\alpha\beta+\beta^2)}\right]^{1/2}. \quad (3.2)$$

Parameters $\alpha(\lambda)$ and $\beta(\lambda)$ are to be determined by minimizing

$$\langle\psi_\nu|\mathbf{H}(\lambda)|\psi_\nu\rangle \geqslant E(\lambda) \quad (3.3)$$

at each λ.

Negative real λ permits formation of a bound ''dielectron.'' In its free state this composite particle has wave function ($\lambda<0$)

$$(|\lambda|^3/8\pi)^{1/2}\exp(-|\lambda|r_{12}/2) \quad (3.4)$$

and binding energy

$$-|\lambda|^2/4. \quad (3.5)$$

Accurate numerical solutions are available for the Hartree–Fock approximation to the two-electron ground state.[7] In principle they could be used to construct the single-particle operators $\mathbf{V}_j^{(c)}$ and $\mathbf{V}_j^{(e)}$. However, that would be ''numerical overkill'' given the modest objective of our elementary variational strategy. Instead we exploit two simplifying approximations. First we use the best effective-charge, single-exponential $1s$ orbital for each nuclear charge Z:[8]

$$\varphi_0(r)=(\alpha_0^3/\pi)^{1/2}\exp(-\alpha_0 r),$$
$$\alpha_0=Z-5/16. \quad (3.6)$$

In this approximation $\mathbf{V}_j^{(c)}$ is a simple closed-form $\mathbf{r}$-space potential

$$\mathbf{V}_j^{(c)}=(2/r_j)[1-\exp(-2\alpha_0 r_j)]-2\alpha_0\exp(-2\alpha_0 r_j). \quad (3.7)$$

The second simplifying assumption involves the following modification of the exchange operator:

$$\mathbf{V}_j^{(e)}\cdot f(r_j)=\varphi_0(r_j)\int d\mathbf{s}\,\varphi(s)f(\mathbf{s})/|\mathbf{r}_j-\mathbf{s}|$$
$$\rightarrow f(\mathbf{r}_j)\int d\mathbf{s}\,\varphi_0(s)\varphi_0(s)/|\mathbf{r}_j-\mathbf{s}|$$
$$\equiv(1/2)\mathbf{V}_j^{(c)}\cdot f(\mathbf{r}_j). \quad (3.8)$$

Note that this is exact when $f=\varphi_0$. Thus we assume

$$\mathbf{V}_j^{(c)}-\mathbf{V}_j^{(e)}\cong(1/2)\mathbf{V}_j^{(c)} \quad (3.9)$$

for illustrative purposes; this is equivalent to the restricted Hartree approximation.[9]

Subject to these simplifications, the Appendix contains the matrix elements needed for the $E(\lambda)$ variational minimization, each of which has a rational algebraic form. These have been used to obtain $\alpha(\lambda)$ and $\beta(\lambda)$ numerically along the real λ axis. Note that the $\mathbf{H}(1)$ matrix element has a relatively simple form:[10]

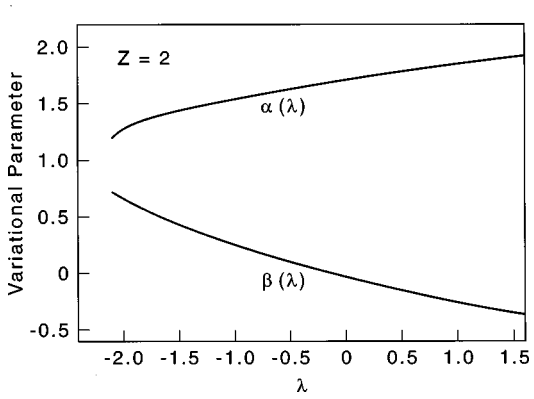

FIG. 1. Coupling constant (λ) dependence of variational parameters α and β in trial wave function ψ_ν, Eq. (3.1), for the helium atom ground state ($Z=2$).

$$\langle\psi_\nu|\mathbf{H}(1)|\psi_\nu\rangle=[(\alpha+\beta)/(8\alpha^2+5\alpha\beta+\beta^2)]$$
$$\times[8\alpha^3+7\alpha^2\beta+4\alpha\beta^2+\beta^3+5\alpha^2$$
$$+4\alpha\beta+\beta^2-4Z\alpha(4\alpha+\beta)]. \quad (3.10)$$

By substituting the $\alpha(\lambda)$ and $\beta(\lambda)$ variational results into this expression one obtains the corresponding $W(\lambda)$.

IV. NUMERICAL RESULTS

Figure 1 shows $\alpha(\lambda)$ and $\beta(\lambda)$ computed for the neutral helium atom, $Z=2$. The corresponding energy functions $E(\lambda)$ and $W(\lambda)$ appear in Fig. 2, which also contains the free dielectron binding energy curve, Eq. (3.5). Numerically it appears possible to locate normalizable wave functions in this approximation for

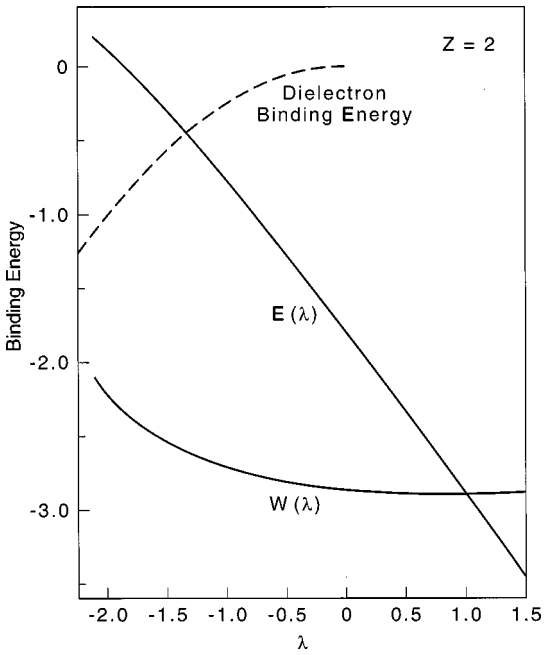

FIG. 2. Variationally determined energy curves $E(\lambda)$ and $W(\lambda)$ for the helium atom ground state. The free dielectron binding energy $-\lambda^2/4$ for $\lambda<0$ has been included to locate the dielectron ionization singularity, its intersection with $E(\lambda)$.

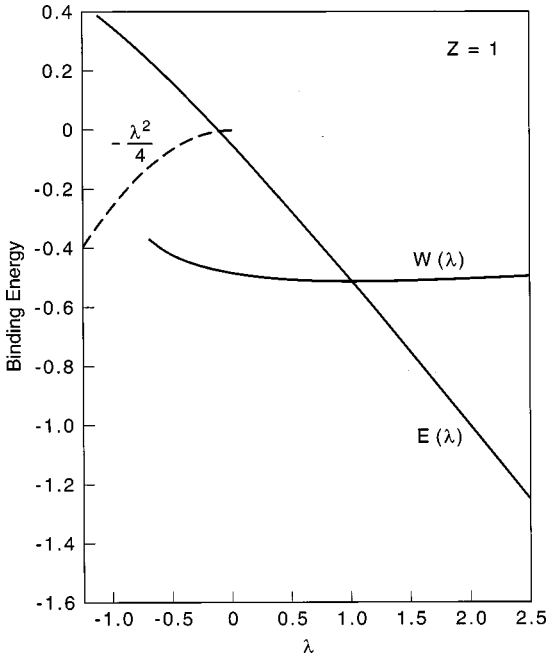

FIG. 3. Variationally determined energy curves $E(\lambda)$ and $W(\lambda)$ for the hydride anion ground state ($Z=1$), along with the free dielectron binding energy $-\lambda^2/4$ for $\lambda<0$.

$$-2.10 \leqslant \lambda \quad (Z=2). \tag{4.1}$$

However, notice in Fig. 2 that $E(\lambda)$ intersects the curve for the dielectron binding energy at the ionization threshold critical value

$$\lambda_c(Z=2) \cong -1.33; \tag{4.2}$$

if the variational wave function had been sufficiently flexible to describe the dielectron ionization explicitly and realistically, no normalizable solution would have been found for $\lambda<\lambda_c$. In view of the fact that the predicted magnitude $|\lambda_c(Z=2)|$ exceeds unity, the Møller–Plesset series for the helium atom ground state is predicted to be absolutely convergent.

Analogous calculations have been performed for $Z=3$ (Li^+) and $Z=4$ (Be^{++}). The critical thresholds for these cases are the following:

$$\begin{aligned} &\lambda_c(Z=3) \cong -2.52, \\ &\lambda_c(Z=4) \cong -3.70. \end{aligned} \tag{4.3}$$

Evidently the ionization singularity moves farther from the origin as Z increases, which would be reflected as a more rapid convergence of the Møller–Plesset series.

The situation is drastically different for $Z=1$ (H^-). The corresponding energy curves appear in Fig. 3. The $E(\lambda)$ and dielectron binding energy functions intersect very close to the origin:

$$\lambda_c(Z=1) \cong -0.08, \tag{4.4}$$

indicative of a strongly divergent Møller–Plesset series.

V. DISCUSSION

Although the variational calculations presented above are admittedly crude and are restricted to two-electron atomic ground states, it is reasonable to suppose that they present qualitatively correct patterns. In particular they lead to the proposition that the Møller–Plesset series for $W(\lambda)$, Eq. (2.4), has a radius of convergence determined by the presence of a singularity on the negative real λ axis. Furthermore this singularity arises from a wave function–distorting phenomenon whereby the electrons are expelled from the region of the nucleus as a free dielectron complex. The variational calculations also indicate that for a fixed number of electrons the singularity position depends strongly on nuclear charge Z: larger Z moves the singularity farther from the origin of the complex λ plane and improves the convergence rate of the Møller–Plesset energy power series.

Obviously it is desirable to strengthen the case by repeating the two-electron ground-state calculations in a more precise manner. One area for improvement involves the $\lambda=0$ Hartree–Fock orbitals, approximated crudely in the present study by a single exponential with effective nuclear charge. A result of this approximation is that the variational calculations at $\lambda=0$ do not quite replicate a product of simple exponentials; instead $\beta(0)$ has a small negative value, and $\alpha(0)$ slightly exceeds effective charge $Z-5/16$, for all Z values investigated. These minor discrepancies would be eliminated upon insertion of a correct Hartree–Fock solution for $\mathbf{H}(0)$.

At the same time it is also desirable to employ a more flexible, and thus potentially more accurate, variational wave function. For the two-electron ground state considered above it would be advantageous to work with linear combinations of Gaussian functions that are appropriately symmetrized:

$$\psi_\nu(\mathbf{r}_1,\mathbf{r}_2) = \sum_i A_i \exp(-a_i r_1^2 - b_i r_2^2 - c_i r_{12}^2). \tag{5.1}$$

The full set of parameters $\{A_i, a_i, b_i, c_i\}$ could in principle be treated as independent variables (subject to ψ_ν normalization), but for practical reasons might be linked into contracted subsets. In any case sufficient flexibility should remain to describe the formation and ionization of the dielectron complex at negative λ.

Using a Gaussian basis would remove another source of imprecision in the present calculations, the replacement (3.8) of exchange operators with their Coulomb operator analogs. The result of $\mathbf{V}^{(e)}$ operating on any Gaussian term can be put into closed form, so approximation (3.8) becomes unnecessary. It is also clear that more than two electrons should be considered, and polyatomic molecules as well, to provide a more comprehensive view of Møller–Plesset convergence issues in quantum chemistry.

Two further matters deserve mention. The first is the nature of the free ''multelectron'' in its ground and excited states. Locating singularity-associated thresholds generally will require determining the energy of these ''self-gravitating'' units composed of given numbers of down-spin and up-spin electrons. The other matter is the precise mathematical characterization of the $E(\lambda)$ and $W(\lambda)$ singularities

themselves; if the analogous $1/Z$-expansion threshold singularity for two electrons is any indication, this mathematical problem will require a deep analysis.[5] Understanding these singularities and how they dominate high-order series coefficients should suggest how best to sum partial series into a closed form, including cases that formally diverge, to provide reliable energy estimates for systems of chemical interest.

ACKNOWLEDGMENTS

The author is pleased to acknowledge helpful discussions with Dr. K. Raghavachari and Dr. C. White.

APPENDIX

Owing to the special choice (3.1) of variational wave function, to the restricted Hartree simplification (3.8)–(3.9), and to the fact that only one nucleus is present, all of the matrix elements needed to evaluate $E(\lambda)$ and $W(\lambda)$ can be expressed as rational algebraic functions of the variational parameters α and β. The kinetic energy has the form

$$\langle K\rangle = \int d\mathbf{r}_1 \int d\mathbf{r}_2\, \psi_\nu(\mathbf{r}_1,\mathbf{r}_2)[-(\boldsymbol{\nabla}_1^2+\boldsymbol{\nabla}_2^2)/2]\psi_\nu(\mathbf{r}_1,\mathbf{r}_2) = \frac{(\alpha+\beta)(8\alpha^3+7\alpha^2\beta+4\alpha\beta^2+\beta^3)}{8\alpha^2+5\alpha\beta+\beta^2}. \tag{A1}$$

Nuclear attractions, proportional to Z, lead to the following:

$$\langle V_N\rangle = \int d\mathbf{r}_1 \int d\mathbf{r}_2\, \psi_\nu^2(-Z/r_1 - Z/r_2) = -\frac{4Z\alpha(\alpha+\beta)(4\alpha+\beta)}{8\alpha^2+5\alpha\beta+\beta^2}. \tag{A2}$$

Direct Coulomb repulsion between the two electrons generates the matrix element:

$$\langle V_e\rangle = \int d\mathbf{r}_1 \int d\mathbf{r}_2\, \psi_\nu^2/r_{12} = \frac{(\alpha+\beta)(5\alpha^2+4\alpha\beta+\beta^2)}{8\alpha^2+5\alpha\beta+\beta^2}. \tag{A3}$$

Equation (3.10) above results from substituting Eqs. (A1), (A2), and (A3) into the expression

$$\langle \mathbf{H}(1)\rangle = \langle K\rangle + \langle V_N\rangle + \langle V_e\rangle. \tag{A4}$$

In order to carry out the variational calculation for arbitrary λ that leads to evaluation of $\alpha(\lambda)$, $\beta(\lambda)$, $E(\lambda)$, and $W(\lambda)$, one also requires a general expression for the Hartree self-consistent-field matrix element. The effective-charge approximation, Eq. (3.6) above, also produces an algebraic matrix element, though rather more complicated than before:

$$\langle V_H\rangle = \int d\mathbf{r}_1 \int d\mathbf{r}_2\, \psi_\nu^2[(\mathbf{V}_1^{(c)}+\mathbf{V}_2^{(c)})/2] = \frac{4\alpha(\alpha+\beta)(4\alpha+\beta)}{8\alpha^2+5\alpha\beta+\beta^2} - \frac{16\alpha^3(\alpha+\beta)^5}{8\alpha^2+5\alpha\beta+\beta^2} \times\{R+\alpha_0 S+4\alpha_0(\alpha+\alpha_0)\alpha\beta T\}, \tag{A5}$$

where

$$R = \frac{\alpha}{[(\alpha+\alpha_0)^2-\beta^2](\alpha^2-\beta^2)^2} + \frac{\beta}{\alpha_0(2\alpha+\alpha_0)(\alpha^2-\beta^2)^2} - \frac{4\alpha^2\beta}{\alpha_0(2\alpha+\alpha_0)(\alpha^2-\beta^2)^3} + \frac{4\alpha\beta^2}{[(\alpha+\alpha_0)^2-\beta^2](\alpha^2-\beta^2)^3} + \frac{2\alpha^2\beta}{\alpha_0^2(2\alpha+\alpha_0)^2(\alpha^2-\beta^2)^2} + \frac{2\alpha\beta^2}{[(\alpha+\alpha_0)^2-\beta^2]^2(\alpha^2-\beta^2)^2} - \frac{2(\alpha+\alpha_0)\alpha\beta}{\alpha_0^2(2\alpha+\alpha_0)^2[(\alpha+\alpha_0)^2-\beta^2]^2}, \tag{A6}$$

$$S = \frac{(\alpha+\alpha_0)\alpha}{[(\alpha+\alpha_0)^2-\beta^2]^2(\alpha^2-\beta^2)^2} + \frac{(\alpha+\alpha_0)\beta}{\alpha_0^2(2\alpha+\alpha_0)^2(\alpha^2-\beta^2)^2} + \frac{\alpha\beta}{\alpha_0^2(2\alpha+\alpha_0)^2[(\alpha+\alpha_0)^2-\beta^2]^2}, \tag{A7}$$

$$T = -\frac{\alpha+\alpha_0}{\alpha_0^3(2\alpha+\alpha_0)^3[(\alpha+\alpha_0)^2-\beta^2]^2} - \frac{\alpha+\alpha_0}{\alpha_0^2(2\alpha+\alpha_0)^2[(\alpha+\alpha_0)^2-\beta^2]^3} + \frac{\alpha}{\alpha_0^3(2\alpha+\alpha_0)^3(\alpha^2-\beta^2)^2} - \frac{\alpha}{\alpha_0^2(2\alpha+\alpha_0)^2(\alpha^2-\beta^2)^3} + \frac{\beta}{[(\alpha+\alpha_0)^2-\beta^2]^3(\alpha^2-\beta^2)^2} + \frac{\beta}{[(\alpha+\alpha_0)^2-\beta^2]^2(\alpha^2-\beta^2)^3}. \tag{A8}$$

The variational eigenvalue estimate $E(\lambda)$ is obtained as follows:

$$E(\lambda) = \min_{(\alpha,\beta)}[\langle K\rangle + \langle V_N\rangle + \lambda\langle V_e\rangle + (1-\lambda)\langle V_H\rangle]. \tag{A9}$$

[1] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).
[2] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1953), Vol. 1, pp. 343–350; K. Gottfried, *Quantum Mechanics, Vol. I Fundamentals* (W. A. Benjamin, New York, 1966), pp. 357–365.
[3] J. Olsen, O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **105**, 5082 (1996).
[4] See R. Courant and D. Hilbert, Ref. 2, pp. 175 and 176.
[5] J. D. Baker, D. E. Freund, R. N. Hill, and J. D. Morgan, Phys. Rev. A **41**, 1247 (1990).
[6] I. A. Ivanov, Phys. Rev. A **52**, 1942 (1995).
[7] C. C. Roothaan, L. M. Sachs, and A. W. Weiss, Rev. Mod. Phys. **32**, 186 (1960); C. A. Stuart, J. F. Toland, and P. G. Williams, Proc. R. Soc. London, Ser. A **388**, 229 (1983).
[8] L. Pauling and E. B. Wilson, *Introduction to Quantum Mechanics* (McGraw-Hill, New York, 1935), pp. 184–185.
[9] D. R. Hartree, *The Calculation of Atomic Structures* (Wiley, New York, 1957).
[10] F. H. Stillinger, J. Chem. Phys. **45**, 3623 (1966), Appendix.