THEORETICAL CHEMISTRY INSTITUTE

THE UNIVERSITY OF WISCONSIN

THE VARIATIONAL METHOD II -- PERTURBATION THEORY

by

Saul T. Epstein

WIS-TCI-338

10 March 1969

MADISON, WISCONSIN

The Variational Method II -- Perturbation Theory[*]

by

Saul T. Epstein

Theoretical Chemistry Institute and Physics Department
University of Wisconsin, Madison, Wisconsin

I WOULD GREATLY APPRECIATE YOUR CRITICISMS, COMMENT, AND SUGGESTIONS CONCERNING THE CONTENT AND ORGANIZATION OF THESE LECTURES.

S.T.E.

---

[*] Research supported by National Aeronautics and Space Administration Grant NGL 50-002-001.

XII. The Variation Principle and Perturbation Theory

There are many connections between the variation method and perturbation theory. In the succeeding sections we will be concerned in the main with using the variational method to approximate solutions of the perturbation equations, and, before that, with the closely related problem of the perturbation analysis of optimal trial functions and energies. However the variation principle can play other roles. For example it can often offer insight into exact perturbation theory results. Thus let ν be the perturbation parameter. Then if $\tilde{\psi}-\psi$ is of order ν^{n+1} where ψ is an eigenfunction (not necessarily of the lowest state) then it is an immediate consequence of the variation principle that $\tilde{E}-E$ is of order ν^{2n+2} and we have the well known result that to calculate the energy through order ν^{2n+1} one needs to know the wave function only through order ν^{n}. This theorem is usually proven by clever rearrangement of the perturbation equations. In Sec. XIV. we will present a generalization of this theorem.

In this section we wish to discuss yet another kind of connection; one which, so to speak, goes from perturbation theory to a variational principle rather than the other way round. The point is simply the following. The mark of a variational principle is that it leads to errors of second order in the quantity to be calculated (we have been concentrating on the energy). This then suggests that if using perturbation theory we can give a formula for calculating a quantity correct through first order, i.e., so that the error is of second order, then from this formula we should be able to infer a variation principle for the quantity in question. (This also suggests that starting from the results of higher order perturbation theory one could infer "super-variation principles". See for example Kikuta, Prog. Theo. Phys. 1210 (1954), 14, 457 (1955). See also Koster and Brooks, J. Math. Phys. 5, 169 (1954).)

Rather than discussing in detail how to carry out this program we will simply illustrate the point by a few examples, others will occur in later sections. (For a general discussion see Pomraning, J. Math. Phys. 8, 149 (1967) and references given there.) First let us consider the energy itself. Then given a zero order wave function $\psi^{(0)}$, a zero order energy $E^{(0)}$, and a zero order Hamiltonian $H^{(0)}$

$$H^{(0)}\psi^{(0)} = E^{(0)}\psi^{(0)}$$

one has the familiar result that through first order the energy for a Hamiltonian H is given by

$$E^{(0)} + \frac{(\psi^{(0)}, (H - H^{(0)})\psi^{(0)})}{(\psi^{(0)}, \psi^{(0)})}$$

which we can write as

$$(\psi^{(0)}, H\psi^{(0)}) / (\psi^{(0)}, \psi^{(0)})$$

In this last form all explicit reference to $H^{(0)}$ has disappeared. Further it is a formula which is correct through first order in the error in $\psi^{(0)}$, whatever $\psi^{(0)}$ may be. Hence we can infer that

$$(\tilde{\psi}, H\tilde{\psi}) / (\tilde{\psi}, \tilde{\psi})$$

will yield a variational approximation to the energy, and we are back where we started from.

To give another example, let us use perturbation theory to derive a variation principle for a Green's function. Using an operator notation one has

$$(H - z)\, G = 1 \qquad \text{(XII-1)}$$

$$(H^{(0)} - z)\, G^{(0)} = 1 \qquad \text{(XII-2)}$$

where z is a real number. We now rewrite (1) as

$$(H^{(0)} - z)\, G = 1 + (H^{(0)} - H)\, G$$

Then multiplying through by $G^{(0)} = (H^{(0)} - z)^{-1}$ we have

$$G = G^{(0)} + G^{(0)}(H^{(0)} - H)\, G$$

or, correct through first order, the Green's function is given by

$$G^{(0)} + G^{(0)}(H^{(0)} - H)\, G^{(0)} \qquad \text{(XII-3)}$$

Following the pattern in the previous case, we now elminate explicit reference to $H^{(0)}$. To do this we use (2) to write (3) as

$$2\, G^{(0)} + G^{(0)}(z - H)\, G^{(0)}$$

whence, since $G^{(0)}$ is arbitrary, we infer that

$$2\tilde{G} + \tilde{G}(z - H)\tilde{G} \qquad \text{(XII-4)}$$

furnishes a stationary approximation to G.

Problem: Prove this. Hint: Write $\tilde{G} = G + \Delta$ and show that this quantity equals G plus terms of order Δ^2 (note that $G(H-z) = 1$).

Finally we give an example from scattering theory. We consider the scattering of a particle of mass M and energy E by a central field and confine attention to a single partial wave. Then there is a well known formula which gives k^{-1} times the tangent of the phase shift correct through first order when we change the potential, namely

$$k^{-1}\tan\delta^{(0)} + 2M\left(u^{(0)}, (H^{(0)} - H)u^{(0)}\right) \qquad \text{(XII-5)}$$

where $k = \sqrt{\frac{2ME}{\hbar^2}}$ and where $H^{(0)}$ and H are radial Hamiltonians for angular momentum ℓ and where $u^{(0)}$ is the zero order radial wave function normalized according to

$$k u^{(0)} \xrightarrow[r\to\infty]{} \sin\left(kr - \frac{\ell\pi}{2}\right) + \tan\delta^{(0)}\cos\left(kr - \ell\pi/2\right) \qquad \text{(XII-6)}$$

We can now eliminate explicit reference to $H^{(0)}$ in (6) by use of $(H^{(0)} - E)u^{(0)} = 0$ and hence infer that

$$k^{-1}\tan\tilde{\delta} + 2M\left(\tilde{u}, (E - H)\tilde{u}\right)$$

will furnish a variational approximation to $k^{-1}\tan\delta$, the $\tilde{u}$ being required to have the asymptotic form

$$k\tilde{u} \xrightarrow[r\to\infty]{} \sin\left(kr - \ell\pi/2\right) + \tan\tilde{\delta}\cos\left(kr - \ell\pi/2\right)$$

This result is known as the Kohn variation principle.

It is of some interest to note that if one sets the first order variation equal to zero the result, after an integration by parts, can be written in a form identical to that which we found for bound states, namely

$$(\delta\tilde{u}, (H - E)\tilde{u}) + ((H - E)\tilde{u}, \delta\tilde{u}) = 0$$

XIII. Perturbation Analysis of the Variational Method

Let us suppose that the Hamiltonian takes the form

$$H = H^{(0)} + \nu H^{(1)} \qquad \text{(XIII-1)}$$

with $H^{(0)}$ and $H^{(1)}$ independent of the real parameter ν, and that we are interested in exhibiting $\hat{\Psi}$ and $\hat{E}$ as power series in ν (we will assume without further comment that this is possible). One way to derive such expansions is simply to carry out the complete calculation and then to expand the results. The other approach, and the one which we wish to explore in this section, is to try to produce the successive terms in the series by an iterative procedure.

> While the second approach is the one which has usually been followed in say calculating the response to external fields, or in the $1/Z$ expansion, in recent years there has been a trend the other way. Thus to calculate the static polarizability, instead of directly calculating the second order energy, one calculates the total energy in a series of weak fields and fits the results to a polynomial. See for example papers by Hurst and collaborators, Phys. Rev. 167, 1 (1968), Cohen and Roothaan, J. Chem. Phys., 43, 534 (1965), and Pople et. al., J. Chem. Phys. 49, 2960 (1968).

Before turning to this however, let us say a few words about the various possibilities for $H^{(0)}$ and $\nu H^{(1)}$ that one encounters in applications. First of all $H^{(0)}$ will usually be either the "accurate" Hamiltonian for an isolated system or some approximation to it. In saying this we qualified the word accurate by quotation marks because it is a relative term. Thus for an isolated atom or a molecule, under many circumstances the fixed nucleus, non-relativistic (no magnetic effects) Hamiltonian, is the standard of accuracy, with corrections adequately taken care of by *first* order perturbation theory. In other circumstances, particularly for heavy elements, this will not be the case and relativistic effects must be treated more accurately.

If $H^{(0)}$ is an accurate Hamiltonian then $\nu H^{(1)}$ will usually represent the effect of an external field, (i.e., the sort of thing we have elsewhere included in μW .) On the other hand if $H^{(0)}$ is approximate then $\nu H^{(1)}$ might represent the effect of an external field but it might also be the difference between the "exact" and approximate Hamiltonians of the ioslated system, i.e., the λV of earlier discussions. Also it might be a quantity for which we would like to calculate an accurate expectation value and which using an interchange theorem, we have introduced as a ficticious perturbation (i.e., again a quantity of the μW type).

For complete generality then we really should replace $\nu H^{(1)}$ by a sum of terms (including terms of different orders), however since the modification in the formalism which this introduces should be quite obvious, at least in low orders, we will forgo complicating the notation at this point.

Let us now return to our problem -- the expansion of $\hat{\psi}$ and $\hat{E}$ in powers of ν. We will confine our attention initially to situations in which the *space of trial functions* $\tilde{\psi}$ *is independent of* ν since this is often the case in practice. If we denote the collection of parameters and/or functions which label the members of the space by $\tilde{\alpha}$, and if we denote a typical trial function by $\phi(\tilde{\alpha})$, then this means that

$$\hat{\psi} = \phi(\hat{\alpha}) \qquad \text{(XIII-2)}$$

will depend on ν only because the $\hat{a}$ depend on ν. (As we have seen this also means that $\hat{\psi}$ will satisfy the Hellmann-Feynman theorem for $\sigma = \nu$ and, generalizing to the case of several perturbations, that interchange theorems will be satisfied.) In accord with our program then we write

$$\hat{\psi} = \hat{\psi}^{(0)} + \nu \hat{\psi}^{(1)} + \nu^2 \hat{\psi}^{(2)} + \cdots \tag{XIII-3}$$

$$\hat{E} = \hat{E}^{(0)} + \nu \hat{E}^{(1)} + \nu^2 \hat{E}^{(2)} + \cdots \tag{XIII-4}$$

Further we will assume that $\hat{a}$ can also be expanded in a power series in ν

$$\hat{a} = \hat{a}^{(0)} + \nu \hat{a}^{(1)} + \nu^2 \hat{a}^{(2)} + \cdots \tag{XIII-5}$$

This is sufficient to yield (3) and (4) but not necessary. However we assume it because it simplifies the discussion. We will comment on more general possibilities at the end of this section.

From our assumption that $\hat{\psi}$ depends on ν only because $\hat{a}$ does, it then follow that

$$\hat{\psi}^{(0)} = \phi(\hat{a}^{(0)}) \tag{XIII-6}$$

$$\hat{\psi}^{(1)} = \left(\frac{\partial \phi}{\partial \tilde{a}}\right)_{\tilde{a} = \hat{a}^{(0)}} \hat{a}^{(1)} \tag{XIII-7}$$

$$\hat{\psi}^{(2)} = \left(\frac{\partial \phi}{\partial \tilde{a}}\right)_{\tilde{a} = \hat{a}^{(0)}} \hat{a}^{(2)} + \frac{1}{2}\left(\frac{\partial^2 \phi}{\partial \tilde{a}^2}\right)_{\tilde{a} = \hat{a}^{(0)}} (\hat{a}^{(1)})^2 \tag{XIII-8}$$

etc.

Now a point of notation. The formulae (6) - (8) are quite explicit (except for our use of the collective symbol $\tilde{a}$. Thus if the $\tilde{a}$ are a set of parameters then the r.h.s. of (7) is short hand for $\sum_i \left(\frac{\partial\phi}{\partial\tilde{\alpha}_i}\right)_{\tilde{\alpha}=\hat{\alpha}^{(0)}} \hat{\alpha}_i^{(1)}$.) However for compactness we will replace $\left(\frac{\partial\phi}{\partial\tilde{\alpha}}\right)_{\tilde{\alpha}=\hat{\alpha}^{(0)}}$, $\left(\frac{\partial^2\psi}{\partial\tilde{\alpha}^2}\right)_{\tilde{\alpha}=\hat{\alpha}^{(0)}}$ etc. by $\partial\hat{\psi}^{(0)}/\partial\hat{\alpha}^{(0)}$, $\partial^2\hat{\psi}^{(0)}/\partial^2\hat{\alpha}^{(0)}$ etc. which we will then further abbreviate by $\hat{\psi}_1^{(0)}$, $\hat{\psi}_2^{(0)}$ etc. Thus we have

$$\hat{\psi} = \hat{\psi}^{(0)} + \nu\,\hat{\psi}_1^{(0)}\,\hat{\alpha}^{(1)} + \nu^2\left[\hat{\psi}_1^{(0)}\,\hat{\alpha}^{(2)} + \tfrac{1}{2}\hat{\psi}_2^{(0)}\left(\hat{\alpha}^{(1)}\right)^2\right] + \cdots \qquad \text{(XIII-9)}$$

This more compact notation should cause no confusion. However, one must keep in mind that some members of the $\hat{\alpha}^{(0)}$ may be zero. (Consider for example the Stark effect for a closed shell atom and let one of the $\tilde{\alpha}$ measure the amount of P state.) Note also that if there is no distinction between $\tilde{\psi}$ and $\tilde{\alpha}$ (c.f. the S-limit in Helium) then $\hat{\psi}_m^{(0)} = \delta_{m,1}$.

Finally we need $\delta\hat{\psi}$. We now remark that we can explore the $\tilde{\alpha}$ in the neighborhood of $\hat{\alpha}$ simply by making arbitrary changes in $\hat{\alpha}^{(0)}$ alone. Thus it will be sufficient to consider

$$\delta\hat{\psi} = \frac{\partial\hat{\psi}}{\partial\hat{\alpha}^{(0)}}\,\delta\hat{\alpha}^{(0)}$$

with $\delta\hat{\alpha}^{(0)}$ arbitrary within the set (i.e. if $\tilde{\alpha}_i$ is an arbitrary number, $\delta\hat{\alpha}_i^{(0)}$ is an arbitrary number; if $\tilde{\alpha}_i$ is an arbitrary function of certain variables, then $\delta\hat{\alpha}_i^{(0)}$ is an arbitrary function of those variables). We now expand $\delta\hat{\psi}$ in powers of ν to find, using (9).

$$\delta\hat{\psi} = \hat{\psi}_1^{(0)}\,\delta\hat{\alpha}^{(0)} + \nu\,\hat{\psi}_2^{(0)}\,\hat{\alpha}^{(1)}\,\delta\hat{\alpha}^{(0)}$$

$$+ \nu^2\left[\hat{\psi}_2^{(0)}\,\hat{\alpha}^{(2)}\,\delta\hat{\alpha}^{(0)} + \tfrac{1}{2}\hat{\psi}_3^{(0)}\left(\hat{\alpha}^{(1)}\right)^2\delta\hat{\alpha}^{(0)}\right] + \cdots \qquad \text{(XIII-10)}$$

$$\equiv \delta\hat{\psi}^{(0)} + \nu\,\delta\hat{\psi}^{(1)} + \nu^2\,\delta\hat{\psi}^{(2)} + \cdots \qquad \text{(XIII-11)}$$

If we now substitute these formulae into the basic equation of the variational method, the results will be in the form of power series in ν equal to zero. Since ν is arbitrary this means that the coefficients of each power of ν must vanish separately. More specifically let us write the equations in the general form (III-5)

$$\delta\hat{J} = 0 \qquad \text{(XIII-12)}$$

where

$$\hat{J} = (\hat{\psi}, (H - \hat{E})\,\hat{\psi}) = 0 \qquad \text{(XIII-13)}$$

If now we insert the expansions, $\hat{J}$ will assume the form

$$\hat{J} = \sum_{n=0}^{\infty} \nu^n\, \hat{J}^{(n)} \qquad \text{(XIII-14)}$$

From which we will derive the sequence of equations

$$\delta\hat{J}^{(n)} = 0 \qquad \text{(XIII-15)}$$

and

$$\hat{J}^{(n)} = 0 \qquad \text{(XIII-16)}$$

The first few $\hat{J}^{(n)}$ are given by

$$\hat{J}^{(0)} = (\hat{\psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\,\hat{\psi}^{(0)}) \qquad \text{(XIII-17)}$$

$$\hat{J}^{(1)} = (\hat\psi^{(0)}, (H^{(0)}-\hat{E}^{(0)})\hat\psi^{(1)}) + (\hat\psi^{(1)}, (H^{(0)}-\hat{E}^{(0)})\hat\psi^{(0)}) + (\hat\psi^{(0)}, (H^{(1)}-\hat{E}^{(1)})\hat\psi^{(0)}) \qquad \text{(XIII-18)}$$

$$\hat{J}^{(2)} = (\hat\psi^{(0)}, (H^{(0)}-\hat{E}^{(0)})\hat\psi^{(2)}) + (\hat\psi^{(2)}, (H^{(0)}-\hat{E}^{(0)})\hat\psi^{(0)}) + (\hat\psi^{(0)}, (H^{(1)}-\hat{E}^{(1)})\hat\psi^{(1)}) + (\hat\psi^{(1)}, (H^{(1)}-\hat{E}^{(1)})\hat\psi^{(0)}) + (\hat\psi^{(1)}, (H^{(0)}-\hat{E}^{(0)})\hat\psi^{(1)}) - \hat{E}^{(2)}(\hat\psi^{(0)}, \hat\psi^{(0)})$$

etc.

The equations (15) and (16) now enable us to determine the $\hat\alpha^{(n)}$ and the $\hat{E}^{(n)}$ recursively. Thus from $\delta\hat{J}^{(0)}=0$, $\hat{J}^{(0)}=0$ we will determine $\hat\alpha^{(0)}$ and $\hat{E}^{(0)}$. Then $\delta\hat{J}^{(1)}$ yields, using (18), (11) and (10).

$$\{(\hat\psi^{(0)}_{,}\delta\hat\alpha^{(0)}, (H^{(0)}-\hat{E}^{(0)})\hat\psi^{(1)}) + (\hat\psi^{(0)}_{,,}\hat\alpha^{(1)}\delta\hat\alpha^{(0)}, (H^{(0)}-\hat{E}^{(0)})\hat\psi^{(0)}) + (\hat\psi^{(0)}_{,}\delta\hat\alpha^{(0)}, (H^{(1)}-\hat{E}^{(1)})\hat\psi^{(0)})\} + c.c. \qquad \text{(XIII-19)}$$

which together with $\hat{J}^{(1)}=0$ will serve to determine the $\hat\alpha^{(1)}$ and $\hat{E}^{(1)}$, etc., etc.

This kind of approach has been systematically pursued by J. N. Silverman and collaborators for cases in which the $\tilde\alpha$ are parameters. (See Phys. Rev. 162, 1175 (1967), and further references there.) However, it is implicit in many applications of perturbation theory within the variational method.

It should be noted however that the assumption that $\hat{\psi}$ can be expanded in integral powers of ν does not require that $\hat{\alpha}$ can also be so expanded. For example $\hat{\alpha}$ might involve $\sqrt{\nu}$, and an example of this sort is mentioned in the next section. Thus our discussion is not completely general. However, the general pattern is correct. Thus we will have $\delta\hat{\psi}$ of the general form (11) and hence one will have (15) and (16). However, the detailed formulae for the $\hat{\psi}^{(n)}$ and the $\delta\hat{\psi}^{(n)}$ as given by (9) and (10) will not apply.

Problem: For $\nu \neq 0$ we could evidently equally well vary $\hat{\alpha}$ by making changes only in $\hat{\alpha}^{(n)}$ say. Show that use of

$$\delta\hat{\psi} = \frac{\partial\hat{\psi}}{\partial\hat{\alpha}^{(n)}}\,\delta\hat{\alpha}^{(n)}$$

with $\delta\hat{\alpha}^{(n)}$ arbitrary yields the same equations as above

In our discussion so far we have assumed the space of trial functions to be independent of ν. It is also of interest to discuss one particular sort of situation in which this is not the case. It arises naturally for example in extending restricted Hartree-Fock to include the external fields. Also it arises in a more general connection which we will discuss in Sec. XV.

Namely suppose that

$$\hat{\psi} = \hat{\psi}^{(0)} + \nu\,\hat{\psi}^{(1)} + \nu^2\,\hat{\psi}^{(2)} + \cdots \qquad \text{(XIII-20)}$$

where $\hat{\psi}^{(0)}$ is a given function (usually derived from some set of trial functions via $\delta\hat{J}^{(0)} = 0$ but not necessarily) and where the $\hat{\psi}^{(n)}$ contain arbitrary parameters and/or functions whose optimal values we require to be independent of ν. (Note that whatever the source of

$\hat\psi^{(10)}$ in notation is then consistent since evidently for $V=0$ the optimal $\tilde\psi$ in the present context is $\hat\psi^{(10)}$.)

Varying these quantities we will then evidently be led again to equations of the form (15) except that now $\delta\hat\psi^{(10)}\equiv 0$ and for the $\delta\hat\psi^{(1n)}$ we as yet have no formulae like (15); for the moment they are simply the variations of the $\hat\psi^{(1n)}$ and nothing more. It is now important to notice that these equations now impose restrictions on the kinds of variational parameters and/or functions we can imbed in the $\hat\psi^{(1n)}$! Thus consider the equation $\delta\hat J^{(11)}=0$. It yields

$$(\delta\hat\psi^{(11)},(H^{(10)}-\hat E^{(10)})\hat\psi^{(10)})+c.c.=0 \qquad \text{(XIII-21)}$$

and hence tells us that $\delta\hat\psi^{(11)}$ must be an allowed variation of $\hat\psi^{(10)}$. Turning now to $\delta\hat J^{(12)}=0$ let us suppose for generality that $\hat\psi^{(12)}$ involves not only new quantities $\hat\alpha^{(12)}$ but also contains the $\hat\alpha^{(11)}$ imbedded in $\hat\psi^{(11)}$. Then

$$\delta\hat\psi^{(12)}=\frac{\partial\hat\psi^{(12)}}{\partial\hat\alpha^{(11)}}\delta\hat\alpha^{(11)}+\frac{\partial\hat\psi^{(12)}}{\partial\hat\alpha^{(12)}}\delta\hat\alpha^{(12)}$$

whence, since the $\hat\alpha^{(11)}$ and $\hat\alpha^{(12)}$ are independent, $\delta\hat J^{(12)}=0$ implies

$$\Big\{(\delta\hat\psi^{(11)},(H^{(10)}-\hat E^{(10)})\hat\psi^{(11)}+(H^{(11)}-\hat E^{(11)})\hat\psi^{(10)}) + \Big(\frac{\partial\hat\psi^{(12)}}{\partial\hat\alpha^{(11)}}\delta\hat\alpha^{(11)},(H^{(10)}-\hat E^{(10)})\hat\psi^{(10)}\Big)\Big\}+c.c.=0 \qquad \text{(XIII-22)}$$

and

$$\Big(\frac{\partial\hat\psi^{(12)}}{\partial\hat\alpha^{(12)}}\delta\hat\alpha^{(12)},(H^{(10)}-\hat E^{(10)})\hat\psi^{(10)}\Big)+c.c.=0 \qquad \text{(XIII-23)}$$

Eq. (23) then requires that $\frac{\partial\hat\psi^{(12)}}{\partial\hat\alpha^{(12)}}\delta\hat\alpha^{(12)}$ be an allowed variation of while (22) with (21) may or may not impose restricts on $\partial\hat\psi^{(12)}/\partial\hat\alpha^{(11)}$.

We have taken the point of view that $\hat{\psi}^{(0)}$ is given and that therefore (21) implies restrictions on $\tilde{\psi}^{(1)}$. However in some cases one might want to consider the $\hat{\psi}^{(n)}$. As given, whence (21) yields requirements on $\hat{\psi}^{(0)}$.

Problem Generalize (20) by replacing $\hat{\psi}^{(0)}$ by $\tilde{\psi}^{(0)}$ and discuss the consistency conditions, show that all is consistent if all the $\hat{\psi}^{(n)}$ are drawn from a common linear space but that then (20) and $\tilde{\psi}$ independent of ν are equivalent.

XIV. A General ν^{2n+1} Theorem

As we have discussed, if one knows an exact eigenfunction correctly through order ν^{n}, then one can calculate the exact eigenvalue correctly through order ν^{2n+1}. We now want to exhibit a similar theorem within the context of the variational method.

For generality we will allow the trial functions to depend explicitly on ν. Thus

$$\tilde{\psi} = \phi(\tilde{\alpha}, \nu) \tag{XIV-1}$$

Now consider the function $\overset{*}{\psi}$ defined by

$$\overset{*}{\psi} = \phi(\overset{*}{\alpha}, \nu) \qquad + \text{ possibly terms of order } \nu^{2n+2} \tag{XIV-2}$$

where

$$\overset{*}{\alpha} - \hat{\alpha} = O(\nu^{n+1}) \tag{XIV-3}$$

Thus the implicit ν dependence of $\overset{*}{\psi}$ differs from that of $\hat{\psi}$ by terms of order ν^{n+1}, while its explicit ν dependence must be the same at least to order ν^{2n+1}.

We now define

$$\Delta = \overset{*}{\psi} - \hat{\psi} \tag{XIV-4}$$

Then one readily finds from

$$\overset{*}{E} = (\overset{*}{\psi}, H\overset{*}{\psi})/(\overset{*}{\psi},\overset{*}{\psi})$$

$$\hat{E} = (\hat{\psi}, H\hat{\psi})/(\hat{\psi},\hat{\psi})$$

That

$$\overset{*}{E} - \hat{E} = \{ (\Delta, (H-\hat{E})\hat{\psi}) + (\hat{\psi}, (H-\hat{E})\Delta]\}/(\overset{*}{\psi},\overset{*}{\psi})$$

$$+ (\Delta, (H-\hat{E})\Delta)/(\overset{*}{\psi},\overset{*}{\psi}) \qquad \text{(XVI-5)}$$

Now from (2) and (3) it follows that

$$\Delta = (\overset{*}{\alpha} - \hat{\alpha})\frac{\partial\hat{\psi}}{\partial\hat{\alpha}} + O(\nu^{2n+2}) \qquad \text{(XIV-6)}$$

Then we note that

$$\Delta\hat{\psi} \equiv (\overset{*}{\alpha} - \hat{\alpha})\frac{\partial\hat{\psi}}{\partial\hat{\alpha}}$$

is (i) of order ν^{n+1} so that the last term in (5) is of order ν^{2n+2}, and (ii) is an allowed variation of $\hat{\psi}$ so that it makes no contribution to the sum of the first two terms in (5). Thus it follows that even if $(H-\hat{E})\hat{\psi}$ is of order zero, still

$$\overset{*}{E} - \hat{E} = O(\nu^{2n+2}) \qquad \text{(XIV-7)}$$

Thus if the $\hat{\alpha}$ involve only integral power of ν we have the result that if we know $\hat{\alpha}$ correctly through order ν^{n} and if we include any explicit ν dependence correctly at least through order ν^{2n+1}, then we can calculate $\hat{E}$ correctly through order ν^{2n+1}. In particular then if the space of trial functions is independent of

ν we have a direct generalization of <u>the</u> ν^{2n+1} theorem. Of course if $(H-\hat{E})\hat{\psi} = O(\nu^{s})$ where $s>0$ we can weaken the additional requirement on the explicit ν dependence. Namely to derive (7) we need only that

$$\overset{*}{\psi} = \Phi\,(\overset{*}{\alpha},\nu) + \text{possibly terms of } O(\nu^{T}) \qquad \text{(XIV-8)}$$

where $T = 2n+2-s$ or $n+1$, whichever is larger.

<u>Problem</u>: Show that (8) works.

<u>Problem</u>: Show that for $s=\infty$ we have <u>the</u> ν^{2n+1} theorem.

<u>Problem</u>: Would you expect the ν^{2n+1} theorem to have useful consequence for calculations of the sort exemplified by XIII-20? How about linear variational calculations in which the basis functions depend on ν ?

If the $\hat{\alpha}$ say involve half integral powers of ν (we will give an example in a moment) then (3) requires that we know $\hat{\alpha}$ correctly through order $\nu^{n+1/2}$. However, in some cases accuracy through order ν^{n} will still suffice. Namely suppose we replace (3) by

$$\overset{*}{\alpha} - \hat{\alpha} = O(\nu^{n+1/2})$$

Then instead of (6) we will have

$$\Delta = \Delta\hat{\psi} + O(\nu^{2n+1})$$

Now as it stands $\Delta\hat{\psi}$ looks to be of order $\nu^{n+1/2}$ but if, as we assume $\hat{\psi}$ and $\overset{*}{\psi}$ can be expanded in <u>integral</u> powers of ν, $\Delta\hat{\psi}$ must in fact be of order ν^{n+1}. Thus the last term in (5) is again of order ν^{2n+2}. Turning to the first two terms, $\Delta\hat{\psi}$ is still an allowed variation, so that only the $O(\nu^{2n+1})$ part of Δ contributes. This contribution will then be of order ν^{2n+2} if

$$(H-\hat{E})\,\hat{\psi} = O(\nu) \tag{XIV-9}$$

As a case in point consider the ground state of Helium with $\nu H^{(1)}$ the electron-electron interaction and $\tilde{\psi}$ that linear combination of $|\tilde{\Delta}\alpha\,\tilde{\Delta}'\beta|$ and $|\tilde{\Delta}'\alpha\,\tilde{\Delta}\beta|$ which yields a spin singlet. Then one finds (see Hibbert, Proc. Phys. Soc. 91, 33 (1967) and references there. See also Hibbert, Proc. Phys. Soc. 1, B1048 (1968) for a similar method applied to Li. For some reason in these papers the method is referred to as URHF) that although $\hat{\psi}$ can be expanded in integral powers of $\nu = \frac{1}{Z}$, the $\hat{a}$, namely $\hat{\Delta}$ and $\hat{\Delta}'$ involve $\sqrt{\nu}$. However $\hat{\psi}$ is exact in the limit $\nu \to 0$ $(Z \to \infty)$ and hence (9) is satisfied. Thus in this case one does find that knowing the orbitals correctly through order ν^{n} enables one to calculate $\hat{E}$ correctly through order ν^{2n+1}.

Problem: Give an analogous discussion for other fractional powers

Problem: Since for a linear space (which may depend on ν), $\hat{\psi}$ is an exact eigenfunction of $\bar{H}$, it follows that if $\chi - \hat{\psi} = O(\nu^{n+1})$ then

$$\bar{E} - \hat{E} = O(\nu^{2n+2})$$

where

$$\bar{E} = (\chi, \bar{H}\chi)/(\chi,\chi)$$

What is the connection between this result and that found above?

XV. The Variational Method as a Means for Approximating Perturbation Theory

As we have shown in Sec. IV, if, in the variational method, $\delta\hat{\psi}$ can be completely arbitrary, then $\hat{\psi}$ and $\hat{E}$ satisfy the Schroedinger equation. Similarly we will now show that if $\delta\hat{\psi}^{(0)}$ can be completely arbitrary, then the $\hat{\psi}^{(m)}$ and $\hat{E}^{(m)}$ satisfy the basic equations of perturbation theory. The proof requires little more then writing out the successive equations $\delta\hat{J}^{(m)}=0$. Thus writing out the equations for $m=0,1,2$ we have

$$(\delta\hat{\psi}^{(0)}, (H^{(0)}-\hat{E}^{(0)})\hat{\psi}^{(0)}) + c.c. = 0 \tag{XV-1}$$

$$\begin{aligned}\{(\delta\hat{\psi}^{(0)}, (H^{(0)}-\hat{E}^{(0)})\hat{\psi}^{(1)} + (H^{(1)}-\hat{E}^{(1)})\hat{\psi}^{(0)})\\ +(\delta\hat{\psi}^{(1)}, (H^{(0)}-\hat{E}^{(0)})\hat{\psi}^{(0)})\} + c.c. = 0\end{aligned} \tag{XV-2}$$

$$\begin{aligned}\{(\delta\hat{\psi}^{(0)}, (H^{(0)}-\hat{E}^{(0)})\hat{\psi}^{(2)} + (H^{(1)}-\hat{E}^{(1)})\hat{\psi}^{(1)}-\hat{E}^{(2)}\hat{\psi}^{(0)})\\ +(\delta\hat{\psi}^{(1)}, (H^{(0)}-\hat{E}^{(0)})\hat{\psi}^{(1)} + (H^{(1)}-\hat{E}^{(1)})\hat{\psi}^{(0)})\\ +(\delta\hat{\psi}^{(2)}, (H^{(0)}-\hat{E}^{(0)})\hat{\psi}^{(0)})\} + c.c. = 0\end{aligned} \tag{XV-3}$$

Now if $\delta\hat{\psi}^{(0)}$ is completely arbitrary then (1) yields

$$(H^{(0)}-\hat{E}^{(0)})\hat{\psi}^{(0)}=0 \tag{XV-4}$$

which of course is just the usual starting equation of perturbation theory, telling us that $\hat{\psi}^{(0)}$ and $\hat{E}^{(0)}$ should be an eigenfunction, and eigenvalue respectively of $H^{(0)}$.

Turning now to (2) and using (4), $\delta\hat{\psi}^{(0)}$ completely arbitrary yields the familiar first order equation

$$(H^{(0)} - \hat{E}^{(0)})\hat{\psi}^{(1)} + (H^{(1)} - \hat{E}^{(1)})\hat{\psi}^{(0)} = 0 , \qquad \text{(XV-5)}$$

and then (3) with (4), (5), and $\delta\hat{\psi}^{(0)}$ completely arbitrary, yields the second order equation

$$(H^{(0)} - \hat{E}^{(0)})\hat{\psi}^{(2)} + (H^{(1)} - \hat{E}^{(1)})\hat{\psi}^{(1)} - \hat{E}^{(2)}\hat{\psi}^{(0)} = 0 \qquad \text{(XV-6)}$$

etc., etc.

Further, in accord with the discussion in Sec. XIII, we get the same equations if we put $\delta\hat{\psi}^{(0)} = 0$ and let $\delta\hat{\psi}^{(1)}$ be completely arbitrary. Thus with $\delta\hat{\psi}^{(0)} = 0$ (1) yields no information, but (2) with $\delta\hat{\psi}^{(0)} = 0$ and $\delta\hat{\psi}^{(1)}$ completely arbitrary yields (4). Then using (4), (3) with $\delta\hat{\psi}^{(0)} = 0$ and $\delta\hat{\psi}^{(1)}$ complete arbitrary, yields (5) etc. Similarly we set the same results if we put $\delta\hat{\psi}^{(0)}$ and $\delta\hat{\psi}^{(1)}$ equal to zero and let $\delta\hat{\psi}^{(2)}$ be completely arbitrary, etc.

Having now seen in detail how the variational method contains perturbation theory, the following possibility immediately presents itself: Just as one uses $\delta\hat{J} = 0$ to derive approximations to the solutions of the Schroedinger equation, so one might try to use the $\delta\hat{J}^{(n)} = 0$ to find approximations to the solutions of the perturbation equations. The next several sections will be devoted to investigating this program.

Thus we assume that $\hat{\psi}^{(0)}$ and $\hat{E}^{(0)}$ are exact, i.e., that (4) is satisfied. Also we will assume for convenience that $\hat{\psi}^{(0)}$ is normalized

$$(\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)}) = 1 \qquad \text{(XV-7)}$$

This being the case we will then denote $\hat{\Psi}^{(0)}$ and $\hat{E}^{(0)}$ by $\Psi^{(0)}$ and $E^{(0)}$ respectively. Further we will implicitly assume that $\Psi^{(0)}$ is essentially unique, i.e. that $E^{(0)}$ is non-degenerate. A discussion of degenerate perturbation theory can be found in J. Chem. Phys. 44, 3266 (1966). Then from $\hat{J}^{(1)} = 0$ it follows, that, as we expect, $\hat{E}^{(1)}$ is also exact, i.e.,

$$\hat{E}^{(1)} = (\Psi^{(0)}, H^{(1)} \Psi^{(0)}) \equiv E^{(1)} \qquad \text{(XV-8)}$$

We now turn to the determination of $\hat{\Psi}^{(1)}$, which we will assume we can't determine exactly. Since we have been "handed" $\hat{\Psi}^{(0)}$ it is natural to base our discussion on (XIII-20). Then since $\hat{\Psi}^{(0)}$ is exact and hence all variations are allowed, there are no restrictions imposed on the $\hat{\Psi}^{(1)}$, and from (XIII-21) we have as the equation to determine $\hat{\Psi}^{(1)}$

$$(\delta\hat{\Psi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\Psi}^{(1)} + (H^{(1)} - E^{(1)})\Psi^{(0)}) + c.c. = 0 \qquad \text{(XV-9)}$$

The equation $\hat{J}^{(2)} = 0$ then determines $\hat{E}^{(2)}$:

$$\hat{E}^{(2)} = (\hat{\Psi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\Psi}^{(1)}) + (\hat{\Psi}^{(1)}, (H^{(1)} - E^{(1)})\Psi^{(0)}) + (\Psi^{(0)}, (H^{(1)} - E^{(1)})\hat{\Psi}^{(1)}) \qquad \text{(XV-10)}$$

XVI. The Hylleraas Variational Method

From their structures and from the remark that there are no *a priori* restrictions on $\delta\hat{\Psi}^{(1)}$, it is clear that (XV-9) and (XV-10) can be viewed as describing a process quite analogous to that of the variational method. Namely define a trial second order energy by

$$\tilde{E}^{(2)} = (\tilde{\Psi}^{(1)}, (H^{(0)} - E^{(0)})\tilde{\Psi}^{(1)}) + (\hat{\Psi}^{(1)}, (H^{(1)} - E^{(1)})\Psi^{(0)}) + (\Psi^{(0)}, (H^{(1)} - E^{(1)})\hat{\Psi}^{(1)}) \qquad \text{(XVI-1)}$$

where $\tilde{\Psi}^{(1)}$ is a trial first order wave function drawn from some set. Then (XV-9) says that $\hat{\Psi}^{(1)}$ is that $\tilde{\Psi}^{(1)}$ which makes $\tilde{E}^{(2)}$ stationary for variations of $\tilde{\Psi}^{(1)}$ within the set, while (XV-10) says that the stationary value of $\tilde{E}^{(2)}$ is $\hat{E}^{(2)}$. This observation also serves as a statement of the Hylleraas variation method, and we will call the functional on the r.h.s. of (1) the Hylleraas functional and denote it by $\tilde{J}_H^{(2)}$. Thus (XV-9) and (XV-10) can be written

$$\delta \hat{J}_H^{(2)} = 0 \qquad , \qquad \hat{J}_H^{(2)} = 0 \qquad \text{(XVI-2)}$$

The (probably impatient) reader will no doubt have noticed that we could have written $\tilde{J}_H^{(2)}$ down directly as yielding the variational equivalent of eq. (XV-5), and thereby could have avoided all the discussion of the preceding section. With this sentiment we agree. Our only plea is that we have thereby placed (2) in a more general setting.

$\tilde{J}^{(2)}_{H}$ can also be related to the variational functional for a Green's function given in (XII-4). Thus consider the equation

$$(H^{(0)} - z)\,\phi + \chi = 0 \qquad \text{(XVI-3)}$$

where the real number z, and the function χ are given. The appropriate variational functional is then

$$(\tilde{\phi}, (H^{(0)} - z)\tilde{\phi}) + (\tilde{\phi}, \chi) + (\chi, \tilde{\phi}) \qquad \text{(XVI-4)}$$

Now the solution of (3) is (we assume that z is not equal to an eigenvalue of $H^{(0)}$)

$$\phi = -G\chi \qquad \text{(XVI-5)}$$

where G is the Green's function for $(H^{(0)} - z)$. (Note that in the present discussion $H^{(0)}$ is playing the role of H in Sec. XII.) Let us now write

$$\tilde{\phi} = -\tilde{G}\chi$$

where $\tilde{G}$ is a Hermitian integral operator (a trial Green's function). Then (4) becomes

$$(\chi, \tilde{G}(H^{(0)} - z)\tilde{G}\chi) + 2(\chi, \tilde{G}\chi) \qquad \text{(XVI-6)}$$

which is just the average of the functional (XII-4)

We will now show that for the ground state

$$\tilde{E}^{(2)} \geq E^{(2)} \qquad \text{(XVI-7)}$$

where $E^{(2)}$ is the exact second order energy. We will discuss the situation for excited states in Sec. XIX. Till then we will be concerned exclusively with the ground state and hence will not keep mentioning this fact explicitly. Thus $\tilde{E}^{(2)}$ is not only a variational approximation to $E^{(2)}$, it is an upper bound and hence the Hylleraas variational method partakes of the benefits that we have discussed earlier in connection with the general variational method.

The proof is simple. We write

$$\tilde{\psi}^{(1)} = \psi^{(1)} + \tilde{\Delta}^{(1)} \qquad \text{(XVI-8)}$$

where $\psi^{(1)}$ is the exact first order function satisfying

$$(H^{(0)} - E^{(0)})\,\psi^{(1)} + (H^{(1)} - E^{(1)})\,\psi^{(0)} = 0 \qquad \text{(XVI-9)}$$

Then using (8) and (9) one readily finds

$$\tilde{E}^{(2)} = E^{(2)} + (\tilde{\Delta}^{(1)}, (H^{(0)} - E^{(0)})\tilde{\Delta}^{(1)}) \qquad \text{(XVI-10)}$$

where $E^{(2)}$ is the exact second order energy:

$$E^{(2)} = (\psi^{(0)}, (H^{(1)} - E^{(1)})\,\psi^{(1)}) \qquad \text{(XVI-11)}$$

For $E^{(0)}$ the lowest eigenvalue of $H^{(0)}$ we then clearly have (7).

The fact that we have a minimal principle may not seem surprising. After all the general variational method yields a minimum under the similar circumstances. However, it is in fact not trivial and is linked to our assumption that $\hat{\psi}^{(0)}$ is exact. Namely we know that for the ground state $\hat{E} \geq E$ and therefore for all ν

$$\hat{E}^{(0)} + \nu \hat{E}^{(1)} + \nu^2 \hat{E}^{(2)} + \cdots \geq E^{(0)} + \nu E^{(1)} + \nu^2 E^{(2)} + \cdots \qquad \text{(XVI-12)}$$

Now in the absence of further information, about all that we can conclude about the individual $\hat{E}^{(n)}$ is what we know anyway, namely that

$$\hat{E}^{(0)} \geq E^{(0)}$$

which follows from (12) when $\nu = 0$. On the other hand with $\hat{\psi}^{(0)}$ exact so that $\hat{E}^{(0)} = E^{(0)}$ and $\hat{E}^{(1)} = E^{(1)}$, we find, dividing by ν^2 and setting $\nu = 0$, that

$$\hat{E}^{(2)} \geq E^{(2)} \tag{XVI-13}$$

as before.

Following the argument further, if $\hat{\psi}^{(1)}$ is also exact then we expect (and can show from $\hat{J}^{(2)} = \hat{J}^{(3)} = 0$) that $\hat{E}^{(2)}$ and $\hat{E}^{(3)}$ will also be exact, whence we infer from (13) that

$$\hat{E}^{(4)} \geq E^{(4)} \tag{XVI-14}$$

The result (14) of course also follows from the variational method under the same conditions. Namely with $\hat{\psi}^{(0)} = \psi^{(0)}$ and $\hat{\psi}^{(1)} = \psi^{(1)}$ we put $\delta\hat{\psi}^{(0)} = \delta\hat{\psi}^{(1)} = 0$. Then $\delta\hat{J}^{(4)} = 0$ yields the first new information, namely an equation for $\hat{\psi}^{(2)}$ of the form (Problem: Fill in the details)

$$\delta\hat{J}_{H}^{(4)} = 0 \tag{XVI-15}$$

where $\tilde{J}_{H}^{(4)}$ is, not surprisingly, just the variational functional that one would associate with the second order equation

$$(H^{(0)} - E^{(0)})\psi^{(2)} + (H^{(1)} - E^{(1)})\psi^{(1)} - E^{(2)}\psi^{(0)} = 0$$

That is

$$\begin{aligned} \tilde{J}_{H}^{(4)} = {} & (\tilde{\psi}^{(2)}, (H^{(0)} - E^{(0)})\tilde{\psi}^{(2)}) + (\tilde{\psi}^{(2)}, (H^{(1)} - E^{(1)})\psi^{(1)}) \\ & + (\psi^{(1)}, (H^{(1)} - E^{(1)})\tilde{\psi}^{(2)}) - E^{(2)}(\tilde{\psi}^{(2)}, \psi^{(0)}) \\ & - E^{(2)}(\psi^{(0)}, \tilde{\psi}^{(2)}) \end{aligned} \tag{XVI-16}$$

Further it follows from $\hat{J}^{(4)} = 0$ that

$$\hat{J}_{H}^{(4)} = \hat{E}^{(4)} \tag{XVI-17}$$

From this we can then directly infer (14). Namely we know that $\tilde{J}^{(2)}_H$ is stationary when $\tilde{\phi}^{(2)} = \psi^{(2)}$ and from (17) it follows that when $\tilde{\phi}^{(2)} = \psi^{(2)}$ its numerical value is $E^{(2)}$. Now if we write

$$\tilde{\phi}^{(2)} = \psi^{(2)} + \tilde{\delta}^{(2)}$$

then the only possible source of 'second order terms' in $\tilde{J}^{(2)}_H$ is first term on the r.h.s. of (16), whence we infer that

$$\tilde{J}^{(2)}_H = E^{(2)} + (\tilde{\delta}^{(2)}, (H^{(0)} - E^{(0)}) \tilde{\delta}^{(2)}) \qquad \text{(XVI - 18)}$$

which proves the point.

The Hylleraas variational method, eqs. (1), (15), and their generalizations to higher even orders, has been widely used to yield approximations to perturbation theory, for example in the calculation of static polarizibilities and susceptibilities, in the calculation of the effect of Coulomb repulsion ($1/Z$ approximation), in the calculation of long range forces etc., etc., and many of these applications will be discussed in detail in later chapters. (Note that since $E^{(2)}$ is less than zero, an upper bound for $E^{(2)}$ yields a lower bound for polarizibilities etc.)

We would emphasize however that (7) holds only if $\hat{\phi}^{(0)} = \psi^{(0)}$, while (18) holds only if $\hat{\phi}^{(0)} = \psi^{(0)}$ and $\hat{\phi}^{(1)} = \psi^{(1)}$ etc. Now even granting that $\hat{\phi}^{(0)} = \psi^{(0)}$, usually one then derives $\hat{\phi}^{(1)}$ from (1) and then goes on to use it in (15) etc., (also $\hat{E}^{(3)}$ is usually calculated from the formulae one gets from $\hat{J}^{(3)}_{2D}$ assuming that $\hat{\phi}^{(1)}$ is exact) and one may well wonder about the legitimacy of such a procedure. In general of course one can only hope that one has achieved a sufficiently accurate approximation to $\psi^{(1)}$, etc., so that all is well numerically.

One should also note that the formulae of the Hylleraas variational method with the symbols suitably reinterpreted also apply to perturbation theory within the variational method as discussed in Sec. XIII if the space of trial function is <u>linear</u>. One can see this in several ways. First of all in a linear space $\hat{\psi}^{(0)}_{n} = 0$ for $n > 1$. Thus (XIII-19) has the form (XV-9) and similarly for the higher orders. Another approach is to note that $\hat{\psi} = \bar{\psi}$ where $\bar{\psi}$ is an eigenfunction of $\bar{H}$. Thus $\hat{\psi}^{(0)} = \bar{\psi}^{(0)}$, the exact $\hat{\psi}^{(n)} = \bar{\psi}^{(n)}$ etc, Further if Π is independent of v then $\bar{H}^{(0)} = \overline{H^{(0)}}$ and $\bar{H}^{(n)} = \overline{H^{(n)}}$ and all higher order perturbation vanish. Thus we may then apply all the above formulae except that we should put bars on $H^{(0)}$, $H^{(n)}$ and the $E^{(2n)}$ and $\psi^{(n)}$. However we may then drop the bars on $H^{(0)}$ and $H^{(n)}$ since the functions all lie in the space.

Now as a practical matter, one can, as we have discussed before, usually completely carry through a calculation in a linear space, i.e. find the $\bar{\psi}^{(n)}$, only if the space is finitely parametrized. If the space is really infinite (the S-limit in He with $v\overline{H^{(n)}}$ the electron-electron interaction is an example) then one must resort to approximation. If one uses a finitely parametrized linear subspace, then the Hylleraas formulae of course again Apply with further reinterpretation of the symbols. If one uses non-linear parameters too, then in any case, the analysis of Sec. XIII applies. If however one does use the formulae of the Hylleraas method, then one must hope that $\hat{\psi}^{(n)}$ is a sufficiently accurate approximation to $\bar{\psi}^{(n)}$ etc.

In the preceding discussion we have assumed that $\hat{\psi}^{(0)}$ is exact and worried about $\hat{\psi}^{(n)}$ etc. However if $H^{(0)}$ is the exact Hamiltonian of an isolated system (with more than 1 electron) then one rarely knows $\psi^{(0)}$ with any precision. Nevertheless, many workers have used (1) and (2) with $\psi^{(0)}$ replaced by an approximate $\hat{\psi}^{(0)}$ and have taken $\hat{J}^{(2)}$ as an approximation to $E^{(2)}$. We will discuss some properties of this approximation in Sec. XXI.

XVII. The Hylleraas Variational Method for the Second Order Energy-Miscellaneous Results

(A) If there are no reality restrictions on $\delta\hat{\psi}^{(1)}$ then we can replace (XV-8) by

$$(\delta\hat{\psi}^{(1)}, (H^{(0)}-E^{(0)})\hat{\psi}^{(1)} + (H^{(1)}-E^{(1)})\psi^{(0)}) = 0 \qquad \text{(XVII-1)}$$

with its obvious interpretation in terms of moments.

(B) If

$$\Delta\hat{\psi}^{(1)} = \delta\hat{\eta}\,\hat{\psi}^{(1)} \qquad \text{(XVII-2)}$$

is an allowed variation with $\delta\hat{\eta}$ a real parameter then (9), together with (10) yields

$$\hat{E}^{(2)} = \frac{1}{2}\left[(\psi^{(0)}, (H^{(1)}-E^{(1)})\hat{\psi}^{(1)}) + c.c.\right] \qquad \text{(XVII-3)}$$

If in addition

$$\Delta\hat{\psi}^{(1)} = i\,\delta\hat{\eta}\,\hat{\psi}^{(1)}. \qquad \text{(XVII-4)}$$

is allowed then one finds

$$\hat{E}^{(2)} = (\psi^{(0)}, (H^{(1)}-E^{(1)})\psi^{(1)}) \qquad \text{(XVII-5)}$$

which is quite analogous to (XVI-10). (Problem: Derive (3) and (5).)

(<u>C</u>) For some purposes it is useful to replace the functional $\tilde{\mathcal{J}}^{(2)}_H$ by another, equivalent, functional which is derived from $\tilde{\mathcal{J}}^{(2)}_H$ as follows: Suppose that the overall scale of the $\tilde{\Psi}^{(1)}$ is arbitrary. That is, suppose that

$$\tilde{\Psi}^{(1)} = \tilde{C}\,\tilde{\Phi}^{(1)} \tag{XVII-6}$$

where $\tilde{C}$ is a variational parameter and $\tilde{\Phi}^{(1)}$ is independent of $\tilde{C}$. If there are no <u>a priori</u> reality restrictions on $\tilde{C}$, then the optimal $\tilde{C}$ for a given $\tilde{\Phi}$ (we will not introduce a special symbol for this $\tilde{C}$) is evidently determined by

$$(\delta\tilde{C})^* \left(\tilde{\Phi}^{(1)}, (H^{(0)}-E^{(0)})\,\tilde{C}\,\tilde{\Phi}^{(1)} + (H^{(1)}-E^{(1)})\,\psi^{(0)}\right) = 0$$

whence we have

$$\tilde{C} = -\frac{(\tilde{\Phi}^{(1)}, (H^{(0)}-E^{(1)})\,\psi^{(0)})}{(\tilde{\Phi}^{(1)}, (H^{(0)}-E^{(0)})\,\tilde{\Phi}^{(1)})}$$

If now we use this value for $\tilde{C}$ in the Hylleraas Functional, we get a new functional which we will call the Dirichlet Functional and denote by $\tilde{\mathcal{J}}^{(2)}_D$. Thus

$$\tilde{\mathcal{J}}^{(2)}_H = \tilde{\mathcal{J}}^{(2)}_D = -\frac{\left|(\tilde{\Phi}^{(1)}, (H^{(1)}-E^{(1)})\,\psi^{(0)})\right|^2}{(\tilde{\Phi}^{(1)}, (H^{(0)}-E^{(0)})\,\tilde{\Phi}^{(1)})} \tag{XVII-7}$$

$\delta\hat{\mathcal{J}}^{(2)}_D$ then yields equations to determine $\hat{\Phi}^{(1)}$ [$\hat{C}$ is then determined by use of (7)] and $\tilde{\mathcal{J}}^{(2)}_D$ and in particular $\hat{\mathcal{J}}^{(2)}_D$ yield upper bounds to $E^{(2)}$. The $\tilde{\mathcal{J}}^{(2)}_D$ furnished by the particular choice

$$\tilde{\Phi}^{(1)} = (H^{(1)}-E^{(1)})\,\psi^{(0)}$$

yields the so-called Hassé bound to $E^{(2)}$.

One sometimes reads that $\tilde{\mathcal{J}}_D^{(2)}$ is a better approximation to $E^{(2)}$ than is $\tilde{\mathcal{J}}_H^{(2)}$. In view of our discussion this sounds puzzling. What is meant is the following: Suppose that in $\tilde{\mathcal{J}}_H^{(2)}$ we insert *not* (6) with (7), but simply

$$\tilde{\Psi}^{(1)} = \tilde{\Phi}^{(1)}$$

Then the conclusion will clearly follow since we have not adjusted the overall scale of $\tilde{\Psi}^{(1)}$ in an optimal way. Thus for the *same* trial function, i.e., $\tilde{\Psi}^{(1)} = \tilde{\Phi}^{(1)}$, we have for the ground state

$$\tilde{\mathcal{J}}_H^{(2)} \geq \tilde{\mathcal{J}}_D^{(2)} \geq E^{(2)} \qquad \text{(XVII-8)}$$

In particular note that like $E^{(2)}$, $\tilde{\mathcal{J}}_D^{(2)}$ is definitely ≤ 0.

Problem: Prove (8) directly by showing that $\tilde{\mathcal{J}}_H^{(2)} - \tilde{\mathcal{J}}_D^{(2)}$ is ≥ 0 for the ground state.

(*D*) If one tries to prove directly that

$$\tilde{\mathcal{J}}_D^{(2)} \geq E^{(2)} \qquad \text{(XVII-9)}$$

by writing $\tilde{\Phi}^{(1)} = \Psi^{(1)} + \tilde{\Delta}$ etc., one finds that one must invoke the Schwartz inequality (*Problem*: Fill in the details). We will now show that in general (9) can be considered as a special case of the Schwartz inequality.

To avoid spurious singularities, let us replace $E^{(0)}$ by z where z is slightly less than $E^{(0)}$ so that $(H^{(0)} - z)$ is a strictly positive operator. At the end of the calculation we can then let $z \to E^{(0)}$. With this understanding, $E^{(2)}$ can be written

$$E^{(2)} = -\left(\Psi^{(0)}, (H^{(1)} - E^{(1)})(H^{(0)} - z)^{-1}(H^{(1)} - E^{(1)})\Psi^{(0)}\right)$$

which we will write as

$$E^{(2)} = -\left(\chi^{(0)}, \Omega^{-1}\chi^{(0)}\right)$$

where

$$\chi^{(0)} \equiv (H^{(1)} - E^{(1)})\psi^{(0)}$$

and

$$\Omega \equiv (H^{(0)} - z)$$

Now from our hypotheses $\Omega^{\pm 1/2}$ exist and are Hermitian, whence it follows from the Schwartz inequality, that for any $\tilde{\phi}^{(1)}$

$$(\chi^{(0)}, \Omega^{-1}\chi^{(0)})(\tilde{\phi}^{(1)}, \Omega\tilde{\phi}^{(1)}) = (\Omega^{-1/2}\chi^{(0)}, \Omega^{-1/2}\chi^{(0)})(\Omega^{1/2}\tilde{\phi}^{(1)}, \Omega^{1/2}\tilde{\phi}^{(1)})$$
$$\geq |(\Omega^{-1/2}\chi^{(0)}, \Omega^{1/2}\tilde{\phi}^{(1)})|^2$$

whence, since the r.h.s. equals $|(\chi^{(0)}, \tilde{\phi}^{(1)})|^2$ we have, after a bit of rearranging that

$$-\frac{|(\chi^{(0)}, \tilde{\phi}^{(1)})|^2}{(\tilde{\phi}^{(1)}, \Omega\tilde{\phi}^{(1)})} \geq E^{(2)}$$

which of course is just

$$\tilde{J}_D^{(2)} \geq E^{(2)}$$

<u>Problem</u>: Generalize the results (A) - (D) to higher orders.

(E) In many situations $vH^{(1)}$ describes the interaction of the system with two fields

$$v H^{(1)} = v_1 H_1^{(1)} + v_2 H_2^{(1)}$$

and one is interested in the cross term in $\nu^2 E^{(2)}$, i.e., the term proportional to $\nu_1\nu_2$ (shielding coefficints, electro-magnetic effects, etc.)

$$\nu^2 E^{(2)} = (\nu_1)^2 E^{(2)}_1 + \nu_1 \nu_2 E^{(2)}_{12} + (\nu_2)^2 E^{(2)}_2$$

Now from our discussion so far we have a variational upper bound to $E^{(2)}$, and hence a variational upper bound to $E^{(2)}_1$, (put $\nu_2 = 0$) and a variational upper bound to $E^{(2)}_2$ (put $\nu_1 = 0$)). However, this yields no hold on $E^{(2)}_{12}$. Thus for example, suppose that we would have

$$\nu^2 \hat{E}^{(2)} = (\nu_1)^2 \hat{E}^{(2)}_1 + \nu_1 \nu_2 \hat{E}^{(2)}_{12} + (\nu_2)^2 \hat{E}^{(2)}_2$$

Then from $\hat{E}^{(2)} \geq E^{(2)}$ we can't conclude anything about the relative sizes of $\hat{E}^{(2)}_{12}$ and $E^{(2)}_{12}$. In this situation several alternatives are open and have been used.

(i) Determine a $\nu^2 \hat{E}^{(2)}$ and take the term in it proportional to $\nu_1\nu_2$ as the "best" approximation to the $E^{(2)}_{12}$. Note however that if we define ζ by $\nu_1 = \nu\zeta$ and $\nu_2 = \nu(1-\zeta)$ then if $\tilde{\psi}^{(1)}$ involves non-linear parameters, then usually $\hat{E}^{(2)}$ will not simply be a quadratic form in ζ..

(ii) Put $\nu_2 = 0$ and determine an $\hat{E}^{(2)}_1$. Then use the $\hat{\psi}^{(1)}_1$ to compute a "best" approximation to $E^{(2)}_{12}$ according to

$$(\psi^{(0)}, (H^{(1)}_2 - E^{(1)}_2)\hat{\psi}^{(1)}_1) + c.c. \qquad \text{(XVII-11)}$$

where

$$E^{(1)}_2 \equiv (\psi^{(0)}, H^{(1)}_2 \psi^{(0)})$$

Rev. 172, 1788 (1968) Sec. II for some comments on this point.) However, we need not use $\tilde{\mathcal{J}}^{(2)}_{12}$ in this way. One could, for example, insert for $\tilde{\psi}^{(1)}_1$ the $\psi^{(1)}_1$ from (ii) and for $\tilde{\psi}^{(1)}_2$ the $\psi^{(1)}_2$ from (iii) with the hope that (14) would then provide a better approximation to $E^{(2)}_{12}$ then either (11) or (12).

> Problem: Show that if in the procedures (i) - (iv) one draws all the trial functions from a common linear space, then all four methods yield the SAME answer.

(v) Find a variational lower bound for $E^{(2)}$ (and hence for $E^{(2)}_1$ and $E^{(2)}_2$). That is, given the second order energy, call it $E^{(2)}_+$, for

$$v H^{(1)} = v\,(H_1^{(1)} + H_1^{(2)})$$

and the second order energy, call it $E^{(2)}_-$, for

$$v H^{(1)} = v\,(H_1^{(1)} - H_1^{(2)})$$

then evidently

$$2E^{(2)}_{12} = E^{(2)}_+ - E^{(2)}_-$$

Thus an upper (lower) bound for $E^{(2)}_+$ combined with a lower (upper) bound for $E^{(2)}_-$ yields an upper (lower) bound for $E^{(2)}_{12}$. We will discuss such lower bounds in chapter ______.

XVIII. The Linear Hylleraas Variational Method

In practice one often takes for $\tilde{\psi}^{(1)}$ the linear variational form

$$\tilde{\psi}^{(1)} = \sum_{k=1}^{M} \tilde{a}_k^{(1)} \phi_k \qquad \text{(XVIII-1)}$$

where the ϕ_k are a given basis set and the $\tilde{a}_k^{(1)}$ are variational parameters. The ϕ_k can of course also contain parameters, however we will consider them fixed for the purposes of the present discussions.

Now clearly we can replace the ϕ_k by a set of M functions χ_k with the properties

$$(\chi_k, \chi_L) = \delta_{kL} \qquad \text{(XVIII-2)}$$

$$(\chi_k, H^{(0)} \chi_L) = \epsilon_k \delta_{kL} \qquad \text{(XVIII-3)}$$

namely the χ_k and ϵ_k are just the $\hat{\psi}_k$ and $\hat{E}_k$ which emerge from the linear variational method of Sec. VI with $H = H^{(0)}$ and the ϕ_k again the basis set. Thus without loss of generality we may replace (1) by

$$\tilde{\psi}^{(1)} = \sum_{k=1}^{M} \tilde{b}_k^{(1)} \chi_k \qquad \text{(XVIII-4)}$$

Since we will impose no reality restrictions on the $\tilde{b}_k^{(1)}$ we can now use (XVII-2) to find, varying $\tilde{b}_k^{(1)}$, and using (2) and (3),

$$(\epsilon_k - E^{(0)}) \hat{b}_k^{(1)} + (\chi_k, (H^{(1)} - E^{(1)}) \psi^{(0)}) = 0$$

or

$$\hat{b}_k^{(1)} = -\frac{(\chi_k, (H^{(1)} - E^{(1)})\psi^{(0)})}{(\epsilon_k - E^{(0)})} \qquad \text{(XVIII-5)}$$

whence

$$\psi^{(1)} = -\sum_{k=1}^{M} \chi_k \frac{(\chi_k, (H^{(1)} - E^{(1)})\psi^{(0)})}{\epsilon_k - E^{(0)}} \qquad \text{(XVIII-6)}$$

and from (XVII-5) we have

$$\hat{E}^{(2)} = -\sum_{k=1}^{M} \frac{|(\chi_k, (H^{(1)} - E^{(1)})\psi^{(0)})|^2}{\epsilon_k - E^{(0)}} \qquad \text{(XVIII-7)}$$

These formulae of course have a close resemblance to the "sum over states" formulae for $\psi^{(1)}$ and $E^{(2)}$.

In the above discussion we have implicitly assumed that none of the ϵ_k is equal to $E^{(0)}$. If we would include $\psi^{(0)}$ in the basis set, say choose

$$\phi_1 = \psi^{(0)}$$

then one of the χ_k, call it χ_1, would also be $\psi^{(0)}$ and ϵ_1 would equal $E^{(0)}$. Under these circumstances (5) holds only for $k \neq 1$ while $\hat{b}_1^{(1)}$ can be chosen quite arbitrarily. In particular then we can take it equal to zero. Then since from (2), $(\chi_k, \psi^{(0)})$ will equal zero for $k \neq 1$ we find

$$\hat{\psi}^{(1)} = -\sum_{k=2}^{M} \chi_k \frac{(\chi_k, H^{(1)}\psi^{(0)})}{\epsilon_k - E^{(0)}} \qquad \text{(XVIII-8)}$$

$$\hat{E}^{(2)} = -\sum_{k=2}^{M} \frac{|(\chi_k, H^{(1)}\psi^{(0)})|^2}{\epsilon_k - E^{(0)}} \qquad \text{(XVIII-9)}$$

which have an even stronger resemblance to the "sum over states" formulae for $\psi^{(11)}$ and $E^{(12)}$.

XIX. The Hylleraas Variational Method for Excited States

In this section we wish to describe several circumstances in which $\hat{E}^{(12)}$ will be an upper bound to $E^{(12)}$ when $\psi^{(10)}$ is an excited state. The discussion will closely parallel that which we gave earlier of conditions under which $\hat{E}$ would be an upper bound to E for excited states.

First of all suppose that there is some symmetry property which is maintained in the presence of the perturbation. Then if $E^{(10)}$ is the lowest zero order eigenvalue associated with that symmetry, and if the $\tilde{\psi}^{(11)}$ all have that symmetry then it immediately follows from (XVI-10) that

$$\tilde{E}^{(12)} \geq E^{(12)}$$

since if we imagine $\tilde{\Delta}^{(11)}$ expanded in terms of the eigenfunctions of $H^{(10)}$, it can involve only states with energy greater than or equal to $E^{(10)}$.

In Sec. II, we also pointed out that more generally we would have $\hat{E} \geq E$ if $\hat{\psi}$ were orthogonal to all the lower states. However we dismissed this as being of no practical interest in the absence of symmetry, since one would not know the lower states exactly. However in the present context the analogous result becomes more interesting. The point is that one may well know the exact zero order wave functions for the lower states, and that knowing this it is then easy to enforce orthogonality to the exact lower states through first order. We will now show that if one has orthogonality through first order then $\hat{E}^{(2)} \geq E^{(2)}$.

Suppose then there are T zero order states with eigenfunctions $\psi_t^{(0)}$ and eigenvalues $E_t^{(0)}$ such that $E_t^{(0)} < E^{(0)}$ (more precisely if there is symmetry involved, then T is the number of states with $E_t^{(0)} < E^{(0)}$ and having the symmetry of $\psi^{(1)}$, however we will not continue to stress this point). We now require that $\hat{\psi}$ be orthogonal to ψ_t through first order. Then, since they are orthogonal in zero order, this means that

$$(\psi_t^{(0)}, \hat{\psi}^{(1)}) + (\psi_t^{(1)}, \psi^{(0)}) = 0 \qquad t = 1 \cdots T \tag{XIX-1}$$

But from the "sum over states" formulae

$$(\psi_t^{(1)}, \psi^{(0)}) = -\frac{(\psi_t^{(0)}, H^{(1)}\psi^{(0)})}{E^{(0)} - E_t^{(0)}}$$

whence from (1) we have

$$(\psi_t^{(0)}, \hat{\psi}^{(1)}) = -\frac{(\psi_t^{(0)}, H^{(1)}\psi^{(0)})}{E_t^{(0)} - E^{(0)}}$$

which in turn equals $(\psi_t^{(0)}, \psi^{(1)})$. Thus if (1) is satisfied then $\tilde{\Delta}^{(1)}$ will be orthogonal to all the $\psi_t^{(0)}$, and hence we will have, as announced, $\hat{E}^{(2)} \geq E^{(2)}$.

Now one way to ensure that (1) is satisfied is simply to use trial functions of the form

$$\tilde{\psi}^{(1)} = -\sum_{t=1}^{T} \psi_t^{(0)} \frac{(\psi_t^{(0)}, H^{(1)}\psi^{(0)})}{E_t^{(0)} - E^{(0)}} + \tilde{\chi}$$

where $\tilde{\chi}$ is orthogonal to the $\psi_t^{(0)}$. (This is the approach suggested by Sinanoglu, Phys. Rev. 122, 491 (1961)). A rather more flexible method (essentially that of Miller, J. Chem. Phys. 44, 2198 (1966)) is to use trial functions of the form

$$\tilde{\psi}^{(1)} = \sum_{t=1}^{T} \tilde{a}_t \psi_t^{(0)} + \tilde{Y}$$

where the $\tilde{a}_t$ are arbitrary variational parameters and where $\tilde{Y}$ can be anything except that it should not involve the $\tilde{a}_t$ explicitly. To see that this works, note that if we ask for the optimal $\tilde{a}_t$ for a given $\tilde{Y}$ then we will have

$$\delta \tilde{a}_t \left[(\psi_t^{(0)}, (H^{(0)} - E^{(0)}) \tilde{\psi}^{(1)}) + (\psi_t^{(0)}, (H^{(1)} - E^{(1)}) \psi^{(0)}) \right] = 0$$

or

$$(\psi_t^{(0)}, \tilde{\psi}^{(1)}) = - \frac{(\psi_t^{(0)}, H^{(1)} \psi^{(0)})}{E_t^{(0)} - E^{(0)}}$$

and hence, in particular, (1) will be satisfied, which proves the point. Note, however, that in general

$$\hat{a}_t \neq - \frac{(\psi_t^{(0)}, H^{(1)} \psi^{(0)})}{E_t^{(0)} - E^{(0)}}$$

unless $\hat{Y}$ is orthogonal to the $\psi_t^{(0)}$.

Finally let us turn to the analog of the variational method for excited states discussed in Sec. VI. Let us suppose that we have used the linear variational method of Sec. XVIII and have chosen M sufficiently large, and the ϕ_k sufficiently well, so that T of the ϵ_k are $< E^{(0)}$. We will now show that under these circumstances $\hat{E}^{(2)} \geq E^{(2)}$. To do this we will imitate the procedure of Sec. VI and add one more function to our basis set. Thus we consider trial functions of the form

$$\tilde{\psi}^{(1)} = \sum_{K=1}^{M} \tilde{C}_K^{(1)} \chi_K + \tilde{C}^{(1)} \phi \qquad \text{(XIX-2)}$$

where without loss of generality we may assume that ϕ is normalized and orthogonal to all the χ_K. It is now a simple matter to derive (Problem: Carry through the details) the result that the new $\hat{E}^{(2)}$, call it $\hat{E}^{(2)}(M+1)$, is given by

$$\hat{E}^{(2)}(M+1) = \hat{E}^{(2)}(M) - \frac{Z}{D(E^{(0)})} \qquad \text{(XIX-3)}$$

where $\hat{E}^{(2)}(M)$ is given by (XVIII-7), where

$$Z = \left| \sum_K \frac{(\psi^{(0)}, (H^{(1)} - E^{(1)})\chi_K)(\chi_K, (H^{(0)} - E^{(0)})\phi)}{E^{(0)} - \epsilon_K} + (\psi^{(0)}, (H^{(1)} - E^{(1)})\phi) \right|^2 \qquad \text{(XIX-4)}$$

is non-negative, and where

$$D(E^{(0)}) = \sum_K \frac{|(\phi, H^{(0)}\chi_K)|^2}{E^{(0)} - \epsilon_K} - [E^{(0)} - (\phi, H^{(0)}\phi)] \qquad \text{(XIX-5)}$$

$$\equiv F(E^{(0)}) - G(E^{(0)}) \qquad \text{(XIX-6)}$$

We will now show that $D(E^{(0)}) \geq 0$ whence we will have

$$\hat{E}^{(2)}(M) \geq \hat{E}^{(2)}(M+1)$$

and hence, assuming convergence, $\hat{E}^{(2)}(M) \geq E^{(2)}$ as announced. To show that $D(E^{(0)})$ is positive we note that from (VI-10) the roots of $D(\epsilon') = 0$ would just be the ϵ_K' that one would derive by using a trial function of the form (2) in the linear variational method as applied to $H = H^{(0)}$. Thus in particular it follows from the separation theorem (VI-11) that

$$\epsilon'_{T+1} \leq \epsilon_{T+1}$$

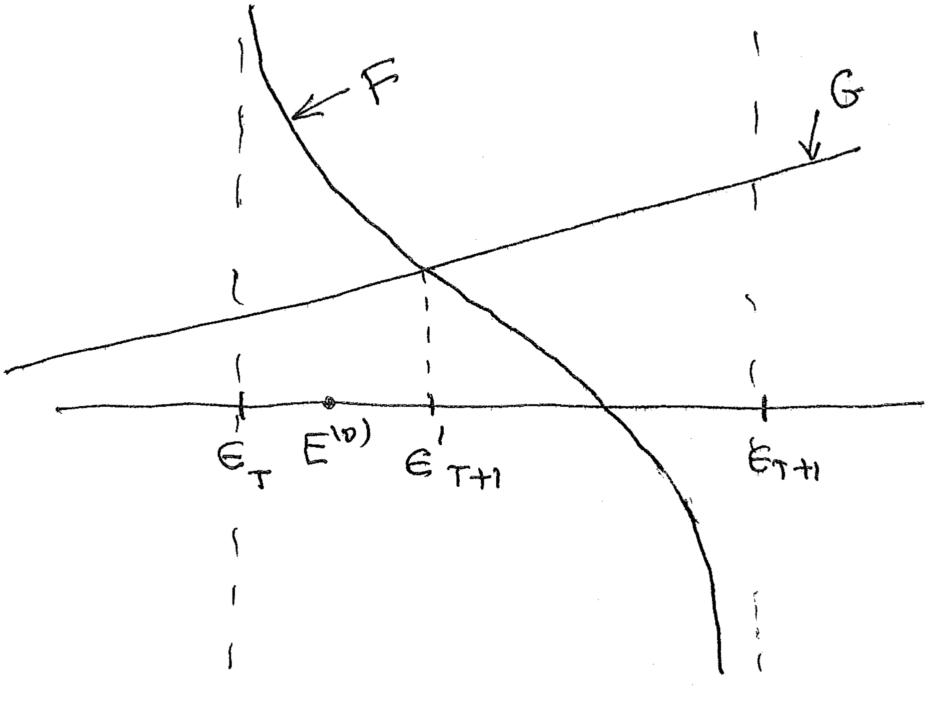
Further we certainly have

$$E^{(\nu)} \leq \epsilon'_{T+1}$$

and from our initial assumption we have

$$\epsilon_T < E^{(\nu)}$$

That $D(E^{(\nu)})$ is positive now follows simply by looking at a graph. Namely ϵ'_{T+1} is given by the appropriate intersection of $F(\epsilon)$ and the straight line $G(\epsilon)$:

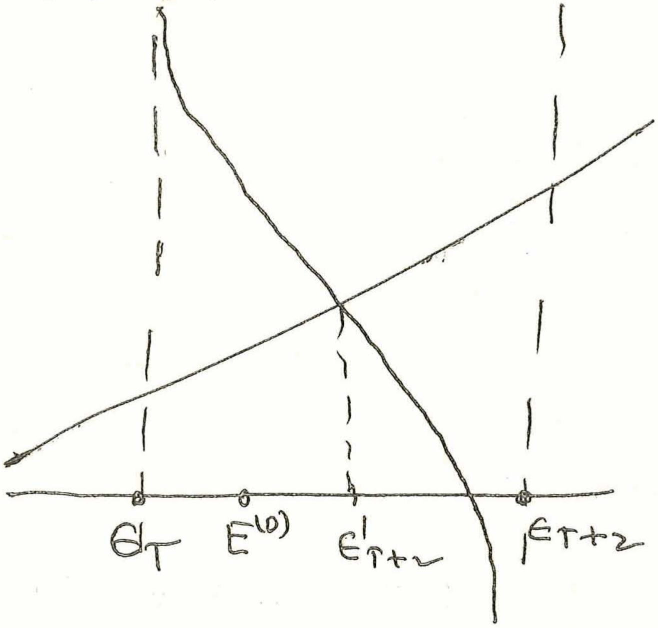


whence we see from the graph that

$$F(E^{(\nu)}) \geq G(E^{(\nu)})$$

which proves the point. Thus we have here an alternative method for getting upper bounds which does not involve use of the $\psi_t^{(\nu)}$. Of course, in accord with the discussion at the end of Sec. VI, one will get a better result if one does require (1). <u>Problem</u>: Prove this.

In this discussion we have implicitly assumed that none of the ϵ_k was equal to $E^{(0)}$. If, changing the notation slightly from Sec. XVIII, $\chi_{T+1} = \psi^{(0)}$ (for an excited state this is not the only way of having an $\epsilon_k = E^{(0)}$ however it is the most common one) then one finds results like (3) - (6) except that the term $k = T+1$ is excluded from the sums. (<u>Problem</u>: Carry through the details.) Comparing with (VI-10) we see that this corresponds to having, in the notation of that section, one of the $(\hat{\psi}^{(k)}, H^{(0)}\phi)$ equal to zero. However, the qualitative result is the same, the appropriate graph being



(<u>Problem</u>: Carry through the details.)

XX. Special Theorems Satisfied by Optimal First Order Hylleraas Variational Wave Functions

Our discussion here will be patterned after that in Sec. X. We are interested in (sufficient) conditions such that $\psi^{(0)} + v\hat{\psi}^{(1)}$ will satisfy various theorems through first (or second) order. The generalization of these results to higher orders we leave as an exercize for the reader.

(A) Generalized Hellmann-Feynman Theorem

If both $H^{(0)}$ and $vH^{(1)}$ contain the parameter σ, then clearly knowing only $\psi^{(0)}$ and $\psi^{(1)}$ we can calculate $\frac{\partial E}{\partial \sigma}$ accurately only through first order. Hence in this case the most we can expect is that $\psi^{(0)} + v\hat{\psi}^{(1)}$ will satisfy the Generalized Hellmann-Feynman theorem through first order. On the other hand if only $vH^{(1)}$ contains σ then we might hope to satisfy the theorem through second order.

Let us first consider the simplest case -- that in which $H^{(0)}$ is independent of σ. Then the generalized Hellmann-Feynman theorem reads

$$\frac{\partial \hat{E}}{\partial \sigma} = \left(\hat{\psi}, \frac{\partial}{\partial \sigma}(vH^{(1)})\hat{\psi}\right) \Big/ (\hat{\psi}, \hat{\psi}) \qquad \text{(XX-1)}$$

which when expanded out through second order yields

$$\frac{\partial E^{(0)}}{\partial \sigma} + \frac{\partial}{\partial \sigma}(vE^{(1)}) + \frac{\partial}{\partial \sigma}(v^2 \hat{E}^{(2)}) + \cdots$$
$$= \left(\psi^{(0)}, \frac{\partial}{\partial \sigma}(vH^{(1)})\psi^{(0)}\right) + v\left\{\left(\hat{\psi}^{(1)}, \left[\frac{\partial}{\partial \sigma}(vH^{(1)} - vE^{(1)})\right]\psi^{(0)}\right) + c.c.\right\}$$
$$+ \cdots$$

or, equating terms of the same order in ν we have

$$\frac{\partial E^{(0)}}{\partial\sigma} = 0 \tag{XX-2}$$

$$\frac{\partial}{\partial\sigma}(\nu E^{(1)}) = \left(\psi^{(0)}, \left[\frac{\partial}{\partial\sigma}(\nu H^{(1)})\right]\psi^{(0)}\right) \tag{XX-3}$$

$$\frac{\partial}{\partial\sigma}(\nu^2 \hat{E}^{(2)}) = \nu\left(\hat{\psi}^{(1)}, \left[\frac{\partial}{\partial\sigma}(\nu H^{(1)} - \nu E^{(1)})\right]\psi^{(0)}\right) + c.c. \tag{XX-4}$$

Since $E^{(0)}$ doesn't depend on σ eq. (2) is trivial and so is Eq. (3) since

$$\nu E^{(1)} = (\psi^{(0)}, \nu H^{(1)} \psi^{(0)})$$

and $\psi^{(0)}$ is independent of σ. (More accurately $\psi^{(0)}$ could involve σ only in a numerical phase factor.) Thus we are left with the question, when will (4) be satisfied? Well, we have

$$\nu^2 \hat{E}^{(2)} = \nu^2 (\hat{\psi}^{(1)}, (H^{(0)} - E^{(0)}) \hat{\psi}^{(1)}) + \nu^2 \{ (\hat{\psi}^{(1)}, (H^{(1)} - E^{(1)}) \psi^{(0)}) + c.c. \}$$

Now let us suppose that $\frac{\partial \hat{\psi}^{(1)}}{\partial\sigma}\delta\sigma$ is an allowed variation of $\hat{\psi}^{(1)}$. Then in calculating $\frac{\partial}{\partial\sigma}(\nu^2 \hat{E}^{(2)})$ we may ignore the terms involving $\frac{\partial \hat{\psi}^{(1)}}{\partial\sigma}$ whence we have

$$\frac{\partial}{\partial\sigma}(\nu^2 \hat{E}^{(2)}) = \frac{\partial \nu^2}{\partial\sigma} \hat{E}^{(2)} + \nu^2 \left\{ \left(\hat{\psi}^{(1)}, \left[\frac{\partial}{\partial\sigma}(H^{(1)} - E^{(1)})\right]\psi^{(0)}\right) + c.c. \right\}$$

If $\partial\nu/\partial\sigma = 0$ then this is the same as (4). If $\frac{\partial\nu}{\partial\sigma} \neq 0$ then to agree with (4) we also must require that

$$\hat{E}^{(2)} = \frac{1}{2}\left[(\hat{\psi}^{(1)}, (H^{(1)} - E^{(1)}) \psi^{(0)}) + c.c. \right] \tag{XX-5}$$

the conditions for which we discussed in Sec. XVII-(B). Thus if we select $\hat{\psi}^{(1)}$ from a set of trial functions which are independent of so that $\frac{\partial \hat{\psi}^{(1)}}{\partial \sigma}\delta\sigma$ is an allowed variation, and which have an arbitrary over all scale so that (5) is satisfied, then (4) will be satisfied automatically. (Again, if $\partial \nu/\partial\sigma = 0$ then we don't need (5).)

Problem: Discuss the content of the Generalized Hellmann-Feynman theorem for the case $\sigma = \nu$

Now let us consider the more general case in which both $H^{(0)}$ and $\nu H^{(1)}$ may involve . The Generalized Hellmann-Feynman theorem now reads

$$\frac{\partial \hat{E}}{\partial \sigma} = \left(\hat{\psi}, \left(\frac{\partial H^{(0)}}{\partial \sigma} + \frac{\partial}{\partial \sigma}(\nu H^{(1)})\right)\hat{\psi}\right) \Big/ (\hat{\psi}, \hat{\psi}) \qquad \text{(XX-6)}$$

Expanding (6) in power of ν and equating terms of like order we then find in particular

$$\frac{\partial \hat{E}^{(0)}}{\partial \sigma} = \left(\psi^{(0)}, \frac{\partial H^{(0)}}{\partial \sigma}\psi^{(0)}\right) \qquad \text{(XX-7)}$$

and

$$\frac{\partial}{\partial \sigma}(\nu \hat{E}^{(1)}) = \nu\left(\psi^{(0)}, \left[\frac{\partial}{\partial \sigma}(H^{(0)} - E^{(0)})\right]\hat{\psi}^{(1)}\right) + \text{c.c.} \qquad \text{(XX-8)}$$
$$+ \left(\psi^{(0)}, \left[\frac{\partial}{\partial \sigma}(\nu H^{(1)})\right]\psi^{(0)}\right)$$

Now (7) is certainly satisfied since $\psi^{(0)}$ is an eigenfunction of $H^{(0)}$. Thus our goal is to find sufficient conditions so that (8) is satisfied.

Before proceeding further it will be useful to rewrite (8) so as to eliminate the $\frac{\partial}{\partial\sigma}(H^{(0)} - E^{(0)})$. Now from

$$(H^{(0)} - E^{(0)})\psi^{(0)} = 0$$

it follows that

$$\left[\frac{\partial}{\partial\sigma}(H^{(0)} - E^{(0)})\right]\psi^{(0)} = -(H^{(0)} - E^{(0)})\frac{\partial\psi^{(0)}}{\partial\sigma}$$

whence we can replace (8) by

$$\frac{\partial}{\partial\sigma}(\nu E^{(1)}) = -\nu\left(\frac{\partial\psi^{(0)}}{\partial\sigma}, (H^{(0)} - E^{(0)})\tilde{\psi}^{(1)}\right) + c.c. \qquad \text{(XX-9)}$$

$$+ \left(\psi^{(0)}, \left[\frac{\partial}{\partial\sigma}(\nu H^{(1)})\right]\psi^{(0)}\right)$$

Now let us assume that $\frac{\partial\psi^{(0)}}{\partial\sigma}\delta\sigma$ is an allowed variation of $\tilde{\psi}^{(1)}$. Then from the Hylleraas variational principle we find

$$\left\{\left(\frac{\partial\psi^{(0)}}{\partial\sigma}, (H^{(0)} - E^{(0)})\tilde{\psi}^{(1)}\right) + \left(\frac{\partial\psi^{(0)}}{\partial\sigma}, H^{(1)}\psi^{(0)}\right)\right\} + c.c. = 0$$

where we have used $\frac{\partial}{\partial\sigma}(\psi^{(0)}, \psi^{(0)}) = 0$. If we insert this into (9) we have

$$\frac{\partial}{\partial\sigma}(\nu E^{(1)}) = \nu\left\{\left(\frac{\partial\psi^{(0)}}{\partial\sigma}, H^{(1)}\psi^{(0)}\right) + c.c.\right\} + \left(\psi^{(0)}, \frac{\partial}{\partial\sigma}(\nu H^{(1)})\psi^{(0)}\right)$$

or

$$\frac{\partial}{\partial\sigma}(\nu E^{(1)}) = \frac{\partial}{\partial\sigma}(\psi^{(0)}, \nu H^{(1)}\psi^{(0)}) \equiv \frac{\partial}{\partial\sigma}(\nu E^{(1)})$$

i.e., an identity and hence eq. (8) is satisfied. Now how might one ensure that $\delta\tilde{\psi}^{(1)} = \frac{\partial\psi^{(0)}}{\partial\sigma}\delta\sigma$ is an allowed variation? The one general way we know of is simply to use trial functions of the form

$$\tilde{\psi}^{(1)} = \tilde{a}\frac{\partial\psi^{(0)}}{\partial\sigma} + \tilde{X}$$

where $\tilde{\alpha}$ is a variational parameter and $\tilde{\chi}$ is anything which does not depend explicitly on $\tilde{\alpha}$.

(B) Virial and Hypervirial Theorems

For simplicity let us assume that g is independent of ν. Then expanding the hypervirial theorem

$$(\hat{\psi}, [g, H]\hat{\psi}) = 0$$

the zero order equation is of course trivially satisfied whence we are led to consider the first order equation

$$(\psi^{(0)}, [g, H^{(1)}]\psi^{(0)}) + \{(\psi^{(0)}, [g, H^{(0)}]\psi^{(1)}) + (\hat{\psi}^{(1)}, [g, H^{(0)}]\psi^{(0)})\} = 0 \qquad \text{(XX-10)}$$

One now readily finds that if

$$\Delta\hat{\psi}^{(1)} = i\,\delta\hat{\eta}\; g\psi^{(0)} \qquad \text{(XX-11)}$$

with $\delta\hat{\eta}$ a real parameter, is an allowed variation, then (7) will be satisifed. (Problem: Prove this.) A general way of ensuring that (8) is allowed is to introduce it via a linear variational parameter, i.e., to use

$$\tilde{\psi}^{(1)} = i\,\tilde{\alpha}\, g\psi^{(0)} + \tilde{\chi}$$

Note that this is just the first order version of (XI-14). Problem: What theorem is satisfied if no reality conditions are imposed on $\tilde{\alpha}$?
Problem : Discuss the situation if $g = g^{(0)} + \nu g^{(1)}$

(C) Orthogonality

Here our question is when are the approximate functions $\psi^{(0)} + \nu\hat{\psi}^{(1)}$ and $\chi^{(0)} + \nu\hat{\chi}^{(1)}$ orthogonal through first order, given that $\psi^{(0)}$ and

$\chi^{(0)}$ are two eigenfunctions of $H^{(0)}$ belonging to different eigenvalues, and hence orthogonal. Thus the question is

$$(\chi^{(0)}, \hat{\psi}^{(1)}) + (\hat{\chi}^{(1)}, \psi^{(0)}) \overset{?}{=} 0 \tag{XX-12}$$

Let us now assume that (XVII-1) and its analog for $\hat{\chi}^{(1)}$ apply. That is, we have

$$(\delta\hat{\psi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\psi}^{(1)} + (H^{(1)} - E^{(1)})\psi^{(0)}) = 0 \tag{XX-13}$$

$$(\delta\hat{\chi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\chi}^{(1)} + (H^{(1)} - E^{(1)})\chi^{(0)}) = 0 \tag{XX-14}$$

Now let us suppose that

$$\Delta\hat{\psi}^{(1)} = \delta\eta\, \chi^{(0)}$$

and

$$\Delta\hat{\chi}^{(1)} = \delta\eta\, \psi^{(0)} \tag{XX-15}$$

are allowed variations of $\hat{\psi}^{(1)}$ and $\hat{\chi}^{(1)}$ respectively. Then it follows from (13) and the complex conjugate of (14) that

$$(\chi^{(0)}, (H^{(0)} - E^{(0)})\hat{\psi}^{(1)}) + (\chi^{(0)}, (H^{(1)} - E^{(1)})\psi^{(0)}) = 0$$

and (XX-16)

$$(\hat{\chi}^{(1)}, (H^{(0)} - E^{(0)})\psi^{(0)}) + (\chi^{(0)}, (H^{(1)} - E^{(1)})\psi^{(0)}) = 0$$

Subtracting and using $(H^{(0)} - E^{(0)})\psi^{(0)} = 0$, $(H^{(0)} - E^{(0)})\chi^{(0)} = 0$, and $(\chi^{(0)}, \psi^{(0)}) = 0$ we then find (9). Note also that from (16) it follows that

$$(\chi^{(0)}, \hat{\psi}^{(1)}) = \frac{(\chi^{(0)}, H^{(1)}\psi^{(0)})}{E^{(0)}-\mathcal{E}^{(0)}} = -(\chi^{(1)}, \psi^{(0)})$$

and similarly

$$(\hat{\chi}^{(1)}, \psi^{(0)}) = -(\chi^{(0)}, \psi^{(1)})$$

i.e., referring to the discussion in Sec. XIX, we see that $\hat{\psi}$ is orthogonal to the exact χ through first order, and likewise $\hat{\chi}$ is orthogonal to ψ through first order. Thus, if $E^{(0)}$ is the ground state and $\mathcal{E}^{(0)}$ the first excited state it will also follow that $\hat{\mathcal{E}}^{(2)} \geq \mathcal{E}^{(2)}$. Further, as was the case in Sec. XIX, the general way to ensure that the variations (15) are allowed is introduce $\chi^{(0)}$ $[\psi^{(0)}]$ into $\hat{\psi}^{(1)}$ $[\hat{\chi}^{(1)}]$ use of a linear parameter.

(E) <u>Gauge Invariance</u>

The conditions on the $\tilde{\psi}^{(1)}$ which will ensure the gauge invariance of the energy through second order can be inferred directly from the discussion in Sec. XI-(E). However it may be worthwhile to furnish some further detail since in cases of interest H is usually not of the simple form $H^{(0)}+\nu H^{(1)}$ which we have been assuming. Thus consider a system in a magnetic field. Then H typically has the form

$$H = H^{(0)} + \nu H^{(1)} + \nu^2 H^{(2)}$$

where $\nu H^{(1)}$ is linear in the field and $\nu^2 H^{(2)}$ quadratic. Thus from the discussion in Sec. XI-(E) it follows, since after a gauge transformation $U = e^{\frac{ie}{c}\Sigma\Lambda} \equiv e^{i\Gamma}$

$$H' = H^{(0)} + \nu H^{(1)'} + \nu^2 H^{(2)'}$$

that (Γ is of order ν).

$$\nu H^{(1)'} = \nu H^{(1)} + i\,[\Gamma, H^{(0)}] \tag{XX-17}$$

$$\nu^2 H^{(2)'} = \nu^2 H^{(2)} + i\,[\Gamma, \nu H^{(1)}] - \frac{1}{2}[\Gamma, [\Gamma, H^{(0)}]] \tag{XX-18}$$

and that the higher order terms which would result from the expansion of $e^{i\Gamma}$ simply vanish.

From (7) it then follows that $E^{(1)'} = E^{(1)}$ for any Γ since $(\psi^{(0)}, [\Gamma, H^{(0)}]\,\psi^{(0)}) = 0$, which expresses complete gauge invariance through first order. Then, in accord with the fact that if $H\psi = E\psi$ and $H'\psi' = E\psi'$ then $\psi' = e^{i\Gamma}\psi$, one can readily show (<u>Problem</u>: Do this) that if we define

$$\tilde{\psi}^{(1)'} = \tilde{\psi}^{(1)} + i\Gamma\psi^{(0)} \tag{XX-19}$$

then

$$\tilde{J}_H^{(2)'} + (\psi^{(0)}, H^{(2)'}\psi^{(0)}) = \tilde{J}_H^{(2)} + (\psi^{(0)}, H^{(2)}\psi^{(0)}) \tag{XX-20}$$

where $\tilde{J}_H^{(2)'}$ is the Hylleraas Functional for $H^{(0)} + \nu H^{(1)'}$, using $\tilde{\psi}^{(1)'}$ as the trial function. Now let us suppose that the set of trial functions which we propose to use for both H and H' is closed under (19) for certain Γ, i.e., that if $\tilde{\psi}^{(1)}$ is in the set then so is $\tilde{\psi}^{(1)'}$. (This is clearly just the first order version of our requirement in Sec. XI-(E).) Then from (20), since given the gauges, $(\psi^{(0)}, H^{(2)'}\psi^{(0)})$ and $(\psi^{(0)}, H^{(2)}\psi^{(0)})$ are <u>fixed</u> numbers independent of the trial functions, it follows that if $\tilde{\psi}^{(1)}$ makes $\delta\tilde{J}_H^{(2)} = 0$ then $\tilde{\psi}^{(1)'}$ will make $\delta\tilde{J}_H^{(2)'} = 0$, whence, using (20) again, we will have

$$\hat{J}_H^{(2)'} + (\psi^{(0)}, H^{(2)'}\psi^{(0)}) = \hat{J}_H^{(2)} + (\psi^{(0)}, H^{(2)}\psi^{(0)})$$

But the quantities on each side are the total second order energies (paramagnetic plus diamagnetic) in each gauge. Therefore under these conditions the energy will be invariant through second order, to the transformation generated by the particular Γ in question.

Problem: More generally discuss invariance to arbitrary unitary transformations.

XXI. The Delves Variational Method

The techniques of double perturbation theory were introduced to deal with the following situation. We are given a (normalized) function $\psi^{(00)}$ which in some sense approximates the exact (normalized) eigenfunction $\psi^{(0)}$ of an isolated system, whose Hamiltonian is $H^{(0)}$, and we would like to improve on the approximation to $(\psi^{(0)}, H^{(1)}\psi^{(0)})$ offered by $(\psi^{(00)}, H^{(1)}\psi^{(00)})$, or we would like to improve on a "polarizibility", or what all, where $H^{(1)}$ is some operator. Then double perturbation theory (or the interchange theorem) offers the following prescription for calculating first order corrections: Find a Hamiltonian such that

$$(H^{(00)} - E^{(00)})\psi^{(00)} = 0 \qquad \text{(XXI-1)}$$

where $E^{(00)}$ is a real number (and such that $E^{(00)}$ is non-degenerate). Then determine a function $\psi^{(01)}$ from,

$$(H^{(00)} - E^{(00)})\psi^{(01)} + (H^{(1)} - E^{(01)})\psi^{(00)} = 0 \qquad \text{(XXI-2)}$$

where

$$E^{(01)} = (\psi^{(00)}, H^{(1)}\psi^{(00)}) \tag{XXI-3}$$

One then shows (interchange theorem) that

$$\lambda E^{(11)} = \lambda(\psi^{(00)}, (V - E^{(10)})\psi^{(00)}) + c.c. \tag{XXI-4}$$

where

$$\lambda V = H^{(0)} - H^{(00)} \tag{XXI-5}$$

and

$$E^{(10)} = (\psi^{(00)}, V\psi^{(00)}) \tag{XXI-6}$$

yields a first order (in the parameter λ) correction to $E^{(01)}$. That is one shows that

$$\lambda E^{(11)} = \lambda(\psi^{(00)}, (H^{(1)} - E^{(01)})\psi^{(10)}) + c.c. \tag{XXI-7}$$

where

$$(H^{(00)} - E^{(00)})\psi^{(10)} + (V - E^{(10)})\psi^{(00)} = 0 \tag{XXI-8}$$

which clearly is the first order correction.

The advantage of the formulation (4) over (7) is that $\psi^{(01)}$ may be easier to deal with than $\psi^{(10)}$, i.e., $H^{(1)}$ may be a simpler perturbation than V. In either case one can use variational techniques to approximate their solution, and we will comment on this later.

The point we want to make now is that this approach, either in the form (4) or (7) hinges on choosing an $H^{(00)}$. Now in some situations there may be a natural choice, however usually there will not be, and since the answers certainly depend on the choice of $H^{(00)}$ it is not really clear what to do.

We now note that the variational method suggests another approach which seems to avoid the problem of choosing $H^{(00)}$. This approach has been most developed by Delves. (See Proc. Phys. Soc. 92, 55 (1967) for a recent discussion and references to his earlier papers. Similar methods have been suggested by Deal, et. al. J. Chem. Phys. 49, 3395 (1968).)

Consider the Hamiltonian

$$H = H^{(0)} + \nu H^{(1)}$$

Then its eigenvalue E is the quantity of interest to us since $E^{(1)} = (\psi^{(0)}, H^{(1)} \psi^{(0)})$, $E^{(2)}$ yields the polarizibility etc. Now suppose that from someplace one produces a function ϕ depending on ν, which is such that for $\nu = 0$, ϕ becomes $\psi^{(00)}$, i.e., $\phi^{(0)} = \psi^{(00)}$ Now whatever ϕ is we know that $\mathcal{E}$ defined by

$$\mathcal{E} = (\phi, (H^{(0)} + \nu H^{(1)})\phi) / (\phi, \phi) \qquad \text{(XXI-9)}$$

will differ from E by terms of the order of the square of the error in ϕ. In general ϕ will be in error, i.e. differ from ψ the exact eigenfunction of H, by terms of order ν and by terms of order ξ where ξ is some measure of $\psi^{(00)} - \psi^{(0)}$. Thus $\mathcal{E}$ will differ from E by terms of order ν^2, λ^2 and $\lambda\nu$. But now suppose that ϕ is such that if ξ were equal to zero then ϕ would be exact. Under these circumstances ϕ will be in error only by terms of order ξ whence $E - \mathcal{E}$, and hence the $E^{(n)} - \mathcal{E}^{(n)}$, will be only of order ξ^2. A case in point is URHF for an atom.

Then if there are no degeneracies for $Z \to \infty$ we can identify ξ as $1/Z$. Further it is clear that under these circumstances URHF is exact when $Z \to \infty$ if $H^{(1)}$ is any one electron operator. Thus we have rederived, in a very direct way, the result we found in Sec. XI-(D) by use of Brillouin's theorem.

Now in double perturbation theory ξ is identified with the parameter λ introduced in (5) and one ensures exactness in the $\lambda = 0$ limit simply by choosing $\phi = \psi^{(0,)}$ where $\psi^{(0,)}$ satisfies the exact $\lambda = 0$ equation, namely

$$(H^{(00)} + \nu H^{(1)})\psi^{(0,)} = E^{(0,)}\psi^{(0,)} \qquad \text{(XXI-10)}$$

Thus in particular writing $\psi^{(0,)} = \sum \nu^n \psi^{(0,n)}$ and $E^{(0,)} = \sum \nu^n E^{(0,n)}$, the first order equation is just (2) again. Further, whatever ϕ is, we have from (9), writing $\phi = \sum \nu^n \phi^{(n)}$ and $\epsilon = \sum \nu^n \epsilon^{(n)}$ that

$$\epsilon^{(0)} = (\psi^{(00)}, H^{(10)}\psi^{(00)}) \qquad \text{(XXI-11)}$$

$$\epsilon^{(1)} = E^{(01)} + \{ (\psi^{(00)}, (H^{(10)} - \epsilon^{(0)})\phi^{(1)}) + c.c. \} \qquad \text{(XXI-12)}$$

etc. and one readily finds (<u>Problem</u>: Fill in the details) that with $\phi^{(1)} = \psi^{(0,1)}$ then $\epsilon^{(1)} - E^{(01)} = \lambda E^{(11)}$ as given in (4).

However now suppose that instead of (2) we consider the equation

$$(H^{(10)} - \epsilon^{(0)})\phi_D^{(1)} + (H^{(1)} - E^{(01)})\psi^{(00)} = 0 \qquad \text{(XXI-13)}$$

Presumably we can't solve it exactly, however let us ignore this little difficulty for the moment. Then the point is that clearly $\Phi_D^{(1)}$ will equal $\psi^{(1)}$ if $\xi \to 0$, and hence, if we would insert $\Phi^{(1)} = \Phi_D^{(1)}$ in (12) we would equally well have an estimate of $E^{(1)}$ which is in error only by terms of order ξ^2.

Now the virtue of (10) is of course that it doesn't involve an $H^{(00)}$, but also, as we have noted, one can't solve it. However, except in special cases (the $1/Z$ expansion or the shielding approximation for example) this will usually also be the case with (2) so this is not a special drawback of the proposed method. In either case one can use variational techniques. For (2) this would simply be the Hylleraas approach with $H^{(00)}$ playing the role which $H^{(0)}$ played in our earlier discussions and all those discussions will apply. For (13) however the situation is not quite the same. First of all $\epsilon^{(10)}$ is not equal to $E^{(10)}$. This has the following consequence: $E^{(10)} - \epsilon^{(10)}$ is of order ξ^2 while $(\psi^{(10)}, (H^{(1)} - E^{(01)})\psi^{(00)})$ is of order ξ. This then means that $(\psi^{(10)}, \Phi_D^{(1)})$, if it doesn't vanish for reasons of symmetry, will be of order $1/\xi$. Such a behavior would seem undesireable. Hence we replace (10) by

$$(H^{(10)} - \epsilon^{(10)})\,\Phi_D^{(1)} + (H^{(1)} - \zeta)\,\psi^{(00)} = 0 \qquad \text{(XXI-14)}$$

where ζ is a number to be chosen so that $(\psi^{(00)}, \Phi_D^{(1)}) = 0$ whence there will be no problems as $\xi \to 0$. Since ζ will differ from $E^{(01)}$ only by terms of order ξ, (14) will still be exact in the $\xi \to 0$ limit.

A further difficulty is the following. Suppose that $\psi^{(00)}$ is such a poor approximation to $\psi^{(10)}$ that $\epsilon^{(10)}$ lies in the continuum of $H^{(0)}$ (such a situation is not at all unusual. See Dougherty, et. al. J Chem. Phys. 45, 1803 (1966)). Then the homogenous equation has a solution and the inhomogenity will usually not be orthogonal to it, so that one has an inconsistent equation. This difficulty can also be avoided, only rather artificially. In practice this difficulty usually doesn't appear explicitly because one usually couples a poor $\psi^{(00)}$ with an equally poor "solution" to (12).

As we have said, the natural way to attack (12) is with a variational approach. Evidently the appropriate functional is

$$\tilde{\hat{D}} = (\tilde{\phi}^{(1)}, (H^{(0)} - \epsilon^{(0)})\tilde{\phi}^{(1)}) + (\tilde{\phi}^{(1)}, (H^{(1)} - \hat{\xi})\psi^{(00)}) + (\psi^{(00)}, (H^{(1)} - \hat{\xi})\tilde{\phi}^{(1)}) \tag{XXI-15}$$

where $\hat{\xi}$ is to be chosen so that $(\psi^{(00)}, \hat{\phi}^{(1)}) = 0$. One could of course also simply restrict the set of trial functions to functions orthogonal to $\psi^{(00)}$ and replace (15) by

$$\tilde{\hat{D}} = (\tilde{\phi}^{(1)}, (H^{(0)} - \epsilon^{(0)})\tilde{\phi}^{(1)}) + (\tilde{\phi}^{(1)}, H^{(1)}\psi^{(00)}) + (\psi^{(00)}, H^{(1)}\hat{\phi}^{(1)}) \tag{XXI-16}$$

To change the subject somewhat for a moment, it is now of interest to note that, as we remarked at the end of Sec. XVI, the functional, (15), though often with $\hat{\xi}$ replaced by $E^{(01)}$, is frequently used to approximate the second order energy, i.e. $\hat{D}$ is used as an approximation to $E^{(2)}$.

The following are some properties of this approximation. In the main we state them without proof (see J. Chem. Phys. 42, 3630 (1965)).

(1) With $\tilde{\phi}^{(1)} = \phi^{(1)}_D$ the value of $\tilde{D}$ is

$$D = (\psi^{(00)}, H^{(1)} \phi^{(1)}_D) \tag{XXI-17}$$

One can then show that $D - E^{(12)}$ is of first order in ξ and can have either sign. (If one uses $E^{(00)}$ instead of ξ then $D - E^{(12)}$ can even be of zero order in ξ !)

(2) As noted above, if $\psi^{(00)}$ is so poor that (12) actually is an inconsistent equation, then usually one is not aware of this, and may even find a reasonable result for $\hat{D}$, because one will probably use a limited and "reasonable" set of $\tilde{\phi}^{(1)}$.

(3) Writing

$$\hat{\phi}^{(1)} - \phi^{(1)}_D \equiv \Delta$$

then clearly

$$\hat{D} - D = (\Delta, (H^{(0)} - E^{(0)}) \Delta)$$

If now symmetry considerations apply, and if $E^{(0)}$ is less than the energy of first state of appropriate symmetry then evidently

$$\hat{D} - D > 0$$

If one does not have symmetry, then one can still show that if $\psi^{(0)}$ is the ground state, then $\hat{D}$ will be an upper bound to D if ξ is small enough.

Returning now to double perturbation theory, the approximation of the solution of eq. (2) [eq. (8)] variationally, followed by use of eq. (4) with $\psi^{(01)}$ replaced by $\hat{\psi}^{(01)}$ [eq. (7) with $\psi^{(10)}$ replaced by $\hat{\psi}^{(10)}$] to approximate $E^{(11)}$ (for an application see Chen and Dalgarno, Proc.

Phys. Soc. 855, 399 (1965). For $H^{(00)}$ they use the Sternheimer Hamiltonian) is of course just the same sort of procedure which we discussed in Sec. XVII-(E) part (ii) [part (iii)] in a different connection. Here, as there, then, there is another alternative which deals with $E^{(11)}$ directly. Namely use of the functional (XVII-12), which in our present notation becomes

$$\left\{ \left(\tilde{\psi}^{(10)}, (H^{(00)} - E^{(00)})\, \tilde{\psi}^{(01)}\right) + \left(\tilde{\psi}^{(10)}, (H^{(01)} - E^{(01)})\, \psi^{(00)}\right) + \left(\psi^{(00)}, (V - E^{(10)})\, \tilde{\psi}^{(01)}\right) \right\} + c.c.$$

and we can simply apply the discussion which we gave there to the present situation.

As we have noted, the use of $\hat{D}$ to approximate $E^{(2)}$, c.f. to approximate polarizibilities, involves a first order error in ξ. One can obviously however get an approximation which involves only a second order error either through the use of $\psi^{(02)}$, i.e., via double perturbation theory, or by the use of a $\phi_D^{(2)}$ defined by an equation like

$$(H - E^{(00)})\, \phi_D^{(2)} + (H^{(1)} - E^{(1)})\, \phi_D^{(1)} + \xi\, \psi^{(00)} = 0$$

and their associated variational principles, etc. We leave the details to the interested reader.

Having come so far we must now point out that in fact the Delves method is just a special case of double perturbation theory, i.e. involves a special (implicit) choice of $H^{(00)}$! Let P be the projection operator onto $\psi^{(00)}$, i.e.e, in Dirac notation

$$P = |\psi^{(00)}\rangle\langle\psi^{(00)}|$$

and let us take for $H^{(00)}$ the "partitioned" form

$$H^{(00)} = P H^{(0)} P + (1-P) H^{(0)} (1-P) \qquad \text{(XXI-18)}$$

Then evidently $\psi^{(00)}$ is an eigenfunction of $H^{(00)}$ and the eigenvalue is just $\epsilon^{(0)}$ (Problem: Prove this). Consider now the associated Hylleraas functional

$$(\tilde{\Psi}^{(01)}, (H^{(00)} - E^{(00)}) \tilde{\Psi}^{(01)}) + (\tilde{\Psi}^{(01)}, (H^{(1)} - E^{(01)}) \psi^{(00)}) \qquad \text{(XXI-19)}$$

$$+ (\psi^{(00)}, (H^{(1)} - E^{(01)}) \tilde{\Psi}^{(01)})$$

Now suppose we restrict the $\tilde{\Psi}^{(01)}$ to be orthogonal to $\psi^{(00)}$ [the value of $(\psi^{(01)}, \psi^{(00)})$ is anyway undetermined and irrelevant]. Then we may drop the $E^{(01)}$ terms in (19) and further since $P\tilde{\Psi}^{(1)} = 0$ and $(1-P)\tilde{\Psi}^{(1)} = \tilde{\Psi}^{(1)}$ we may replace $H^{(00)}$ in (19) by H. But then we see that we have $\tilde{D}$ in the form (16) in a slightly different notation, i.e. $\tilde{\Psi}^{(01)}$ instead of $\tilde{\Phi}^{(1)}$. Thus one may say that the Delves procedure, with the orthogonality requirement, is equivalent to double perturbation theory with $H^{(00)}$ being the particular operator (18). A natural question at this point is then, is this $H^{(00)}$ in some sense a best choice? We will not pursue this question here

The choice (18) has been made by Aranoff and Percus Phys. Rev. 166, 1255 (1968) who also note that (19) can be replaced by (16). (18) has also been considered by Adams, J. Chem. Phys. 45, 3422 (1966), in another connection. Evidently this $H^{(00)}$ is not a simple operator. In particular it is not obvious what its other eigenfunctions are. However, referring to our earlier discussion, we can now say for what it is worth that if $\epsilon^{(0)}$ is the lowest eigenvalue of this $H^{(00)}$, then we will have $\hat{D} \geq D$.

The procedure which we have called the Delves procedure is only one of several that he has proposed (see the reference given earlier). However, it seems to be about the best of the lot. Some of his earlier efforts were essentially based on the observation that $E^{(n)} - E^{(01)}$ is in a somewhat different notation ($\psi^{(00)}$ instead of $\hat{\phi}^{(0)}$, $\phi^{(n)}$ instead of $\tilde{\phi}^{(n)}$), just the $\tilde{\mathcal{J}}^{(n)}$ that one derives from (XIII-20). In view of the restrictions on the $\tilde{\phi}^{(n)}$ which we discussed there, it is not surprising that trying to use $\delta\tilde{\mathcal{J}}^{(n)} = 0$ to determine a "best" $\phi^{(n)}$ did not prove very satisfactory.