Complement E_{XI}

THE VARIATIONAL METHOD

The perturbation theory studied in chapter XI is not the only general approximation method applicable to conservative systems. We shall give a concise description here of another of these methods, which also has numerous applications, especially in atomic and molecular physics, nuclear physics, and solid state physics. First of all, we shall indicate, in §1, the principle of the variational method. Then we shall use the simple example of the one-dimensional harmonic oscillator to bring out its principal features (§2), which we shall briefly discuss in §3. Complements F_{XI} and G_{XI} apply the variational method to simple models which enable us to understand the behavior of electrons in a solid and the chemical bond.

1. Principle of the method

Consider an arbitrary physical system whose Hamiltonian H is time-independent. To simplify the notation, we shall assume that the entire spectrum of H is discrete and non-degenerate:

$$H \mid \varphi_n \rangle = E_n \mid \varphi_n \rangle ; \; n = 0, 1, 2, \ldots \tag{1}$$

Although the Hamiltonian H is known, this is not necessarily the case for its eigenvalues E_n and the corresponding eigenstates $\mid \varphi_n \rangle$. The variational method is, of course, most useful in the cases in which we do not know how to diagonalize H exactly.

a. A PROPERTY OF THE GROUND STATE OF A SYSTEM

Choose an arbitrary ket $\mid \psi \rangle$ of the state space of the system. The mean value of the Hamiltonian H in the state $\mid \psi \rangle$ is such that:

$$\langle H \rangle = \frac{\langle \psi \mid H \mid \psi \rangle}{\langle \psi \mid \psi \rangle} \geqslant E_0 \tag{2}$$

(where E_0 is the smallest eigenvalue of H), equality occuring if and only if $\mid \psi \rangle$ is an eigenvector of H with the eigenvalue E_0.

To prove inequality (2), we expand the ket $| \psi \rangle$ on the basis of eigenstates of H:

$$| \psi \rangle = \sum_n c_n | \varphi_n \rangle \tag{3}$$

We then have:

$$\langle \psi | H | \psi \rangle = \sum_n |c_n|^2 E_n \geqslant E_0 \sum_n |c_n|^2 \tag{4}$$

with, of course:

$$\langle \psi | \psi \rangle = \sum_n |c_n|^2 \tag{5}$$

which proves (2). For inequality (4) to become an equality, it is necessary and sufficient that all the coefficients c_n be zero, with the exception of c_0; $| \psi \rangle$ is then an eigenvector of H with the eigenvalue E_0.

This property is the basis for a method of approximate determination of E_0. We choose (in theory, arbitrarily, but in fact, by using physical criteria) a family of kets $| \psi(\alpha) \rangle$ which depend on a certain number of parameters which we symbolize by α. We calculate the mean value $\langle H \rangle(\alpha)$ of the Hamiltonian H in these states, and we minimize $\langle H \rangle(\alpha)$ with respect to the parameters α. The minimal value so obtained constitutes an approximation of the ground state E_0 of the system. The kets $| \psi(\alpha) \rangle$ are called *trial kets*, and the method itself, the *variational method*.

COMMENT:

The preceding proof can easily be generalized to cases in which the spectrum of H is degenerate or includes a continuous part.

b. GENERALIZATION: THE RITZ THEOREM

We shall show that, more generally, *the mean value of the Hamiltonian H is stationary in the neighborhood of its discrete eigenvalues.*

Consider the mean value of H in the state $| \psi \rangle$:

$$\langle H \rangle = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \tag{6}$$

as a functional of the state vector $| \psi \rangle$, and calculate its increment $\delta \langle H \rangle$ when $| \psi \rangle$ becomes $| \psi \rangle + | \delta\psi \rangle$, where $| \delta\psi \rangle$ is assumed to be infinitely small. To do so, it is useful to write (6) in the form:

$$\langle H \rangle \langle \psi | \psi \rangle = \langle \psi | H | \psi \rangle \tag{7}$$

and to differentiate both sides of this relation:

$$\begin{aligned} \langle \psi | \psi \rangle \delta \langle H \rangle + \langle H \rangle [\langle \psi | \delta\psi \rangle + \langle \delta\psi | \psi \rangle] \\ = \langle \psi | H | \delta\psi \rangle + \langle \delta\psi | H | \psi \rangle \end{aligned} \tag{8}$$

that is, since $\langle H \rangle$ is a number:

$$\langle \psi | \psi \rangle \delta \langle H \rangle = \langle \psi | [H - \langle H \rangle] | \delta\psi \rangle + \langle \delta\psi | [H - \langle H \rangle] | \psi \rangle \tag{9}$$

The mean value $\langle H \rangle$ will be stationary if:

$$\delta \langle H \rangle = 0 \tag{10}$$

which, according to (9), means that:

$$\langle \psi | [H - \langle H \rangle] | \delta\psi \rangle + \langle \delta\psi | [H - \langle H \rangle] | \psi \rangle = 0 \tag{11}$$

We set:

$$| \varphi \rangle = [H - \langle H \rangle] | \psi \rangle \tag{12}$$

Relation (11) can then be written simply:

$$\langle \varphi | \delta\psi \rangle + \langle \delta\psi | \varphi \rangle = 0 \tag{13}$$

This last relation must be satisfied for any infinitesimal ket $| \delta\psi \rangle$. In particular, if we choose:

$$| \delta\psi \rangle = \delta\lambda \, | \varphi \rangle \tag{14}$$

(where $\delta\lambda$ is an infinitely small real number), (13) becomes:

$$2 \langle \varphi | \varphi \rangle \, \delta\lambda = 0 \tag{15}$$

The norm of the ket $| \varphi \rangle$ is therefore zero, and $| \varphi \rangle$ must consequently be zero. With definition (12) taken into account, this means that:

$$H | \psi \rangle = \langle H \rangle | \psi \rangle \tag{16}$$

Consequently, the mean value $\langle H \rangle$ is stationary if and only if the state vector $| \psi \rangle$ to which it corresponds is an eigenvector of H, and the stationary values of $\langle H \rangle$ are the eigenvalues of the Hamiltonian.

The variational method can therefore be generalized and applied to the approximate determination of the eigenvalues of the Hamiltonian H. If the function $\langle H \rangle(\alpha)$ obtained from the trial kets $| \psi(\alpha) \rangle$ has several extrema, they give the approximate values of some of its energies E_n (*cf.* exercise 10 of complement H_{XI}).

c. A SPECIAL CASE WHERE THE TRIAL FUNCTIONS FORM A SUBSPACE

Assume that we choose for the trial kets the set of kets belonging to a vector subspace $\mathscr{F}$ of $\mathscr{E}$. In this case, the variational method reduces to the *resolution of the eigenvalue equation of the Hamiltonian H inside* $\mathscr{F}$, and no longer in all of $\mathscr{E}$.

To see this, we simply apply the argument of § 1-b, limiting it to the kets $| \psi \rangle$ of the subspace $\mathscr{F}$. The maxima and minima of $\langle H \rangle$, characterized by $\delta \langle H \rangle = 0$, are obtained when $| \psi \rangle$ is an eigenvector of H in $\mathscr{F}$. The corresponding eigenvalues constitute the variational method approximation for the true eigenvalues of H in $\mathscr{E}$.

We stress the fact that the restriction of the eigenvalue equation of H to a subspace $\mathscr{F}$ of the state space $\mathscr{E}$ can considerably simplify its solution. However, if $\mathscr{F}$ is badly chosen, it can also yield results which are rather far from the true eigenvalues and eigenvectors of H in $\mathscr{E}$ (*cf.* § 3). The subspace $\mathscr{F}$ must therefore be chosen so as to simplify the problem enough to make it soluble, without too greatly altering the physical reality. In certain cases, it is possible to reduce the study of a complex system to that of a two-level system (*cf.* chap. IV), or at least, to that of a system of a limited number of levels. Another important example of this procedure is the method of the *linear combination of atomic orbitals*, widely used in molecular physics. This method consists essentially (*cf.* complement G_{XI}) of the determination of the wave functions of electrons in a molecule in the form of linear combinations of eigenfunctions associated with the various atoms which constitute the molecule, treated as if they were isolated. It therefore limits the search for the molecular states to a subspace chosen using physical criteria. Similarly, in complement F_{XI}, we shall choose as a trial wave function for an electron in a solid a linear combination of atomic orbitals relative to the various ions which constitute this solid.

COMMENT:

Note that first-order perturbation theory fits into this special case of the variational method: $\mathscr{F}$ is then an eigensubspace of the unperturbed Hamiltonian H_0.

2. Application to a simple example

To illustrate the discussion of § 1 and to give an idea of the validity of the approximations obtained with the help of the variational method, we shall apply this method to the one-dimensional harmonic oscillator, whose eigenvalues and eigenstates we know (*cf.* chap. V). We shall consider the Hamiltonian:

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \tag{17}$$

and we shall solve its eigenvalue equation approximately by variational calculations.

a. EXPONENTIAL TRIAL FUNCTIONS

Since the Hamiltonian (17) is even, it can easily be shown that its ground state is necessarily represented by an even wave function. To determine the characteristics of this ground state, we shall therefore choose even trial functions. We take, for example, the one-parameter family:

$$\psi_\alpha(x) = e^{-\alpha x^2} \quad ; \quad \alpha > 0 \tag{18}$$

The square of the norm of the ket $| \psi_\alpha \rangle$ is equal to:

$$\langle \psi_\alpha | \psi_\alpha \rangle = \int_{-\infty}^{+\infty} dx\, e^{-2\alpha x^2} \tag{19}$$

and we find:

$$\begin{aligned}\langle \psi_\alpha | H | \psi_\alpha \rangle &= \int_{-\infty}^{+\infty} dx\, e^{-\alpha x^2} \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 x^2 \right] e^{-\alpha x^2} \\ &= \left[\frac{\hbar^2}{2m}\alpha + \frac{1}{8} m\omega^2 \frac{1}{\alpha} \right] \int_{-\infty}^{+\infty} dx\, e^{-2\alpha x^2} \end{aligned} \tag{20}$$

so that:

$$\langle H \rangle(\alpha) = \frac{\hbar^2}{2m}\alpha + \frac{1}{8} m\omega^2 \frac{1}{\alpha} \tag{21}$$

The derivative of the function $\langle H \rangle(\alpha)$ goes to zero for:

$$\alpha = \alpha_0 = \frac{1}{2}\frac{m\omega}{\hbar} \tag{22}$$

and we then have:

$$\langle H \rangle(\alpha_0) = \frac{1}{2}\hbar\omega \tag{23}$$

The minimum value of $\langle H \rangle(\alpha)$ is therefore exactly equal to the energy of the ground state of the harmonic oscillator. This result is due to the simplicity of the problem that we are studying: the wave function of the ground state happens to be precisely one of the functions of the trial family (18), the one which corresponds to value (22) of the parameter α. The variational method, in this case, gives the exact solution of the problem (this illustrates the theorem proven in § 1-a).

If we want to calculate (approximately, in theory) the first excited state E_1 of the Hamiltonian (17), we should choose trial functions which are orthogonal to the wave function of the ground state. This follows from the discussion of § 1-a, which shows that $\langle H \rangle$ has a lower bound of E_1, and no longer of E_0, if the coefficient c_0 is zero. We therefore choose the trial family of odd functions:

$$\psi_\alpha(x) = x\, e^{-\alpha x^2} \tag{24}$$

In this case:

$$\langle \psi_\alpha | \psi_\alpha \rangle = \int_{-\infty}^{+\infty} dx\, x^2\, e^{-2\alpha x^2} \tag{25}$$

and:

$$\langle \psi_\alpha | H | \psi_\alpha \rangle = \left[\frac{\hbar^2}{2m} \times 3\alpha + \frac{1}{2} m\omega^2 \times \frac{3}{4\alpha} \right] \int_{-\infty}^{+\infty} dx\, x^2\, e^{-2\alpha x^2} \tag{26}$$

which yields:

$$\langle H \rangle(\alpha) = \frac{3\hbar^2}{2m}\alpha + \frac{3}{8} m\omega^2 \frac{1}{\alpha} \tag{27}$$

This function, for the same value α_0 as above [formula (22)], presents a minimum equal to:

$$\langle H \rangle(\alpha_0) = \frac{3}{2}\hbar\omega \tag{28}$$

Here again, we find exactly the energy E_1 and the associated eigenstate because the trial family includes the correct wave function.

b. RATIONAL WAVE FUNCTIONS

The calculations of §2-a enabled us to familiarize ourselves with the variational method, but they do not really allow us to judge its effectiveness as a method of approximation, since the families chosen always included the exact wave function. Therefore, we shall now choose trial functions of a totally different type, for example*:

$$\psi_a(x) = \frac{1}{x^2 + a} \qquad ; \quad a > 0 \tag{29}$$

A simple calculation then yields:

$$\langle \psi_a | \psi_a \rangle = \int_{-\infty}^{+\infty} \frac{\mathrm{d}x}{(x^2 + a)^2} = \frac{\pi}{2a\sqrt{a}} \tag{30}$$

and, finally:

$$\langle H \rangle(a) = \frac{\hbar^2}{4m}\frac{1}{a} + \frac{1}{2} m\omega^2 a \tag{31}$$

The minimum value of this function is obtained for:

$$a = a_0 = \frac{1}{\sqrt{2}}\frac{\hbar}{m\omega} \tag{32}$$

and is equal to:

$$\langle H \rangle(a_0) = \frac{1}{\sqrt{2}}\hbar\omega \tag{33}$$

This minimum value is therefore equal to $\sqrt{2}$ times the exact ground state energy $\hbar\omega/2$. To measure the error committed, we can calculate the ratio of $\langle H \rangle(a_0) - \hbar\omega/2$ to the energy quantum $\hbar\omega$:

$$\frac{\langle H \rangle(a_0) - \frac{1}{2}\hbar\omega}{\hbar\omega} = \frac{\sqrt{2} - 1}{2} \simeq 20\,\% \tag{34}$$

* Our choice here is dictated by the fact that we want the necessary integrals to be analytically calculable. Of course, in most real cases, one resorts to numerical integration.

3. Discussion

The example of §2-b shows that it is easy to obtain the ground state energy of a system, without significant error, starting with arbitrarily chosen trial kets. This is one of the principal advantages of the variational method. Since the exact eigenvalue is a minimum of the mean value $\langle H \rangle$, it is not surprising that $\langle H \rangle$ does not vary very much near this minimum.

On the other hand, as the same reasoning shows, the "approximate" state can be rather different from the true eigenstate. Thus, in the example of §2-b, the wave function $1/(x^2 + a_0)$ [where a_0 is given by formula (32)] decreases too rapidly for small values of x and much too slowly when x becomes large. Table I gives quantitative support for this qualitative assertion. It gives, for various values of x^2, the values of the exact normalized eigenfunction:

$$\varphi_0(x) = (2\alpha_0/\pi)^{1/4}\, e^{-\alpha_0 x^2}$$

[where α_0 was defined in (22)] and of the approximate normalized eigenfunction:

$$\sqrt{\frac{2}{\pi}}(a_0)^{3/4}\psi_{a_0}(x) = \sqrt{\frac{2}{\pi}}\frac{(a_0)^{3/4}}{x^2 + a_0} = \sqrt{\frac{2}{\pi}}(2\sqrt{2}\alpha_0)^{1/4}\frac{1}{1 + 2\sqrt{2}\alpha_0 x^2} \tag{35}$$

$x\sqrt{\alpha_0}$	$\left(\frac{2}{\pi}\right)^{1/4} e^{-\alpha_0 x^2}$	$\sqrt{\frac{2}{\pi}}\frac{(2\sqrt{2})^{1/4}}{1 + 2\sqrt{2}\alpha_0 x^2}$
0	0.893	1.034
1/2	0.696	0.605
1	0.329	0.270
3/2	0.094	0.140
2	0.016	0.083
5/2	0.002	0.055
3	0.000 1	0.039

TABLE I

It is therefore necessary to be very careful when physical properties other than the energy of the system are calculated using the approximate state obtained from the variational method. The validity of the result obtained varies enormously depending on the physical quantity under consideration. In the particular problem which we are studying here, we find, for example, that the approximate mean value of the operator X^2* is not very different from the exact value:

$$\frac{\langle \psi_{a_0} | X^2 | \psi_{a_0} \rangle}{\langle \psi_{a_0} | \psi_{a_0} \rangle} = \frac{1}{\sqrt{2}}\frac{\hbar}{m\omega} \tag{36}$$

* The mean value of X is automatically zero, as is correct since we have chosen even trial functions.

which is to be compared with $\hbar/2m\omega$. On the other hand, the mean value of X^4 is infinite for the wave function (35), while it is, of course, finite for the real wave function. More generally, table I shows that the approximation will be very poor for all properties which depend strongly on the behavior of the wave function for $x \gtrsim 2/\sqrt{\alpha_0}$.

The drawback we have just mentioned is all the more serious as it is very difficult, if not impossible, to evaluate the error in a variational calculation if we do not know the exact solution of the problem (and, of course, if we use the variational method, it is because we do not know this exact solution).

The variational method is therefore a very flexible approximation method, which can be adapted to very diverse situations and which gives great scope to physical intuition in the choice of trial kets. It gives good values for the energy rather easily, but the approximate state vectors may present certain completely unpredictable erroneous features, and we cannot check these errors. This method is particularly valuable when physical arguments give us an idea of the qualitative or semi-quantitative form of the solutions.

References and suggestions for further reading:

The Hartree-Fock method, often used in physics, is an application of the variational method. See references of chapter XI.

The variational method is of fundamental importance in molecular physics. See references of complement G_{XI}.

For a simple presentation of the use of variational principles in physics, see Feynman II (7.2), chap. 19.