40

The Variational Method

Our final perturbation technique is the variational technique, which is particularly well-suited for finding approximations to ground-state energies and wave functions. We shall apply it, in particular, to find approximations to the ground-state energy and wave functions for the He atom.

The variational method employs the following functional: Let the solutions of an energy eigenvalue problem be part of a function space, part or all of our Hilbert space, and let ψ be a particular function of that space. Then, we define the functional, a function of functions ψ to be varied, through

$$E([\psi]) = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}, \tag{1}$$

where $|\psi\rangle$ is square-integrable, but not necessarily normalized to unity; i.e., $\langle\psi|\psi\rangle$ is finite (nonzero), but not necessarily normalized to unity. Also $E([\psi])$ gives the average energy of the system in the state $|\psi\rangle$. The variational method is based on the following theorem.

Theorem: Any state vector, $|\psi\rangle$, for which the average energy, considered as a functional of the vectors of the full state vector space, is stationary is an eigenvector of the discrete spectrum of H; and the corresponding energy eigenvalue is the stationary value of $E([\psi])$.

If we could actually carry out a variation of ψ, starting with some assumed value of ψ and varying ψ by small steps by rotating ψ by small amounts in the infinite-dimensional Hilbert space, such that we could end by finding the true minimum of the functional, we would attain the ground state eigenvector and eigenvalue exactly. Similarly, a local minimum or maximum would give us excited-state eigenvectors and eigenvalues. In actual practice, of course, we cannot do this

K. T. Hecht, *Quantum Mechanics*

even with modern computers. The way the method is actually used is to restict the variation of ψ, not to the set of all functions of the full space, but to a judiciously chosen subset of functions, perhaps a finite number of eigenfunctions of a closely related exactly soluble problem, which we guess to be the most important subspace of our full space. Alternatively, we might restrict the functions ψ to those of a specific analytical form that we guess to be good candidates. (Our physical intuition comes into play here!) These analytical functions should contain a few undetermined (variation) parameters, which are fixed at the values making the variation $\delta E = 0$. If we denote these functions, Φ, and vary $E([\Phi])$ such that $\delta E([\Phi]) = 0$, we may get a good approximation to the exact energy if $E([\Phi_0])$ and $E([\psi_0)]$ differ by only a small amount.

A Proof of the Variational Theorem

To prove the variational principle theorem, let us put $E([\psi])$ in the form

$$E([\psi])\langle\psi|\psi\rangle = \langle\psi|H|\psi\rangle \tag{2}$$

and make the variation to give

$$\delta E\langle\psi|\psi\rangle + E\langle\delta\psi|\psi\rangle + E\langle\psi|\delta\psi\rangle = \langle\delta\psi|H|\psi\rangle + \langle\psi|H|\delta\psi\rangle, \tag{3}$$

or

$$\langle\psi|\psi\rangle\delta E = \langle\delta\psi|(H - E)|\psi\rangle + \langle\psi|(H - E)|\delta\psi\rangle. \tag{4}$$

Hence, we see: If $|\psi\rangle$ is an exact solution of $(H - E)|\psi\rangle = 0$, and if $\langle\psi|\psi\rangle$ is finite and nonzero, $\delta E = 0$.

Also, If $\delta E = 0$,

$$\langle\delta\psi|(H - E)|\psi\rangle + \langle\psi|(H - E)|\delta\psi\rangle = 0. \tag{5}$$

Now, the variation of $|\psi\rangle$ and the variation of $\langle\psi|$ are not independent, but because our vector space is a complex vector space and the variation of $|\psi\rangle$ is completely arbitrary, we could vary the pure real and pure imaginary parts of $|\psi\rangle$ separately and independently of each other. Equivalently, we could vary first $|\psi\rangle$ to give eq. (5), and then vary $i|\psi\rangle$ to give

$$-i\langle\delta\psi|(H - E)|\psi\rangle + i\langle\psi|(H - E)|\delta\psi\rangle = 0. \tag{6}$$

Eqs. (5) and (6), together, then yield: If $\delta E = 0$,

$$\begin{aligned}\langle\delta\psi|(H - E)|\psi\rangle &= 0\\ \langle\psi|(H - E)|\delta\psi\rangle &= 0\end{aligned} \tag{7}$$

separately, leading to the conclusion $|\psi\rangle$ is an exact eigenvector of H with eigenvalue E.

B Bounds on the Accuracy of the Variational Method

Because, in actual practice, we cannot vary the $|\psi\rangle$ of the full vector space, but will vary instead the $|\Phi\rangle$ of a highly restricted subspace of our vector space, it will be very useful to try to get a bound on the variational values of the energy obtained by this technique. We will show, in the subspace of the $|\Phi\rangle$, the variation of the $|\Phi\rangle$, which gives $\delta E([\Phi]) = 0$, leads to an $E([\Phi])$ such that

$$E([\Phi]) \geq E_0, \tag{8}$$

where E_0 is the exact value of the ground-state energy. To prove this, calculate

$$E([\Phi]) - E_0 = \frac{\langle\Phi|(H - E_0)|\Phi\rangle}{\langle\Phi|\Phi\rangle} = \sum_{\text{all } n} \frac{\langle\Phi|(H - E_0)|n\rangle\langle n|\Phi\rangle}{\langle\Phi|\Phi\rangle}, \tag{9}$$

where we have merely inserted the unit operator, $\sum_{\text{all } n}|n\rangle\langle n|$, into the equation and $|n\rangle$ is the exact eigenvector of H with eigenvalue E_n. The sum in this equation may include an integral over continuous energy values if the spectrum of eigenvalues E contains both a discrete part and a continuum. From the above, we get

$$E([\Phi]) - E_0 = \sum_n (E_n - E_0)\frac{\langle\Phi|n\rangle\langle n|\Phi\rangle}{\langle\Phi|\Phi\rangle} = \sum_n (E_n - E_0)\frac{|\langle n|\Phi\rangle|^2}{\langle\Phi|\Phi\rangle} \geq 0. \tag{10}$$

Thus, we know the sign of the error made in the variational technique, although the magnitude of the error is difficult to estimate, which is one of the drawbacks of the technique. However, we can proceed as follows: If we have found an approximate $E([\Phi_1])$ by varying the parameters in the set of functions Φ_1 to obtain an $E([\Phi_1])$, and if we subsequently take a more sophisticated (and hopefully better) set of trial functions, Φ_2, perhaps with a larger number of variational parameters, then if $E([\Phi_2]) < E([\Phi_1])$, we know $E([\Phi_2])$ is a better approximation to the ground-state energy. In principle, this process could be continued to further improve our approximation. The absolute value of the error in the final approximation, however, may still be quite uncertain.

C An Example: The Ground-State Energy of the He Atom

The first step of the calculation involves a good choice of functions, Φ. Here, the ingenuity of the calculator comes to the fore. We want a family of Φ, simple enough for ease of calculation, yet complicated enough to get an accurate result. In our perturbation calculation for the He atom, our zeroth-order wave function, a simple product of hydrogenic $1s$ wave functions, took no account of the presence of the "second" electron. This electron partially shields the nucleus with charge Z from the "first" electron, so this one effectively sees a charge λZ rather than the full charge Z. Here, the shielding factor λ could be introduced as a very simple variational parameter. Thus, the zeroth-order single-particle wave functions,

$\psi^{(0)}_{n=1,l=0}$ of perturbation theory could be replaced by single-particle functions

$$\phi(\lambda) = R(r_i)Y_{00}(\theta_i, \phi_i) = \left(\frac{\lambda Z}{a_0}\right)^{\frac{3}{2}} 2e^{-\lambda Z r_{\text{phys.}i}/a_0} \frac{1}{\sqrt{4\pi}}. \tag{11}$$

If we introduce, the dimensionless r_i, as in Chapter 38, viz. $r_{\text{phys.}i} = \frac{a_0}{Z} r_i$, in terms of these dimensionless r_i, we will build the variational trial function from single-particle functions

$$\phi(\lambda) = \lambda^{\frac{3}{2}} 2e^{-\lambda r_i} \frac{1}{\sqrt{4\pi}}, \tag{12}$$

so the two-particle variational functions are

$$\Phi(\lambda) = 4\lambda^3 e^{-\lambda r_1 - \lambda r_2} \frac{1}{4\pi}. \tag{13}$$

We have chosen the $\Phi(\lambda)$ such that the normalization is $\langle\Phi(\lambda)|\Phi(\lambda)\rangle = 1$ (because it was extremely easy to do so!). Now, using the Hamiltonian, $H^{(0)} + H^{(1)}$ in the form of eq. (14) of Chapter 38, we have

$$\langle\Phi(\lambda)|H|\Phi(\lambda)\rangle$$

$$= 2Z^2\langle\phi(\lambda)|T_{s.p.}|\phi(\lambda)\rangle + 2Z^2\langle\phi(\lambda)|V_{s.p.}|\phi(\lambda)\rangle + Z\langle\Phi(\lambda)|\frac{1}{r_{12}}|\Phi(\lambda)\rangle, \tag{14}$$

where the single-particle expectation values for the single-particle kinetic energy operator, $T_{s.p.}$, and the single-particle potential energy operator, $V_{s.p.}$, follow from

$$\begin{aligned}
\langle\phi(\lambda)|T_{s.p.}|\phi(\lambda)\rangle &= -\frac{1}{2}\langle\phi(\lambda)|\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}|\phi(\lambda)\rangle = \frac{\lambda^2}{2(1)^2}, \\
\langle\phi(\lambda)|V_{s.p.}|\phi(\lambda)\rangle &= -\langle\phi(\lambda)|\frac{1}{r}|\phi(\lambda)\rangle = -\lambda\frac{1}{(1)^2},
\end{aligned} \tag{15}$$

where we have simply made the substitutions, $r = r'/(\lambda)$ in the operators in these equations and have made the transformation $\lambda r = r'$ in the single-particle functions, $\phi(\lambda)$, to relate the integrals to the standard hydrogenic expectation value results, $\langle T_{s.p.}\rangle = +\frac{1}{2n^2}$, and $\langle\frac{1}{r}\rangle = \frac{1}{n^2}$, with $n = 1$. In the same way, we get

$$Z\langle\Phi(\lambda)|\frac{1}{r_{12}}|\Phi(\lambda)\rangle = \frac{5}{8}Z\lambda, \tag{16}$$

where we have again made the substitution $\lambda r_{12} = r'_{12}$ to convert the integral to the form evaluated through eq. (30) of Chapter 38. Thus,

$$E([\Phi(\lambda)]) = \langle\Phi(\lambda)|H|\Phi(\lambda)\rangle = \lambda^2 Z^2 - 2\lambda Z^2 + \frac{5}{8}\lambda Z. \tag{17}$$

With

$$\frac{\partial E(\lambda)}{\partial\lambda} = 2\lambda Z^2 - 2Z^2 + \frac{5}{8}Z = 0, \tag{18}$$

we get

$$\lambda = 1 - \frac{5}{16Z}, \tag{19}$$

or for He, with $Z = 2$, $\lambda Z = Z - \frac{5}{16} = 1.6875$, which is a reasonable value for the shielded nuclear charge. With this value of λ, the variational approximation, $E([\Phi(\lambda)])$, is

$$E([\Phi(\lambda)]) = -Z^2 + \frac{5}{8}Z - \frac{25}{256} = -\left(Z - \frac{5}{16}\right)^2. \tag{20}$$

For, He, with $Z = 2$, this gives $E = -2.8477$. Recall the experimental value was $E = -2.90351$. Considering the extreme simplicity of our trial function, with a single variational parameter, this is a marked improvement over the first-order perturbation theory result of Chapter 38.

To make an improvement over the simple one-parameter variational wave function used here, Hylleraas in (1928) used a six-parameter variational function of the type

$$\Phi = e^{-\lambda(r_1+r_2)}(1 + c_1 u + c_2 t^2 + c_3 s + c_4 s^2 + c_5 u^2)\frac{1}{4\pi}, \tag{21}$$

with $u = r_{12}, s = (r_1 + r_2), t = (r_1 - r_2)$, defined in terms of the dimensionless coordinates, r_i, Φ being a function of the six variational parameters, $\lambda, c_1, c_2, c_3, c_4, c_5$. These trial Φ are now not normalized. Variation of the six parameters in these Φ led to an $E([\Phi]) = -2.90362$. This result violates our theorem, $E([\phi]) \geq E_0(\text{exact})$, $E([\Phi])_{\text{Hylleraas}} < E_{\text{exp.}}$. Recall $E_{\text{exp.}} = -2.90351$. Hylleraas did not, however, include the spin-dependent fine structure terms $H_{\text{f.s.}}$ in his calculation. These terms are of order $\alpha^2 = (1/137)^2$ times the dominant terms in H and would have to be included if we want to compare with experimental results to the order of the 4^{th} decimal place.

D The Ritz Variational Method

So far, we have illustrated the variational method with the use of trial functions built in terms of a few physically motivated parameters. In Ritz's use of the variational technique, the trial function is built in terms of an expansion in a finite number of zeroth-order eigenfunctions of a simpler related problem, with known eigenvalues and eigenfunctions.

$$|\Phi\rangle = \sum_{n=0}^{N} |n\rangle\langle n|\Phi\rangle = \sum_{n=0}^{N} c_n |n\rangle, \tag{22}$$

where the $N + 1$ parameters, c_n are now to be considered as the variational parameters, and it is now easy to choose these such that $\sum_n |c_n|^2 = 1$ and therefore $\langle\Phi|\Phi\rangle = 1$. Now,

$$\delta(\langle\Phi|H|\Phi\rangle) = 0, \tag{23}$$

subject to the subsidiary condition

$$\langle\Phi|\Phi\rangle = 1 = \sum_n |c_n|^2, \tag{24}$$

yields

$$\delta\Big(\sum_{n,m}^{N}\langle\Phi|m\rangle\langle m|H|n\rangle\langle n|\Phi\rangle\Big) = \delta\Big(\sum_{n,m}^{N} c_m^* c_n \langle m|H|n\rangle\Big) = 0, \tag{25}$$

subject to the subsidiary condition, arising from the normalization constraint

$$\delta\Big(\sum_{n,m}^{N} c_m^* \delta_{mn} c_n\Big) = 0. \tag{26}$$

Multiplying this second constraint equation by the Lagrange multiplier, named suggestively $-E$, and adding this to the first (variational) equation, we get

$$\begin{aligned}&\sum_{n,m}^{N}\Big[\delta c_m^*(\langle m|H|n\rangle - E\delta_{mn})c_n + c_m^*(\langle m|H|n\rangle - E\delta_{nm})\delta c_n\Big] = 0\\ &= \sum_{n,m}^{N}\Big[\delta c_m^*(\langle m|H|n\rangle - E\delta_{nm})c_n + \delta c_m\langle m|H|n\rangle^* - E\delta_{mn})c_n^*\Big] = 0,\end{aligned} \tag{27}$$

where we have used the hermitian character of the matrix elements of H and have renamed $n \leftrightarrow m$ in the second term of this equation. Now, the c_n are complex numbers, so we can vary separately the real and imaginary parts of c_n and combine these variations such that, separately, the real and imaginary parts of the above equation can be set equal to zero. This process leads to

$$\begin{aligned}&\sum_{n,m}^{N}\delta c_m^*(\langle m|H|n\rangle - E\delta_{nm})c_n = 0,\\ &\sum_{n,m}^{N}\delta c_m(\langle m|H|n\rangle^* - E\delta_{nm})c_n^* = 0.\end{aligned} \tag{28}$$

Because we can vary the individual c_n separately, we can set all $\delta c_k = 0$, except for one particular δc_m, leading to

$$\sum_{n}^{N}(\langle m|H|n\rangle - E\delta_{nm})c_n = 0. \tag{29}$$

This relation is the usual eigenvalue–eigenvector equation for a finite-dimensional basis. The eigenvalues E_i, with eigenvectors given by the $c_n^{(i)}$ with $i = 0, 1, \ldots, N$ are the variational approximate eigenvalues and eigenvectors. If our original basis was a "good guess," we might expect the lowest few eigenvalues to be good approximations. Improved approximations might then be obtained by expanding the basis from an N-dimensional one to one of slightly higher dimensionality, and if our original guess was indeed a good one, this process should converge to the exact eigenvalues.