A coordinate free analysis of the modified Hylleraas functionals for establishing bounds on the second-order corrections to the energies of excited states of atoms

1973 J. Phys. A: Math. Nucl. Gen. 6 459

(http://iopscience.iop.org/0301-0015/6/4/009)

J. Phys. A: Math., Nucl. Gen., Vol. 6, April 1973. Printed in Great Britain.

A coordinate free analysis of the modified Hylleraas functionals for establishing bounds on the second-order corrections to the energies of excited states of atoms

C S Sharma† and I Rebelo‡

†Department of Mathematics, Birkbeck College, Malet Street, London WC1E 7HX, UK
‡Mathematical Institute, Oxford University, Oxford OX1 3LB, UK

MS received 28 September 1972, in revised form 20 October 1972

Abstract. Upper and lower bounds are established on the second-order energies of excited stationary states of an atom by a coordinate free analysis of suitably modified Hylleraas functionals. The work proves more rigorously and generalizes some of the earlier results of Sharma. A misunderstanding regarding the correct orthogonality condition for the first-order wavefunction of an excited state is corrected.

1. Introduction

Sharma (1967, 1969) has introduced modified Hylleraas functionals whose extrema respectively provide upper and lower bounds to the second-order energies of the excited stationary states of atoms. The purpose of the present paper is to give a coordinate free proof of a generalized version of the earlier results. The generalization removes the requirement of nondegeneracy in the spectrum of the zero-order Hamiltonian. A result stated without proof by Miller (1966) is also formally proved. These proofs are simpler, more elegant and more rigorous than the earlier ones and being coordinate free are automatically valid for any particular choice of coordinates (basis). The method used is that of the calculus on Banach spaces developed by Frechet and others in the 1920s and eloquently described in a recent book by Cartan (1971). Quantum mechanics is formulated in a Hilbert space over the complex field. However, in calculations on atomic energies it is often possible to complete the integration over the angular parts which reduces the problem to one in the Hilbert space of real square integrable functions on $[0, \infty]$. We shall, therefore, restrict our proofs to problems in a Hilbert space over the real field. Complications arise in generalizing these results to a complex Hilbert space. These complications will be discussed and resolved in a later paper.

2. Formalities

All vector spaces are over the real field. The space of functions from a vector space V_1 to another vector space V_2 will be denoted by (V_1, V_2) and the subspace of continuous linear functions from V_1 to V_2 will be denoted by $L(V_1, V_2)$. The space of continuous bilinear functions from $V_1 \times V_1$ to V_2 will be denoted by $BL(V_1 \times V_1, V_2)$. Let $f \in (V_1, V_2)$ and let $M \subset V_1$. The restriction of f to M will be denoted by $f \upharpoonright M$. The restriction

$f \wedge \mathscr{G}$ of a linear map f on a Hilbert space $\mathscr{H}$ to a subspace $\mathscr{G}$ can be extended to $\mathscr{H}$ by

$$x \in \mathscr{G}^{\perp} \Rightarrow x \xrightarrow{f \wedge \mathscr{G}} 0$$

and there are considerable notational advantages in denoting *this* extension also by $f \wedge \mathscr{G}$. Thus if E is the orthogonal projection on $\mathscr{G}$, in our notation $f \wedge \mathscr{G} = f \circ E$.

2.1. *Some elementary definitions*

Definition 1. Let $\mathscr{B}$ be a real Banach space and let $f: \mathscr{B} \to \mathbb{R}$ be a functional. f is said to be differentiable at a point x if there exists an $f'_x \in L(\mathscr{B}, \mathbb{R})$ such that

$$\lim_{\|h\| \to 0} \frac{|f(x+h) - f(x) - f'_x(h)|}{\|h\|} = 0$$

where $\| \ \|$ is a norm on $\mathscr{B}$ (for the definition of a norm see Halmos 1957). The linear map f'_x is called the derivative of f at x. If f is differentiable at every point $x \in \mathscr{B}$ then f is said to be differentiable in $\mathscr{B}$. Note that in this case the derivative $f' \in (\mathscr{B}, L(\mathscr{B}, \mathbb{R}))$ and its value f'_x at a point $x \in \mathscr{B}$ is a linear functional on $\mathscr{B}$.

Definition 2. (The coordinate free definition of the gradient.) Let $\mathscr{H}$ be a real Hilbert space (note that a Hilbert space is necessarily also a Banach space). Let $f: \mathscr{H} \to \mathbb{R}$ be a functional which is differentiable at $x \in \mathscr{H}$. Then $f'_x \in L(\mathscr{H}, \mathbb{R})$ (observe that $L(\mathscr{H}, \mathbb{R})$ is the dual of $\mathscr{H}$). Then from the Riesz representation theorem we know that there exists a unique $y \in \mathscr{H}$ such that

$$f'_x(p) = \langle y, p \rangle \qquad \forall p \in \mathscr{H}$$

where $\langle , \rangle$ is an inner product (for a definition see Halmos 1957) on $\mathscr{H}$ and satisfies $\|p\|^2 = \langle p, p \rangle$. The unique vector y is called the gradient of f at x and is denoted by

$$y = \operatorname{grad} f(x).$$

2.2. *The higher derivatives of a functional and its Taylor expansion*

Let $f: \mathscr{H} \to \mathbb{R}$ be differentiable in $\mathscr{H}$. Then $f' \in (\mathscr{H}, L(\mathscr{H}, \mathbb{R}))$. If f' is also differentiable in $\mathscr{H}$, then its derivative f'', called the second derivative of f, belongs to $(\mathscr{H}, L(\mathscr{H}, L(\mathscr{H}, \mathbb{R})))$. It is shown by Cartan (1971) that $L(\mathscr{H}, L(\mathscr{H}, \mathbb{R}))$ is isomorphic with the space $BL(\mathscr{H} \times \mathscr{H}, \mathbb{R})$. Hence f''_x can be regarded as a bilinear functional on $\mathscr{H}$.

The higher derivatives are defined analogously and the nth derivative at $x \in \mathscr{H}$ can be regarded as an n-linear functional on $\mathscr{H}$.

Let $f: \mathscr{H} \to \mathbb{R}$ be a functional which is $(n+1)$ times differentiable in $\mathscr{H}$, then

$$\left| f(a+p) - f(a) - f'_a(p) - \frac{1}{2!} f''_a(p, p) - \ \ldots \ - \frac{1}{n!} f^{(n)}_a(p, \overset{n \text{ times}}{\ldots}, p) \right| \leqslant \frac{\| f^{(n+1)}_a \| \ \|p\|^{n+1}}{(n+1)!}.$$

This inequality is known as the Taylor formula with the Lagrange remainder. For a proof see Cartan (1971).

2.3. *Stationary points, maxima and minima of a functional*

Let $f: \mathscr{H} \to \mathbb{R}$ be a functional. Suppose f is twice differentiable in $\mathscr{H}$. A point $x \in \mathscr{H}$

such that $f'_x = 0$ is called a *stationary point* of f. It follows from the Taylor formula that if $f''_x(p, p) > 0$; $\forall p \in \mathscr{H}$ then f has a minimum at the stationary point, if $f''_x(p, p) < 0$; $\forall p \in \mathscr{H}$ then f has a maximum at the stationary point and if $f''_x(p, p)$ has both positive and negative values for particular choices of $p \in \mathscr{H}$ then f has a saddle point at the stationary value. If $f''_x = 0$ one has to study higher derivatives to determine the nature of the stationary point. We shall not be concerned with such cases.

2.4. The Hamiltonian H_0 and its spectral resolution

It is supposed that the Hamiltonian H_0 of the system is a self-adjoint endomorphism (a continuous linear function from $\mathscr{H}$ to $\mathscr{H}$) on a real Hilbert space $\mathscr{H}$. It follows from the continuity of H_0 that its spectrum $\Lambda(H_0)$ is bounded. It will be further assumed that the lower part of the spectrum is purely discrete. Let the eigenvalues of the lower part of the spectrum arranged in a monotonically increasing sequence be denoted by $\{\lambda_i\}$ $(i = 0, 1, 2, \ldots)$ and let E be the spectral measure of H_0. Then from the spectral theorem we have

$$H_0 = \int_\Lambda \lambda \, dE_\lambda. \tag{1}$$

It should be noted that E_{λ_i} is the projection operator on the eigenspace $\mathscr{H}_{\lambda_i}$ of H_0 belonging to the eigenvalue λ_i and the dimension of $\mathscr{H}_{\lambda_i}$ is the degeneracy of λ_i.

It should be pointed out that atomic Hamiltonians satisfy the requirements imposed on H_0 except that they are unbounded above so that they are not self-adjoint endomorphisms but nevertheless they are essentially self-adjoint (for a definition see Kato 1966). However, any variational calculation is done in a finite dimensional subspace; if P is the projection on this subspace then PH_0P is necessarily bounded. Furthermore, by truncating the true spectral resolution of the actual Hamiltonian at a sufficiently large value of λ we get a bounded self-adjoint operator on $\mathscr{H}$ which, for all practical purposes, is a good enough approximation to H_0.

The following properties of H_0 will be used in later sections: (i) H_0 commutes with E, (ii) if P is a projection on a subspace which reduces H_0, then H_0 commutes with P, and (iii) if a subspace M is invariant under H_0, then M reduces H_0.

These properties follow from the self-adjointness of H_0 (for proofs see Halmos 1957).

2.5. The second-order energy and the Hylleraas functional

Let a state represented by a normalized vector Φ_i belonging to the eigenspace $\mathscr{H}_{\lambda_i}$ be subjected to a self-adjoint perturbation H_1. The first-order correction $\lambda_i^{(1)}$ to the eigenvalue is given by

$$\lambda_i^{(1)} = \langle \Phi_i, H_1 \Phi_i \rangle. \tag{2}$$

If the eigenvalue λ_i is degenerate, we shall suppose that if Φ'_i is any other eigenvector belonging to λ_i which is orthogonal to Φ_i then it is also orthogonal to $H_1\Phi_i$. We know from the degenerate perturbation theory that states of physical interest satisfy this requirement.

The second-order correction $\lambda_i^{(2)}$ to the eigenvalue is given by

$$\lambda_i^{(2)} = -\langle \Psi, (H_0 - \lambda_i)\Psi \rangle = \langle \Psi, (H_1 - \lambda_i^{(1)})\Phi_i \rangle \tag{3}$$

where Ψ is the first-order correction to the state vector and is a solution of

$$(H_0-\lambda_i)\Psi+(H_1-\lambda_i^{(1)})\Phi_i = 0. \tag{4}$$

It is easy to verify that Ψ has a unique component in $\mathscr{H} \ominus \mathscr{H}_{\lambda_i}$ and an arbitrary component in $\mathscr{H}_{\lambda_i}$; this arbitrariness in Ψ is removed by a normalization condition on the perturbed state vector but is of no significance as far as the calculation of $\lambda_i^{(2)}$ is concerned. If λ_j belongs to the point spectrum of H_0 and $i \neq j$, then since E_{λ_j} and H_0 commute, we have

$$E_{\lambda_j}\Psi = -\frac{E_{\lambda_j}(H_1-\lambda_i^{(1)})\Phi_i}{\lambda_j-\lambda_i}. \tag{5}$$

Let S denote the continuous spectrum of H_0, then $E_S\Psi$ is the unique solution of

$$(H_0-\lambda_i)E_S\Psi = -E_S(H_1-\lambda_i^{(1)})\Phi_i. \tag{6}$$

The Hylleraas functional $h:\mathscr{H} \to \mathbb{R}$ is chosen in such a way that its gradient is twice the left hand side of equation (4) (this is easily achieved by using rules for differentiating some elementary functionals given by Cartan (1971) and Lang (1970)). Thus

$$h(\chi) = \langle\chi, (H_0-\lambda_i)\chi\rangle + 2\langle\chi, (H_1-\lambda_i^{(1)})\Phi_i\rangle. \tag{7}$$

It follows that any solution Ψ of equation (4) is a stationary point of h and the value of h at any such Ψ is $\lambda_i^{(2)}$. The second derivative of h is computed quite easily and its value at $(p, p) \in \mathscr{H} \times \mathscr{H}$ is given by

$$h''_\chi(p, p) = 2\langle p, (H_0-\lambda_i)p\rangle \tag{8}$$

which is positive if $p \in \mathscr{H}_S$ or $\mathscr{H}_{\lambda_j}$ where $j > i$ and is negative if $p \in \mathscr{H}_{\lambda_j}$ where $j < i$. Thus unless λ_i is the lowest eigenvalue Ψ is a saddle point of h (for an alternative proof see Sharma 1967).

In the next section we modify the Hylleraas functional to obtain two functionals such that any solution Ψ of equation (4) is a stationary point of each of the functionals and $\lambda_i^{(2)}$ is the maximum value of one of the functionals and the minimum value of the other.

3. The theorems

Theorem 1. In the notation of the preceding section the following is valid:

$$m(\chi) = h(\chi) - \frac{1}{\lambda_{i+1}-\lambda_i}\|\tfrac{1}{2}\operatorname{grad} h(\chi)\|^2 + \sum_{n<i}(\lambda_i-\lambda_n)\frac{\lambda_{i+1}-\lambda_n}{\lambda_{i+1}-\lambda_i}\left\|E_{\lambda_n}\left(\chi+\frac{H_1\Phi_i}{\lambda_n-\lambda_i}\right)\right\|^2$$
$$\leqslant \lambda_i^{(2)} \leqslant h(\chi) + \sum_{n<i}(\lambda_i-\lambda_n)\left\|E_{\lambda_n}\left(\chi+\frac{H_1\Phi_i}{\lambda_n-\lambda_i}\right)\right\|^2 = M(\chi). \tag{9}$$

Proof. We first consider the functional m. By setting its first derivative equal to zero, we get

$$\left(1-\frac{H_0-\lambda_i}{\lambda_{i+1}-\lambda_i}\right)\{(H_0-\lambda_i)\chi+(H_1-\lambda_i^{(1)})\Phi_i\}$$
$$+\sum_{n<i}(\lambda_i-\lambda_n)\frac{\lambda_{i+1}-\lambda_n}{\lambda_{i+1}-\lambda_i}E_{\lambda_n}\left(\chi+\frac{H_1\Phi_i}{\lambda_n-\lambda_i}\right) = 0. \tag{10}$$

We shall first show that a solution of this equation has a unique component in

$\mathscr{H} \ominus (\oplus_{j=1}^{i+1} \mathscr{H}_{\lambda_j})$ and this unique component is identical with the unique component in $\mathscr{H} \ominus (\oplus_{j=1}^{i+1} \mathscr{H}_{\lambda_j})$ of a solution Ψ of equation (4). Let $j > i+1$, we operate on equation (10) with E_{λ_j} and remembering that E_{λ_j} commutes with H_0 and that E_{λ_j} is a projection on an eigenspace of H_0 belonging to the eigenvalue λ_j, we have

$$\left(1-\frac{\lambda_j-\lambda_i}{\lambda_{i+1}-\lambda_i}\right)\{(\lambda_j-\lambda_i)E_{\lambda_j}\chi+E_{\lambda_j}(H_1-\lambda_i^{(1)})\Phi_i\} = 0. \tag{11}$$

Since

$$1-\frac{\lambda_j-\lambda_i}{\lambda_{i+1}-\lambda_i} \neq 0,$$

this implies that

$$E_{\lambda_j}\chi = \frac{E_{\lambda_j}(H_1-\lambda_i^{(1)})}{\lambda_i-\lambda_j} \tag{12}$$

which is identical with $E_{\lambda_j}\Psi$.

Finally we operate on equation (10) with E_S and get

$$\left(1-\frac{H_0-\lambda_i}{\lambda_{i+1}-\lambda_i}\right)\{(H_0-\lambda_i)E_S\chi+E_S(H_1-\lambda_i^{(1)})\Phi_i\} = 0 \tag{13}$$

or

$$\frac{\lambda_{i+1}-H_0}{\lambda_{i+1}-\lambda_i}\{(H_0-\lambda_i)E_S\chi+E_S(H_1-\lambda_i^{(1)})\Phi_i\} = 0. \tag{14}$$

Since the restriction of $(\lambda_{i+1}-H_0)$ to $\mathscr{H}_S$ is invertible, we can conclude that

$$(H_0-\lambda_i)E_S\chi+E_S(H_1-\lambda_i^{(1)})\Phi_i = 0 \tag{15}$$

which is identical with equation (6). This proves the assertion regarding the uniqueness and the nature of the component of the solution of equation (10) in $\mathscr{H} \ominus (\oplus_{j=1}^{i+1} \mathscr{H}_{\lambda_j})$.

By using the invariance of the subspaces $\mathscr{H}_{\lambda_j}$ and $\mathscr{H}_S$ under H_0 and the Fourier expansion theorem for Hilbert spaces we have

$$m(\chi) = m(E_S\chi)+\sum_j m(E_{\lambda_j}\chi) \tag{16}$$

and

$$\lambda_i^{(2)} = \langle E_S\Psi, H_1\Phi_i\rangle + \sum_j \lambda_{ij}^{(2)} \tag{17}$$

where

$$\lambda_{ij}^{(2)} = \frac{1}{\lambda_i-\lambda_j}\|E_{\lambda_j}H_1\Phi_i\|^2 \qquad \text{if } i \neq j; \qquad \lambda_{ii}^{(2)} = 0 \tag{18}$$

and Ψ is any solution of equation (4). A straightforward evaluation of $m(E_{\lambda_j}\chi)$ for $j < i$ or $j = i+1$ leads to $m(E_{\lambda_j}\chi) = \lambda_{ij}^{(2)}$ which is independent of χ and is a constant. Furthermore it is easy to verify that $m(E_{\lambda_i}\chi) = \lambda_{ii}^{(2)} = 0$. From these facts it follows that a solution of equation (10) has an arbitrary component in $\oplus_{j=1}^{i+1} \mathscr{H}_j$. Some further elementary computations enable us to conclude that the value of m at any solution χ of equation (10) is $\lambda_i^{(2)}$. Finally we show that this value of m is its maximum value. For

this we compute m'' at $(p, p) \in \mathscr{H} \times \mathscr{H}$; we find

$$m''(p,p) = 2\left\langle p, \left(\frac{(\lambda_{i+1} - H_0)(H_0 - \lambda_i)}{\lambda_{i+1} - \lambda_i} + \sum_{n<i} \frac{(\lambda_{i+1} - \lambda_n)(\lambda_i - \lambda_n)}{\lambda_{i+1} - \lambda_i} E_{\lambda_n} \right) p \right\rangle. \tag{19}$$

Using the spectral decomposition of H_0 in the form

$$H_0 = \sum_n \lambda_n E_{\lambda_n} + \int_S \lambda \, \mathrm{d}E_\lambda \tag{20}$$

equation (19) can be written as

$$m''(p,p) = 2\left\langle p, \left(\sum_{n>i} \frac{(\lambda_{i+1} - \lambda_n)(\lambda_n - \lambda_i)}{\lambda_{i+1} - \lambda_i} E_{\lambda_n} + \int_S \frac{(\lambda_{i+1} - \lambda)(\lambda - \lambda_i)}{\lambda_{i+1} - \lambda_i} \mathrm{d}E_\lambda \right) p \right\rangle \tag{21}$$

from which it follows that $m''(p, p) \leqslant 0$ and $m''(p, p) = 0$ if and only if $p \in \mathscr{H}_{\lambda_i} \oplus \mathscr{H}_{\lambda_{i+1}}$, but we have already seen that the value of m at $x \in \mathscr{H}$ does not depend on the component of x in $\mathscr{H}_{\lambda_i} \oplus \mathscr{H}_{\lambda_{i+1}}$.

By setting the first derivative of M equal to zero, one gets an equation whose solution has a unique component in $\mathscr{H} \ominus (\oplus_{n=1}^{i} \mathscr{H}_{\lambda_n})$ and this unique component is identical with the unique component in this subspace of a solution Ψ of equation (4). The value of M is constant in $\oplus_{n=1}^{i} \mathscr{H}_{\lambda_n}$ and at any stationary point of M its value is $\lambda_i^{(2)}$. By examining the second derivative one is able to conclude that this value is a minimum. The details of the analysis are exactly similar to those for m and are therefore omitted.

Theorem 2. Let $\mathscr{G}$ be a subspace of $\mathscr{H}$ invariant under H_0. Let $\mathscr{G}'$ be another subspace such that $\mathscr{G} \subset \mathscr{G}'$. Then the restriction of the Hylleraas functional h to $\mathscr{G}'$ can be written as

$$h \wedge \mathscr{G}' = h \wedge \mathscr{G} + h \wedge (\mathscr{G}' \ominus \mathscr{G}). \tag{22}$$

Furthermore if $\mathscr{H}_{\lambda_j} \subset \mathscr{G}$ $(j \neq i)$ and χ is a stationary point of $h \wedge \mathscr{G}'$ then

$$E_{\lambda_j} \chi = E_{\lambda_j} \Psi \tag{23}$$

where Ψ is a solution of equation (4).

Proof. We recall that the projection E on $\mathscr{G}$ commutes with H_0. Let E' be the projection on $\mathscr{G}'$. Since $\mathscr{G} \subset \mathscr{G}'$ it follows that $EE' = E'E = E = E^2$ and $E' = E + E' - E$. The restriction of h to $\mathscr{G}'$ can be written as

$$\begin{aligned} h(E'\chi) &= \langle E'\chi, (H_0 - \lambda_i) E'\chi \rangle + 2\langle E'\chi, (H_1 - \lambda_i^{(1)})\Phi_i \rangle \\ &= \langle E\chi, (H_0 - \lambda_i) E\chi \rangle + 2\langle E\chi, (H_1 - \lambda_i^{(1)})\Phi_i \rangle + \langle (E' - E)\chi, (H_0 - \lambda_i)(E' - E)\chi \rangle \\ &\quad + 2\langle (E' - E)\chi, (H_1 - \lambda_i^{(1)})\Phi_i \rangle + \langle (E' - E)\chi, (H_0 - \lambda_i) E\chi \rangle \\ &\quad + \langle E\chi, (H_0 - \lambda_i)(E' - E)\chi \rangle. \end{aligned} \tag{24}$$

By using the properties of E described in the first two sentences of the proof, we see that the last two terms in the above expression are zero. Hence we can write

$$h(E'\chi) = h(E\chi) + h((E' - E)\chi) \tag{25}$$

or, in other words (cf the last sentence of the first paragraph of § 2)

$$h \wedge \mathscr{G}' = h \wedge \mathscr{G} + h \wedge (\mathscr{G}' \ominus \mathscr{G}). \tag{26}$$

From this equation it follows that

$$(h \wedge \mathscr{G}')'_x = (h \wedge \mathscr{G})'_x + (h \wedge (\mathscr{G}' \ominus \mathscr{G}))'_x \tag{27}$$

where the subscript denotes the point at which the value of the function is being considered. $L(\mathscr{G}', \mathbb{R})$ has a unique decomposition as $L(\mathscr{G}, \mathbb{R}) \oplus L((\mathscr{G}' \ominus \mathscr{G}), \mathbb{R})$ and $(h \wedge \mathscr{G}')'_x = 0$ implies that

$$(h \wedge \mathscr{G})'_x = 0 = (h \wedge (\mathscr{G}' \ominus \mathscr{G}))'_x. \tag{28}$$

Hence

$$\operatorname{grad}(h \wedge \mathscr{G})(x) = \operatorname{grad}(h \wedge (\mathscr{G}' \ominus \mathscr{G}))(x) = 0. \tag{29}$$

From this the last part of the assertion follows by observing that $\operatorname{grad}(h \wedge \mathscr{G})$ remains unaltered if $\mathscr{G}' = \mathscr{H}$.

Corollary. Let $\mathscr{G} = \oplus_{n=1}^{i-1} \mathscr{H}_{\lambda_n} \subset \mathscr{G}'$, then the stationary value of $h \wedge \mathscr{G}'$ is an upper bound to $\lambda_i^{(2)}$.

Proof. We have seen that h has a maximum in $\mathscr{G}$ and a minimum in $\mathscr{G}^\perp$ and the value of h at any point x is equal to the sum of the values of h at the components of x in $\mathscr{G}$ and $\mathscr{G}^\perp$ respectively. From the preceding theorem it follows that under the condition in the hypothesis, at a stationary point of $h \wedge \mathscr{G}'$ its component in $\mathscr{G}$ is the maximum value of h in $\mathscr{G}$. On the other hand $\mathscr{G}' \ominus \mathscr{G}$ is a subspace of $\mathscr{G}^\perp$, hence the stationary value in this subspace is either greater than or equal to the minimum value of h in $\mathscr{G}^\perp$. Thus the stationary value of h is an upper bound to $\lambda_i^{(2)}$. This result was first stated without proof by Miller (1966).

We conclude this section with the observation that all results concerning the Hylleraas functional proved by Sinanoglu (1961), Miller (1966) and Sharma (1967, 1969) follow as immediate and trivial corollaries of these two theorems.

4. The orthogonality condition for a first-order wavefunction for an excited state

Sharma (1968) had tried to remove a longstanding misunderstanding regarding the orthogonality condition which must be imposed on the first-order wavefunction of an excited state of an atom in order that the second-order energy calculated with its help may be an upper bound to the true value. It should be observed that the energy corresponding to an approximate normalized wavefunction is just the expectation value of the Hamiltonian in the state represented by the approximate wavefunction. If Φ_i is a normalized wavefunction with an expansion of the form

$$\Phi_i = \Phi_i^{(0)} + \frac{1}{Z}\Phi_i^{(1)}\mathrm{A} + \frac{1}{Z^2}\Phi_i^{(2)}\mathrm{A} + \ldots \tag{30}$$

where Z^{-1} is a perturbation parameter (ie the Hamiltonian is $H_0 + Z^{-1}H_1$), $\Phi_i^{(0)}$ is an exact eigenfunction of H_0 belonging to the value λ_i and the letter A after $\Phi_i^{(1)}$ and $\Phi_i^{(2)}$ denotes the fact that these are approximate, then the corresponding second-order energy is

$$\lambda_i^{(2)}\mathrm{A} = \langle \Phi_i^{(1)}\mathrm{A}, H_0\Phi_i^{(1)}\mathrm{A}\rangle + 2\langle \Phi_i^{(1)}\mathrm{A}, H_1\Phi_i^{(0)}\rangle + 2\langle \Phi_i^{(2)}\mathrm{A}, H_0\Phi_i^{(0)}\rangle = h(\Phi_i^{(1)}\mathrm{A}) \tag{31}$$

where we have used the fact that $\Phi_i^{(0)}$ is an exact eigenfunction of H_0 and that the

normalization of Φ_i implies that

$$\langle \Phi_i^{(1)}\mathrm{A}, \Phi_i^{(0)} \rangle = 0 \tag{32}$$

and

$$2\langle \Phi_i^{(2)}\mathrm{A}, \Phi_i^{(0)} \rangle = -\langle \Phi_i^{(1)}\mathrm{A}, \Phi_i^{(1)}\mathrm{A} \rangle. \tag{33}$$

Thus, whether or not the approximate first-order wavefunction is determined through a variational calculation, the value of the corresponding second-order energy is just the value of the Hylleraas functional at the approximate first-order wavefunction. We have already seen that if $\mathscr{G}$ is the subspace spanned by the eigenfunctions of H_0 belonging to eigenvalues less than λ_i, then h has a maximum in $\mathscr{G}$ and a minimum in $\mathscr{G}^{\perp}$. $\mathscr{G}$ is finite dimensional and it is easy to calculate this maximum, but the minimum value in $\mathscr{G}^{\perp}$ is not known. Thus the only way to ensure that the approximate first-order wavefunction gives an upper bound to the second-order energy is to make sure that the component of $\Phi_i^{(1)}\mathrm{A}$ in $\mathscr{G}$ is the point where h has its maximum in $\mathscr{G}$ and this happens when the component of $\Phi_i^{(1)}\mathrm{A}$ in $\mathscr{G}$ is identical with the component in $\mathscr{G}$ of the exact first-order wavefunction. From equation (5) we see that if $\Phi_j^{(0)}$ is any eigenvector of H_0 belonging to an eigenvalue λ_j less than λ_i, then $\Phi_i^{(1)}\mathrm{A}$ must satisfy

$$\langle \Phi_i^{(1)}\mathrm{A}, \Phi_j^{(0)} \rangle = \frac{\langle \Phi_j^{(0)}, (H_1 - \lambda_i^{(1)})\Phi_i^{(0)} \rangle}{\lambda_i - \lambda_j} = \frac{\langle \Phi_j^{(0)}, H_1 \Phi_i^{(0)} \rangle}{\lambda_i - \lambda_j}. \tag{34}$$

This is the orthogonality condition of Sharma (1968).

Midtdal *et al* (1969) suggest that $\mathscr{G}$ should always be a subspace of the trial subspace and the correct orthogonality condition should be simply

$$\langle \Phi_i^{(1)}\mathrm{A}, \Phi_j^{(0)} \rangle + \langle \Phi_i^{(0)}, \Phi_j^{(1)}\mathrm{A} \rangle = 0. \tag{35}$$

They claim to have established the equivalence of the orthogonality conditions (34) and (35). This claim is both incorrect and misleading. If $\mathscr{G}$ is a subspace of the trial space, then it follows from the corollary to theorem 2 that the corresponding $\lambda_i^{(2)}\mathrm{A}$ is an upper bound to the true value and the condition (34) is satisfied. It is also true that if $\Phi_j^{(1)}$ is determined with the help of a variational calculation involving the Hylleraas functional in the same trial space, then equation (35) is also satisfied. This, however, does not make the two conditions equivalent. We wish to make several points on this issue. Firstly, condition (34) is applicable irrespective of the method used in finding the approximate first-order wavefunction and there are methods, even variational ones, other than the method for which condition (35) is applicable, the Hartree–Fock method being a good example. The Hartree–Fock first-order wavefunction for the $2s^2\ {}^1\mathrm{S}$ state of helium satisfies condition (35) with respect to the ground state if the approximate first-order wavefunction for the ground state is taken to be the first-order Hartree–Fock wavefunction, but it does not satisfy condition (34). The projection operator usually used in calculating the energies of autoionizing states of atoms violates the condition (34) with respect to infinitely many lower lying states but can be made to satisfy condition (35) with a suitable choice of $\Phi_j^{(1)}\mathrm{A}$ for each lower lying state. Secondly, it is completely pointless to include $\mathscr{G}$ in the trial space, this involves evaluating the already known quantities given by equation (34) by a variational calculation and therefore is clearly wasteful. Thirdly, as far as the orthogonality condition for $\Phi_i^{(1)}\mathrm{A}$ is concerned, the approximate first-order wavefunctions of lower lying states $\Phi_j^{(1)}\mathrm{A}$ ($j < i$) are completely irrelevant. Finally, the choice of the same trial space for the calculation of the first-order wavefunctions for different states is highly unlikely to be a judicious one.

Acknowledgments

One of us (IR) would like to express her gratitude to the International Federation of University Women for the award of the Dorothy Leet Grant and to the British Federation of University Women for their Memorial Award. We wish to express our gratitude to the two referees for help in removing a few inaccuracies from an earlier draft of the paper.

References

Cartan H 1971 *Differential Calculus* (Paris: Hermann)
Halmos P 1957 *Introduction to Hilbert Space and the Theory of Spectral Multiplicity* (New York: Chelsea)
Kato T 1966 *Perturbation Theory for Linear Operators* (Berlin: Springer-Verlag)
Lang S 1970 *Analysis* II (New York: Addison-Wesley)
Midtdal J, Lyslo G and Aashamar K 1969 *Phys. Norveg.* **3** 163–78
Miller W H 1966 *J. chem. Phys.* **44** 2198–9
Sinanoglu O 1961 *Phys. Rev.* **122** 491–2
Sharma C S 1967 *Proc. Phys. Soc.* **91** 50–3
—— 1968 *Can. J. Phys.* **46** 43–7
—— 1969 *J. Phys.* A: *Gen. Phys.* **2** 140–1