PHYSICAL REVIEW VOLUME 71, NUMBER 9 MAY 1, 1947

Two Applications of the Variational Method to Quantum Mechanics

WALTER KOHN
*Department of Applied Mathematics, University of Toronto, Toronto, Canada**

(Received September 23, 1946)

By using the unperturbed wave function with a variable scale factor as trial-function for a perturbed Schroedinger equation, a lower bound for the second-order perturbation energy of the ground state in terms of the unperturbed and first-order perturbation energy is obtained. A minimum property of the sum of the n lowest proper values is utilized in a new method for obtaining higher proper values and functions.

1. INTRODUCTION

IN most practical problems of quantum mechanics the Schroedinger equation,

$$H\psi=E\psi, \tag{1}$$

cannot be solved explicitly, and one must take recourse to approximation methods.

The present note, which is based on the method of Ritz, contains (1) an estimate of the second-order perturbation energy in terms of the unperturbed and first-order perturbation energy; and (2) a new approach to the evaluation of the proper values and functions of higher energy states.

2. SCALE FACTOR AND SECOND-ORDER PERTURBATION ENERGY

We write the Hamiltonian of a perturbed system as

$$H=T(x)+V(x)+kv(x), \tag{2}$$

where

x	represents the coordinates of the system;
$T(x)=-\Sigma\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$	is the kinetic energy operator, m representing the masses of the particles;
$V(x)$	is the potential energy operator of the unperturbed system; and
$kv(x)$	is the perturbing potential energy operator.

It follows from the definition of $T(x)$ that

$$T(x/\lambda)=\lambda^2T(x), \tag{3}$$

and if we confine ourselves to electrostatic potentials, then

$$V(x/\lambda)=\lambda V(x),\quad v(x/\lambda)=\lambda v(x). \tag{4}$$

The first-order perturbation energy, ϵ_1, of the ground state is obtained by substituting its unperturbed, normalized wave function, ψ_0, into the expression

$$E=\int\psi^*H\psi dx.^1 \tag{5}$$

Thus, if we denote the unperturbed energy by E_0, we have

$$\int\psi_0^*H\psi_0dx=E_0+k\epsilon_1, \tag{6}$$

where

$$\epsilon_1=\int\psi_0^*v(x)\psi_0dx. \tag{7}$$

Now let us take as trial-function the (normalized) function $\lambda^{3n/2}\psi_0(\lambda x)$, where $3n$ is the dimension of the configuration space and λ is to be adjusted to minimize E. Making use of (3) and (4) we obtain

$$E=\int\psi_0^*(\lambda x)[T(x)+V(x)+kv(z)]\psi_0(\lambda x)d(\lambda x)$$
$$=\int\psi_0^*(x)\left[T\left(\frac{x}{\lambda}\right)+V\left(\frac{x}{\lambda}\right)+kv\left(\frac{x}{\lambda}\right)\right]\psi_0(x)dx \tag{8a}$$

or

$$E=\lambda^2\int\psi_0^*(x)T(x)\psi_0(x)dx$$
$$+\lambda\int\psi_0^*(x)V(x)\psi_0(x)dx$$
$$+\lambda k\int\psi_0^*(x)v(x)\psi_0(x)dx. \tag{8b}$$

---

* At present in the Department of Physics, Harvard University, Cambridge, Massachusetts.

[1] The integral to be taken over the configuration space.

The first and second integrals in (8b) represent, respectively, the average kinetic and potential energy of the unperturbed system. According to the Virial theorem in quantum-mechanics,[2] these are equal to $-E_0$ and $2E_0$, respectively. Thus (8b) may be rewritten as

$$E=-E_0\lambda^2+(2E_0+k\epsilon_1)\lambda. \tag{9}$$

Minimizing with respect to λ we obtain

$$\lambda=1+k\epsilon_1/2E_0, \quad E=E_0+k\epsilon_1+k^2\epsilon_2', \tag{10}$$

where

$$\epsilon_2'=\epsilon_1^2/4E_0. \tag{11}$$

The term $k\epsilon_1$ is just the first-order perturbation energy. The next term, $k^2\epsilon_2'$, is negative and therefore improves our estimate of E.

Furthermore, ϵ_2' is (in absolute value) a lower bound of the second-order perturbation energy, ϵ_2. For, if we subtract the true perturbed energy,

$$E_p=E_0+k\epsilon_1+k^2\epsilon_2+k^3\epsilon_3+\cdots \tag{12}$$

from E, as given by (10), we find

$$E-E_p=k^2(\epsilon_2'-\epsilon_2)-k^3\epsilon_3-\cdots, \tag{13}$$

and remembering now that E is an upper bound of E_p we obtain, on letting k tend to 0, $\epsilon_2'\geqslant\epsilon_2$. Since ϵ_2' is negative this proves our statement.

In summary we may say that if one chooses as trial-function for a perturbed system the unperturbed wave function of the normal state, but with an adjustable scale-factor, the variational method leads to an energy which agrees with the perturbed energy to the first order and in addition contains a correction in terms of the first order and unperturbed energy (see (11)), which is a lower bound of the second-order perturbation energy.

Example: The Helium 1^1S Term

Hylleraas[3] derived the expression

$$E_p=-2Z^2+1.25Z^2\left(\frac{1}{Z}\right)-0.3149Z^2\left(\frac{1}{Z^2}\right)+\cdots \tag{14}$$

for the normal state of a helium-like atom of nuclear charge Ze.

Here $1/Z$ corresponds to our k so that $E_0=-2Z^2$, $\epsilon_1=1.25Z^2$, and $\epsilon_2=-0.3149Z^2$. Our correction term,

$$\epsilon_2'=(1.25Z^2)^2/(-8Z^2)=-0.1953Z^2, \tag{15}$$

accounts for 62 percent of ϵ_2.

3. HIGHER ENERGY STATES

In applying the Ritz method to a higher energy state only functions orthogonal to all lower proper functions may be admitted for competition. Unfortunately, the task of assuring orthogonality is, in general, prohibitively complicated.

Hylleraas and Undheim[4] have used the fact that the nth root of a secular determinant is an upper bound of the nth proper value, to approximate the 2^1S term of helium. However, their method is in practice quite cumbersome and only slowly convergent.

As an alternative we shall derive an expression for the nth energy state whose properties are analogous to (5): Its minimum value is the nth proper value and the function which leads to this minimum gives us the nth proper function.

Our method is based on the following

Lemma[5]

Let $\psi_1, \psi_2, \cdots\psi_n$ be the n lowest normalized proper functions of (1) so that

$$E_i=\int\psi_i^*H\psi_i dx, \quad i=1, 2, \cdots n. \tag{16}$$

Further, let

$$H_{ii}=\int\varphi_i^*H\varphi_i dx, \quad i=1, 2, \cdots n, \tag{17}$$

where $\varphi_1, \varphi_2, \cdots\varphi_n$ are any normalized orthogonal functions. Then

$$E_1+E_2+\cdots E_n=\min.(H_{11}+H_{22}+\cdots H_{nn}), \tag{18}$$

where degenerate proper values are to be counted multiply. The minimum is attained whenever

---

[2] A. Sommerfeld, *Wave Mechanics* (Methuen and Company, London, 1930) p. 248.

[3] E. A. Hylleraas, Zeits. f. Physik 60, 624 (1930).

[4] E. A. Hylleraas and R. Undheim, Zeits. f. Physik 60, 759 (1930).

[5] Courant-Hilbert, *Meth. d. Math. Phys.* (Verlagsbuchhandlung, Julius Springer, Berlin, 1931), second edition, p. 399. No proof is given there but it can be easily supplied by expanding the $\varphi-s$ in terms of the $\psi-s$ and thence deriving expressions for H_{ii} in terms of the E_i.

$\varphi_1, \varphi_2, \cdots \varphi_n$ arise out of $\psi_1, \psi_2, \cdots \psi_n$ by an orthogonal transformation.

An application to the estimation of higher proper values suggests itself immediately.

Let φ_1 be a normalized approximation to the lowest proper function of (1) and let φ_2 be an arbitrary function. Then the functions φ_1 and

$$\bar{\varphi}_2 = (\varphi_2 - I_1\varphi_1)/(N^2 - I_1^2)^{\frac{1}{2}}, \tag{19}$$

with

$$N^2 = \int \varphi_2^*\varphi_2 dx, \quad I_1 = \int \varphi_1^*\varphi_2 dx \tag{20}$$

are normal and orthogonal. Consequently, by our lemma,

$$E_1 + E_2 = \min\left[\int \varphi_1^* H\varphi_1 dx + \int \bar{\varphi}_2^* H\bar{\varphi}_2 dx\right] \tag{21}$$

which leads to

$$E_2 = \min\left[H_{11} - E_1 + \frac{H_{22} - 2I_1H_{12} + I_1^2H_{11}}{N^2 - I_1^2}\right], \tag{22}$$

where

$$H_{ij} = \int \varphi_i^* H\varphi_j dx. \tag{23}$$

We note that (22) does not involve the correct ψ_1 but only E_1 which can be easily and accurately estimated by the Ritz method.

The $\bar{\varphi}_2$'s for which the minimum is nearly reached are seen to be approximations to ψ_2.

For practical purposes (22) is rather complicated. It simplifies considerably when $\varphi_1 = \psi_1$, becoming

$$E_2 = \min\left[E + \frac{(E - E_1)I_1^2}{N^2 - I_1^2}\right], \text{ where } E = \frac{H_{22}}{N^2}. \tag{24}$$

Strictly speaking we are now obliged to use the exact wave function, ψ_1, for the computation of I_1. But fortunately, if φ_2 is well chosen, the "un-orthogonality integral" I_1 and consequently also the term following E in (24) are small. Therefore, to estimate E_2 great accuracy in the evaluation of I_1 is not necessary and a simple approximation of ψ_1 is quite sufficient.

For the nth energy state one obtains, in an analogous notation,

$$\bar{\varphi}_n = \left(\varphi_n - \sum_{i=1}^{n-1} I_i\varphi_i\right) \Big/ \left(N^2 - \sum_{i=1}^{n-1} I_i^2\right), \tag{25}$$

$$E_n = \min\left[\sum_{i=1}^{n-1}(H_{ii} - E_i) + \left(H_{nn} - 2\sum_{i=1}^{n-1} I_iH_{in} + \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} I_iI_jH_{ij}\right) \Big/ \left(N^2 - \sum_{i=1}^{n-1} I_i^2\right)\right], \tag{26}$$

where $\varphi_1, \varphi_2 \cdots \varphi_{n-1}$ are any orthogonal functions and φ_n is arbitrary. In the special case when $\varphi_1 = \psi_1, \cdots \varphi_{n-1} = \psi_{n-1}$, we have

$$\bar{\varphi}_n = \left(\varphi_n - \sum_{i=1}^{n-1} I_i\psi_i\right) \Big/ \left(N^2 - \sum_{i=1}^{n-1} I_i^2\right), \tag{27}$$

$$E_n = \min\left[E + \left(\sum_{i=1}^{n-1}(E - E_i)I_i^2\right) \Big/ \left(N^2 - \sum_{i=1}^{n-1} I_i^2\right)\right]. \tag{28}$$

Example: The Helium 2^1S Term

To obtain an approximation to the wave function of the 2^1S state it was assumed that the "inner" and "outer" electrons move in coulomb fields corresponding, respectively, to $Z=2$ and $Z=1$. This led to the function

$$\varphi_2 = (\tfrac{1}{2}r_1 - 1)\exp - [\tfrac{1}{2}r_1] \exp - [2r_2] + (\tfrac{1}{2}r_2 - 1)\exp[\tfrac{1}{2}r_2] \exp - [2r_1], \tag{29}$$

where r_1 and r_2 denote distances from the nucleus. For the wave function of the ground state the approximation

$$\psi_1 \approx \frac{(27/16)^3}{\pi} \exp - \left[\frac{27}{16}(r_1 + r_2)\right] \tag{30}$$

was used.

Substitution of (29) and (30) into (24) gave -4.24 Ry as approximation to E_2. This is still considerably higher than the observed -4.29 Ry, but in view of the crudeness of our trial-function the improvement over the first-order perturbation value of -4.07 Ry may be considered satisfactory.

ACKNOWLEDGMENT

The author takes great pleasure in thanking Professor A. F. C. Stevenson for his kind advice and interest during the course of this investigation.