THEORETICAL CHEMISTRY INSTITUTE

THE UNIVERSITY OF WISCONSIN

VARIATIONAL METHODS AND DEGENERATE PERTURBATION THEORY

by

Saul T. Epstein

WIS-TCI-143

29 December 1965

MADISON, WISCONSIN

VARIATIONAL METHODS AND DEGENERATE PERTURBATION THEORY

by

Saul T. Epstein*

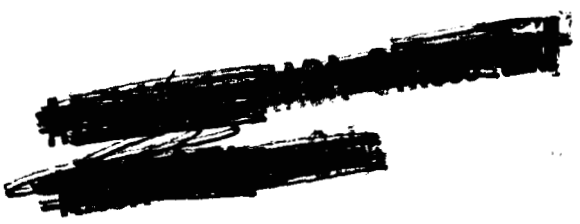
Theoretical Chemistry Institute and Department of Physics
University of Wisconsin, Madison, Wisconsin

Abstract

Variational methods related to degenerate perturbation theory are discussed in a general way through second order in the energy and first order in the wave function.

Author

* This research was supported by the following grant:
National Aeronautics and Space Administration Grant NsG-275-62.



I. Introduction

Variational methods have been much used and studied as a means of approximately calculating the effects of a perturbation on a non-degenerate eigenvalue[1]. In this note we wish to make a start on extending the discussion to the degenerate case by trying to get an idea of the types of problems which will be encountered. Specifically we will consider situations in which the degeneracy is broken in either first or second order.

It is a prime characteristic of degenerate perturbation theory that a priori we do not know the correct zero order wave function. Thus the appropriate tools to use are the variational principles given on page 299 of reference 1 which arise from direct expansion of the variation principle for the total energy, and which allow one to vary the zero order function.

Let $\mathcal{H}_0$ be the zero order Hamiltonian, $\mathcal{H}_1$, the perturbation and $\tilde{\Psi}^{(0)}$, $\tilde{\Psi}^{(1)}$, etc. variational trial functions which are to approximate the zero order wave function, first order correction, etc.

Now define $J^{(0)}$, $J^{(1)}$, $J^{(2)}$ etc. by

$$J^{(0)} \equiv \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_0 - \tilde{E}^{(0)}) \tilde{\Psi}^{(0)} \rangle \tag{1}$$

$$J^{(1)} \equiv \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)}) \tilde{\Psi}^{(0)} \rangle + \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_0 - \tilde{E}^{(0)}) \tilde{\Psi}^{(1)} \rangle + \langle \tilde{\Psi}^{(1)}, (\mathcal{H}_0 - \tilde{E}^{(0)}) \tilde{\Psi}^{(0)} \rangle \tag{2}$$

$$J^{(2)} \equiv \langle \tilde{\Psi}^{(0)}, (-\tilde{E}^{(2)})\tilde{\Psi}^{(0)}\rangle + \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\Psi}^{(1)}\rangle$$
$$+ \langle \tilde{\Psi}^{(1)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\Psi}^{(0)}\rangle + \langle \tilde{\Psi}^{(1)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(1)}\rangle \quad (3)$$
$$+ \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(2)}\rangle + \langle \tilde{\Psi}^{(2)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(0)}\rangle$$

$$J^{(3)} \equiv \langle \tilde{\Psi}^{(0)}, (-\tilde{E}^{(3)})\tilde{\Psi}^{(0)}\rangle + \langle \tilde{\Psi}^{(0)}, (-\tilde{E}^{(2)})\tilde{\Psi}^{(1)}\rangle$$
$$+ \langle \tilde{\Psi}^{(1)}, (-\tilde{E}^{(2)})\tilde{\Psi}^{(0)}\rangle + \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\Psi}^{(2)}\rangle$$
$$+ \langle \tilde{\Psi}^{(2)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\Psi}^{(0)}\rangle + \langle \tilde{\Psi}^{(1)}, (\mathcal{H}_1 - \tilde{E}^{(0)})\tilde{\Psi}^{(1)}\rangle \quad (4)$$
$$+ \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(3)}\rangle + \langle \tilde{\Psi}^{(3)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(0)}\rangle$$
$$+ \langle \tilde{\Psi}^{(1)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(2)}\rangle + \langle \tilde{\Psi}^{(2)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(1)}\rangle$$

$$J^{(4)} \equiv \langle \tilde{\Psi}^{(0)}, (-\tilde{E}^{(4)})\tilde{\Psi}^{(0)}\rangle + \langle \tilde{\Psi}^{(0)}, (-\tilde{E}^{(3)})\tilde{\Psi}^{(1)}\rangle + \langle \tilde{\Psi}^{(1)}, (-\tilde{E}^{(3)})\tilde{\Psi}^{(0)}\rangle$$
$$+ \langle \tilde{\Psi}^{(0)}, (-\tilde{E}^{(2)})\tilde{\Psi}^{(2)}\rangle + \langle \tilde{\Psi}^{(2)}, (-\tilde{E}^{(2)})\tilde{\Psi}^{(0)}\rangle$$
$$+ \langle \tilde{\Psi}^{(1)}, (-\tilde{E}^{(2)})\tilde{\Psi}^{(1)}\rangle + \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\Psi}^{(3)}\rangle$$
$$+ \langle \tilde{\Psi}^{(3)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\Psi}^{(0)}\rangle + \langle \tilde{\Psi}^{(1)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\Psi}^{(2)}\rangle \quad (5)$$
$$+ \langle \tilde{\Psi}^{(2)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\Psi}^{(1)}\rangle + \langle \tilde{\Psi}^{(0)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(4)}\rangle$$
$$+ \langle \tilde{\Psi}^{(4)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(0)}\rangle + \langle \tilde{\Psi}^{(1)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(3)}\rangle$$
$$+ \langle \tilde{\Psi}^{(3)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(1)}\rangle + \langle \tilde{\Psi}^{(2)}, (\mathcal{H}_0 - \tilde{E}^{(0)})\tilde{\Psi}^{(2)}\rangle$$

From the discussion in reference 1 it then follows that the $J^{(\alpha)}$ are to be made stationary with respect to variations of the $\tilde{\psi}^{(\beta)}$, the constants $\tilde{E}^{(\alpha)}$ being held fixed. The latter are then to be determined from $J^{(\alpha)}=0$, and provide variational approximations to the zero order energy, first order energy correction, etc.

Now although we don't know the correct $\psi^{(0)}$ we will assume that we do know a complete set of degenerate eigenfunctions of $\mathcal{H}_0$ belonging to the zero order eigenvalue in question. Denoting these eigenfunctions by χ_α where $\alpha = 1, 2, \cdots D$, D being the extent of the degeneracy, and the common eigenvalue by $E^{(0)}$, then the only $\tilde{\psi}^{(0)}$ we will consider will be of the form

$$\tilde{\psi}^{(0)} = \sum_{\alpha=1}^{D} \chi_\alpha C_\alpha \tag{6}$$

where the C_α are constants to be determined.

Then from $J^{(0)}=0$ we have $\tilde{E}^{(0)} = E^{(0)}$ i.e. $\tilde{E}^{(0)}$ <u>is exact</u>. Further we clearly have that not only is

$$(\mathcal{H}_0 - E^{(0)})\,\tilde{\psi}^{(0)} = 0 \tag{7}$$

but, if $\delta\tilde{\psi}^{(0)}$ is any variation of $\tilde{\psi}^{(0)}$ then also

$$(\mathcal{H}_0 - E^{(0)})\,\delta\tilde{\psi}^{(0)} = 0 \tag{8}$$

From these observations it then follows that we may dispense with $J^{(0)}$, and replace $J^{(1)}$, $J^{(2)}$, $J^{(3)}$ and $J^{(4)}$ by the simpler expressions

$$J^{(1)\prime} \equiv \langle \tilde{\psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(0)} \rangle \tag{9}$$

$$\begin{aligned} J^{(2)\prime} \equiv\ & \langle \tilde{\psi}^{(0)}, (-\tilde{E}^{(2)})\tilde{\psi}^{(0)} \rangle + \langle \tilde{\psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(1)} \rangle \\ & + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(0)} \rangle + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_0 - E^{(0)})\tilde{\psi}^{(1)} \rangle \end{aligned} \tag{10}$$

$$\begin{aligned} J^{(3)\prime} \equiv\ & \langle \tilde{\psi}^{(0)}, (-\tilde{E}^{(3)})\tilde{\psi}^{(0)} \rangle + \langle \tilde{\psi}^{(0)}, (-\tilde{E}^{(2)})\tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, (-\tilde{E}^{(2)})\tilde{\psi}^{(0)} \rangle \\ & + \langle \tilde{\psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(2)} \rangle + \langle \tilde{\psi}^{(2)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(0)} \rangle \\ & + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_0 - E^{(0)})\tilde{\psi}^{(2)} \rangle \\ & + \langle \tilde{\psi}^{(2)}, (\mathcal{H}_0 - E^{(0)})\tilde{\psi}^{(1)} \rangle \end{aligned} \tag{11}$$

$$\begin{aligned} J^{(4)\prime} \equiv\ & \langle \tilde{\psi}^{(0)}, (-\tilde{E}^{(4)})\tilde{\psi}^{(0)} \rangle + \langle \tilde{\psi}^{(0)}, (-\tilde{E}^{(3)})\tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, (-\tilde{E}^{(3)})\tilde{\psi}^{(0)} \rangle \\ & + \langle \tilde{\psi}^{(0)}, (-\tilde{E}^{(2)})\tilde{\psi}^{(2)} \rangle + \langle \tilde{\psi}^{(2)}, (-\tilde{E}^{(2)})\tilde{\psi}^{(0)} \rangle \\ & + \langle \tilde{\psi}^{(1)}, (-\tilde{E}^{(2)})\tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(0)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(3)} \rangle \\ & + \langle \tilde{\psi}^{(3)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(0)} \rangle + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(2)} \rangle \\ & + \langle \tilde{\psi}^{(2)}, (\mathcal{H}_1 - \tilde{E}^{(1)})\tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(1)}, (\mathcal{H}_0 - E^{(0)})\tilde{\psi}^{(3)} \rangle \\ & + \langle \tilde{\psi}^{(3)}, (\mathcal{H}_0 - E^{(0)})\tilde{\psi}^{(1)} \rangle + \langle \tilde{\psi}^{(2)}, (\mathcal{H}_0 - E^{(0)})\tilde{\psi}^{(2)} \rangle \end{aligned} \tag{12}$$

After these preliminaries we now turn to a detailed discussion of $\tilde{E}^{(1)}$, $\tilde{E}^{(2)}$, $\tilde{\psi}^{(0)}$ and $\tilde{\psi}^{(1)}$. In the process we will encounter various complexities all of which will be recognized as the analogs, within our variational approach, of well known complications in the exact sum over states method.

II. The first order energy correction

Since we have exhausted the content of $J^{(0)}$ we now turn to $J^{(1)\prime}$. If we insert (6) into (11) and vary the C_a so as to make $\delta J^{(1)\prime}=0$ then we find[2] the set of homogeneous equations

$$\sum_{b=1}^{D} \langle \chi_a, (\mathcal{H}_1 - \tilde{E}^{(1)}) \chi_b \rangle C_b = 0 \qquad (13)$$

to determine the C_a and the $\tilde{E}^{(1)}$. These of course are precisely the equations of the exact theory[3]. In particular then the $\tilde{E}^{(1)}$ will be exact, hence we will drop the tilde.

We will now go on to investigate two extreme cases:

Case I - all $E^{(1)}$ different, and hence the corresponding $\tilde{\psi}^{(0)}$'s are uniquely defined (and exact).

Case II - all $\tilde{E}^{(1)}$ identical.

Intermediate situations can be handled by analogous methods.

III. $\tilde{E}^{(2)}$ and $\tilde{\psi}^{(1)}$ for Case I.

The previous considerations evidently exhaust the content of $J^{(0)}$ and $J^{(1)'}$. We now turn to $J^{(2)'}$. Denoting a particular normalized solution of (9) by $\psi_a^{(0)}$, $E_a^{(1)}$, and denoting the corresponding $\tilde{\psi}^{(1)}$ by $\tilde{\psi}_a^{(1)}$ we now consider

$$J_a^{(2)'} \equiv \langle \psi_a^{(0)}, (-\tilde{E}_a^{(2)})\psi_a^{(0)} \rangle + \langle \psi_a^{(0)}, (\mathcal{H}_1 - \tilde{E}_a^{(1)})\tilde{\psi}_a^{(1)} \rangle$$

$$+ \langle \tilde{\psi}_a^{(1)}, (\mathcal{H}_1 - \tilde{E}_a^{(1)})\psi_a^{(0)} \rangle \qquad (14)$$

This now has the form of the familiar Hylleraas variation principle for the second order energy, hence we just proceed as usual[1] and vary $\tilde{\psi}_a^{(1)}$; the conditions $\delta J_a^{(2)'} = 0$, $J_a^{(2)'} = 0$ then yielding an optimal $\tilde{\psi}_a^{(1)}$ and an optimal $\tilde{E}_a^{(2)}$. Further one can show by the standard argument (reference 1, page 278) that if $E^{(0)}$ is the lowest eigenvalue of $\mathcal{H}_0$ then $\tilde{E}_a^{(2)}$ will be an upper bound to the exact $E_a^{(2)}$.

However, note that we do not completely determine $\tilde{\psi}_a^{(1)}$ in this way. Namely, not only do we have the usual innocuous ambiguity that $\langle \psi_a^{(0)}, \tilde{\psi}_a^{(1)} \rangle$ may be chosen arbitrarily, but we also don't determine $\langle \psi_b^{(0)}, \tilde{\psi}_a^{(1)} \rangle$ for $b \neq a$. One sees the most easily by noting that one can replace $\tilde{\psi}_a^{(1)}$ in (14) by $\tilde{\psi}_a^{(1)} + \sum_{b=1}^{b} \psi_b^{(0)} F_{ba}$, where the F_{ba} are arbitrary constants, and get the same value for $J_a^{(2)'}$, since, as a

consequence of (9)[3]

$$\langle \psi_b^{(0)}, (\mathcal{H}_1 - E_a^{(1)})\psi_a^{(0)} \rangle = 0 \qquad ; \; a,b = 1,2\cdots D \tag{15}$$

To determine $\langle \psi_b^{(0)}, \tilde{\psi}_a^{(1)} \rangle$ then we must evidently go beyond $J_a^{(2)'}$. Let us then write

$$\tilde{\psi}_a^{(1)} = \tilde{\psi}_{a\perp}^{(1)} + \sum_b{}' \psi_b^{(0)} F_{ba} \tag{16}$$

where $\tilde{\psi}_{a\perp}^{(1)}$ has been determined from $J_a^{(2)'}$ and is to be uniquely defined by the requirement that it is orthogonal to all $\psi_b^{(0)}$. Also we have chosen the convenient normalization $\langle \psi_a^{(0)}, \tilde{\psi}_a^{(1)} \rangle = 0$ so that the prime means no term $b = a$ occurs in the sum. If we now insert this into $J_a^{(3)'}$, then the terms which explicitly involve the F_{ba} are

$$\sum_b{}' \sum_c{}' F_{bc}^* F_{ca} \langle \psi_c^{(0)}, (\mathcal{H}_1 - E_a^{(1)})\psi_b^{(0)} \rangle$$

$$+ \sum_b{}' F_{ba}^* \langle \psi_b^{(0)}, \mathcal{H}_1, \tilde{\psi}_{a\perp}^{(1)} \rangle + \sum_b{}' F_{ba} \langle \tilde{\psi}_{a\perp}^{(1)}, \mathcal{H}_1, \psi_b^{(0)} \rangle$$

where we have omitted terms which vanish identically. If now we vary F_{ba}^* we find

$$\sum_c{}' F_{ca} \langle \psi_c^{(0)}, (\mathcal{H}_1 - E_a^{(1)})\psi_b^{(0)} \rangle + \langle \psi_b^{(0)}, \mathcal{H}_1, \tilde{\psi}_{a\perp}^{(1)} \rangle = 0$$

But from (15) -- remember the $\psi_a^{(0)}$ are normalized and of course are orthogonal --

$$\langle \psi_c^{(0)}, (\mathcal{H}_1 - E_a^{(1)}) \psi_b^{(0)} \rangle = (E_b^{(1)} - E_a^{(1)}) \delta_{bc}$$

So we have

$$F_{ba} (E_b^{(1)} - E_a^{(1)}) + \langle \psi_b^{(0)}, \mathcal{H} \tilde{\psi}_{a-1}^{(1)} \rangle = 0$$

and hence the explicit solution

$$F_{ba} = - \frac{\langle \psi_b^{(0)}, \mathcal{H}_1 \tilde{\psi}_{a-1}^{(1)} \rangle}{E_b^{(1)} - E_a^{(1)}} \quad , \quad b \neq a$$

Thus $\tilde{\psi}_a^{(1)}$ has now been determined.

IV. $\tilde{E}^{(2)}$ for case II.

For case II

$$\langle \chi_a, (\mathcal{H}_1 - E^{(1)}) \chi_b \rangle \equiv 0 \qquad (17)$$

so that the C_a are as yet completely arbitrary. Since we have evidently exhausted the content of $\mathcal{J}^{(1)'}$ we turn to $\mathcal{J}^{(2)'}$. Inserting (6) into (10) and varying C_a^* one finds

$$-\tilde{E}^{(2)} C_a + \langle \chi_a, (\mathcal{H}_1 - E^{(1)}) \tilde{\psi}^{(1)} \rangle = 0$$

which, in the notation of the previous section, can, from (17), be written as

$$-\tilde{E}^{(2)} C_a + \langle \chi_a, (\mathcal{H}_1 - E^{(1)}) \tilde{\psi}_\perp^{(1)} \rangle = 0 \qquad (18)$$

So here we have a set of relations between the C_a, $\tilde{E}^{(2)}$ and $\tilde{\psi}_\perp^{(1)}$. To get others we vary $\tilde{\psi}_\perp^{(1)*}$ in $\mathcal{J}^{(2)'}$ to find

$$\sum_a C_a \langle \delta\tilde{\psi}_\perp^{(1)}, (\mathcal{H}_1 - E^{(1)}) \chi_a \rangle + \langle \delta\tilde{\psi}_\perp^{(1)}, (\mathcal{H}_0 - E^{(0)}) \tilde{\psi}_\perp^{(1)} \rangle = 0 \qquad (19)$$

In general the set of equations (18) and (19) provides us with a non-linear coupled problem which must be solved in some self-consistent way. (If, however, one restricts oneself to linear variational parameters and a fixed basis set the problem becomes a simple linear one[4]). One might, for example, guess a set of C_a's normalized so that $\sum_a |C_a|^2 = 1$. One would then solve (16) for $\tilde{\psi}_\perp^{(1)}$. Then with this $\tilde{\psi}_\perp^{(1)}$, (15) would provide us with an approximate $\tilde{E}^{(2)}$ and a new normalized set of C_a's. One would then iterate until a consistent solution is obtained. Having obtained one solution one would then go on and look for another.

This then raises the question, how many $\tilde{E}^{(2)}$ values will one find? In particular if one uses non linear parameters in $\tilde{\Psi}^{(1)}_{\perp}$ then it does not seem obvious that one cannot get more than D. If this is indeed the case then one would be faced with the problem of choosing the "best" D solutions.

In any case (15) and (16) provide us with an in principal method for determining the $\tilde{E}^{(2)}$ and $\tilde{\Psi}^{(1)}_{\perp}$. Let us now go on and discuss how one would complete the determination of $\tilde{\Psi}^{(1)}$.

V. $\tilde{\Psi}^{(1)}$ for a special case of Case II.

Let us suppose that we have D solutions $\tilde{\Psi}^{(1)}_{a\perp}$, with the associated $\tilde{E}^{(2)}_{a}$ and $\tilde{\Psi}^{(0)}_{a}$ being distinct and unique, though preseumably <u>not</u> exact[5]. We will also assume that the $\tilde{\Psi}^{(0)}_{a}$ are normalized and orthogonal[6]. It then remains to determine the $\langle \tilde{\Psi}^{(0)}_{b}, \tilde{\Psi}^{(1)}_{a} \rangle$.

Since for Case II, $E_a^{(1)} = E_b^{(1)}$, it is clear that $T_a^{(3)'}$ will be of no help so we must turn to $T_a^{(4)'}$. Writing

$$\tilde{\Psi}^{(1)}_{a} = \tilde{\Psi}^{(1)}_{a\perp} + \sum_b{}' \tilde{\Psi}^{(0)}_{b} F_{ba}$$

the relevant terms are then

$$\sum_b{}' E^{*}_{ba} F_{ba}(-\tilde{E}^{(2)}_{a}) + \sum_b{}' F^{*}_{ba} \langle \tilde{\Psi}^{(0)}_{b}, \mathcal{H}_1, \tilde{\Psi}^{(2)}_{a\perp} \rangle + \sum_b{}' F_{ba} \langle \tilde{\Psi}^{(1)}_{a\perp}, \mathcal{H}_1, \tilde{\Psi}^{(0)}_{b} \rangle$$

where we have again dropped terms which vanish identically. In addition $\tilde{\psi}^{(2)}_{a\perp}$ is also involved in the terms

$$\langle \tilde{\psi}^{(2)}_{a\perp}, (H_0 - E^{(0)}) \tilde{\psi}^{(2)}_{a\perp} \rangle + \langle \tilde{\psi}^{(2)}_{a\perp}, (H_1 - E^{(1)}) \tilde{\psi}^{(1)}_{a\perp} \rangle + \langle \tilde{\psi}^{(1)}_{a\perp}, (H_1 - E^{(1)}) \tilde{\psi}^{(1)}_{a\perp} \rangle$$

If now we vary F^{*}_{ba} we find

$$-\tilde{E}^{(2)}_a F_{ba} + \langle \tilde{\psi}^{(0)}_b, H_1 \tilde{\psi}^{(2)}_{a\perp} \rangle = 0 \qquad (20)$$

while if we vary $\tilde{\psi}^{(2)*}_{a\perp}$ we find

$$\langle \delta\tilde{\psi}^{(2)}_{a\perp}, (H_0 - E^{(0)}) \tilde{\psi}^{(2)}_{a\perp} \rangle + \langle \delta\tilde{\psi}^{(2)}_{a\perp}, (H_1 - E^{(1)}) \tilde{\psi}^{(1)}_{a\perp} \rangle + \sum_b{}' F_{ba} \langle \delta\tilde{\psi}^{(2)}_{a\perp}, H_1 \tilde{\psi}^{(0)}_b \rangle = 0$$

and we must now solve this coupled set to determine the F_{ba} (and in addition $\tilde{\psi}^{(2)}_{a\perp}$). If one uses a fixed basis set and linear variational parameters[4] the solution is simple. However, since it does not seem possible to give an explicit solution in general, we leave the discussion at this point.

VI. Discussion

It is clear from this preliminary survey that a variational approach to degenerate perturbation problems will involve one in potentially quite complicated problems of solving coupled functional equations in a self consistent way once the degeneracy

persists beyond first order. Also we have noted that there may be difficulties in interpreting the solutions.

As we have mentioned the solutions to these problems are straight forward if one uses linear variational parameters with a fixed basis set (indeed explicit use of the variation principle is really superfluous in this case), but how one handles other parametrizations will presumably vary from case to case.

It is a pleasure to acknowledge helpful discussions with Dr. Hojing Kim.

Footnotes and References

1. For a recent review see J. O. Hirschfelder, W. Byers Brown, and S. T. Epstein, Advan. Quantum Chem. 1, 255-374 (1964).
2. Note that with this sort of trial function $\delta\tilde{\Psi}^{(0)} = \eta\tilde{\Psi}^{(0)}$ where η is a small constant, is a possible variation whence $\delta J^{(1)\prime}=0$ implies $J^{(1)\prime}=0$. Similar remarks will apply to other variations which we will be using later, so that we will never need to use $J^{(1)\prime}=0$ explicitly.
3. See for example L. Pauling and E. B. Wilson, Jr., Introduction to Quantum Mechanics (McGraw Hill Book Co., New York, 1935, especially section 24).
4. With a fixed basis set one is doing perturbation theory exactly for the operator $\bar{\mathcal{H}}_0+\bar{\mathcal{H}}_1$, where $\bar{\mathcal{H}}_0 = P\mathcal{H}_0P$ and $\bar{\mathcal{H}}_1 = P\mathcal{H}_1P$ and where P is the projection operator onto the basis set. If then one chooses the basis, as one may always do, so as to diagonalize $\bar{\mathcal{H}}_0$, then all formulae look like the exact sum over states formulae except the sums are finite (See also H. J. Kolker and H. H. Michels, J. Chem. Phys. 43, 1027 (1965)).
5. If $\tilde{\Psi}_a^{(0)}$ is not exact then all we can prove in general concerning the relation between the $\tilde{E}_a^{(2)}$ and the exact values is that for the ground state $\tilde{E}_a^{(2)}$ is an upper bound to the lowest $E^{(2)}$ associated with the same symmetry

(see footnote 11, page 300 of reference 1). If one uses linear variational parameters then, of course, one can say more (J.K.L. Mac Donald, Phys. Rev. 43, 830 (1933). See also footnote 6).

6. This orthogonality assumption is not essential for the analysis which follows. However, if $\eta \tilde{\Psi}^{(1)}_{a\perp}$ is a possible variation of $\Psi^{(1)}_{b\perp}$ and conversely then it is easy to prove from (18) and (19) that $\tilde{\Psi}^{(0)}_a$ and $\tilde{\Psi}^{(0)}_b$ will be orthogonal. Thus given a set $\tilde{\Psi}^{(1)}_{a\perp}$ which did not yield $\langle \tilde{\Psi}^{(0)}_a, \tilde{\Psi}^{(0)}_b \rangle = \delta_{ab}$ we can always get orthogonality by entering the variational principle again with the new variational function for $\tilde{\Psi}^{(1)}_\perp$:

$$\sum_{a=1}^{\nu} \tilde{\Psi}^{(1)}_{a\perp} D_a$$

where now the $\tilde{\Psi}^{(1)}_{a\perp}$ are fixed and the D_a are linear variational parameters which we are to determine. Incidentally, from MacDonald's theorem[5] this will also ensure, at least for the ground state, that the (new) $\tilde{E}^{(1)}_a$, arranged in order of increasing value, will be upper bounds to the similarly ordered exact $E^{(1)}_a$.