# THEORETICAL CHEMISTRY INSTITUTE

# THE UNIVERSITY OF WISCONSIN

## THE VARIATIONAL METHOD IV - LOWER BOUNDS FOR THE ENERGY

by

Saul T. Epstein



WIS-TCI-361

21 October 1969

MADISON, WISCONSIN

# The Variational Method IV - Lower Bounds for the Energy*

by

Saul T. Epstein

Theoretical Chemistry Institute and
Physics Department
University of Wisconsin

I WOULD GREATLY APPRECIATE YOUR CRITICISMS, COMMENTS, 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.

## 40. Introduction

In the preceding work, what bounds were found were upper bounds. In what follows we will be concerned with lower bounds. In Sec. 17, Part E we already found a need for such bounds in order to get any bounds at all on "Mixed second order properties". More fundamentally, if we have both upper and lower bounds we can make purely theoretical predictions complete with ± error.

Although lower bound formulae have been know for a long time, they have not been extensively applied. The main reason is that many of them involve calculating $\widetilde{E^2} \equiv (\tilde{\Psi}, H^2\tilde{\Psi})/(\tilde{\Psi},\tilde{\Psi})$ (more precisely, $\frac{(H\tilde{\Psi},H\tilde{\Psi})}{(\tilde{\Psi},\tilde{\Psi})}$, however we will make no distinction) which on the one hand is usually an unpleasant task and on the other hand tends to magnify errors in $\tilde{\Psi}$ with the result that the same $\tilde{\Psi}$ will usually give a considerably poorer lower bound than an upper bound, and thus give a wide gap between upper and lower bounds - facts which of course discourage the calculation of the lower bounds. However as computational techniques become more powerful, interest in lower bound calculations seems to be reviving. In what follows we will review some, but by no means all, of this work.

## 41. The Weinstein-MacDonald Bracketing Theorem

Given a real number $\alpha$ there will be some eigenvalue of $H$ to which it will be closest. Let us denote this eigenvalue by $E_n$. Then clearly (Problem: Show this) for any $\tilde{\Psi}$

$$(\tilde{\Psi}, (H-\alpha)^2\tilde{\Psi})/(\tilde{\Psi},\tilde{\Psi}) \geq (E_n-\alpha)^2 \tag{41-1}$$

(The possibility that $\alpha$ lies midway between two eigenvalues can be considered as a limiting case and causes no difficulty.)

or

$$\widetilde{E}^2 - 2\alpha\widetilde{E} + \alpha^2 \geq (E_n - \alpha)^2 \tag{41-2}$$

which implies that

$$-\sqrt{\widetilde{E}^2 - 2\alpha\widetilde{E} + \alpha^2} \leq (E_n - \alpha) \leq \sqrt{\widetilde{E}^2 - 2\alpha\widetilde{E} + \alpha^2}$$

or finally

$$\alpha - \sqrt{\widetilde{E}^2 - 2\alpha\widetilde{E} + \alpha^2} \leq E_n \leq \alpha + \sqrt{\widetilde{E}^2 - 2\alpha\widetilde{E} + \alpha^2} \tag{41-3}$$

This is the basic result. It is due to Weinstein (Proc. Nat. Acad. Sci. 20, 529 (1934)) who discussed the choice $\alpha = \widetilde{E}$ in detail, and to MacDonald, (Phys. Rev. 46, 828 (1934)). One way to read it is as a "bracketing theorem" - that is it says that there is an eigenvalue of $H$ between the limits indicated so that given any $\widetilde{\Psi}$ and any $\alpha$ one will bracket some eigenvalue of $H$ in this way. However in order to say that it is precisely the n-th eigenvalue which is bracketed we must evidently also know that

$$\frac{E_{n-1} + E_n}{2} \leq \alpha \leq \frac{E_n + E_{n+1}}{2} \tag{41-4}$$

A point of notation: In what follows $E_1$, will denote the ground state. One may then check that in all formulae it will be consistent to put

$$E_0 = -\infty \tag{41-5}$$

One further point: Clearly in (41-4) we can replace $E_{n\pm1}$ by the energies of those states closest to $\alpha$ whose wave functions are not orthogonal to $\tilde{\Psi}$. In particular if $H$ is symmetric and $\tilde{\Psi}$ has a definite symmetry then n-1, n, n+1 can be taken as the successive levels of that symmetry.

Some applications of this bracketing theorem have been made. In particular J. Goodisman, J. Chem. Phys. 47, 5247 (1967), in a paper which also contains reference to earlier work, has made the Weinstein choice $\alpha = \tilde{E}$ so that (3) becomes

$$\tilde{E} - \sqrt{\tilde{E}^2 - \tilde{E}^2} \leq E_n \leq \tilde{E} + \sqrt{\tilde{E}^2 - \tilde{E}^2} \qquad (41\text{-}6)$$

and has then determined an optimal trial function by minimizing $\tilde{E}^2 - \tilde{E}^2$ thereby minimizing the difference between the upper and lower bounds. (Note that even if the space of trial functions is linear this leads to a non-linear problem). We refer the reader to his paper for details and for some numerical results. We now turn our attentions to another use for (3) and (4) - namely as a source of separate upper and lower bounds. We will first discuss lower bounds.

(For specific states)

Minimizing the energy "variance" $\tilde{E}^2 - \tilde{E}^2$ has been suggested on other grounds by several authors as an alternative to the usual variational method for determining a best wave function. For a further discussion and references see the paper by Goodisman mentioned above and also his earlier paper J. Chem. Phys. 45, 3659 (1966). Also mention should be made of the papers of Frager and Birss, J. Chem. Phys. 40, 3207 and 3212 (1964) in which an SCF scheme is developed on the basis of (6).

## 42. The Stevenson-Crawford Lower Bound

We first note that the left hand inequality in (41-3) will remain an inequality even if we ignore the left hand inequality in (41-4), i.e. if we require only that

$$\alpha \leq \frac{E_n + E_{n+1}}{2} \tag{42-1}$$

since this is already sufficient to ensure that the eigenvalue to which $\alpha$ is closest is certainly $\leq E_n$ and since the l.h.s. of (41-3) will be less than or equal to that eigenvalue, it will necessarily also be $\leq E_n$. Thus we have that if (1) is satisfied then

$$E_n^L(\alpha) \equiv \alpha - \sqrt{\tilde{E}^2 - 2\alpha\tilde{E} + \alpha^2} \leq E_n \tag{42-2}$$

We now note (Stevenson and Crawford, Phys. Rev. 54, 375 (1938)) that the derivative with respect to $\alpha$ of l.h.s. of (2) can be written as

$$1 - \frac{(\alpha - \tilde{E})}{\sqrt{(\alpha-\tilde{E})^2 + \tilde{E}^2 - \tilde{E}^2}}$$

which, since $\tilde{E}^2 - \tilde{E}^2 \geq 0$, is positive. Hence to get a best lower bound one tries to choose $\alpha$ as large as possible consistent with (1). Having chosen $\alpha$, one may simply use a given $\tilde{\Psi}$ to compute a lower bound or one may choose that $\tilde{\Psi}$ from among a set of trial functions which minimizes $\tilde{E}^2 - 2\alpha\tilde{E}$ (Note that if the space of trial functions is linear this leads to a linear problem). Calculations of this sort for simple model problems and for two electron problems can be found in the following references

Stevenson and Crawford, Phys. Rev. 54, 374 (1938)
Kinoshita, Phys. Rev. 115, 306 (1969)
Caldow and Coulson, Proc. Camb. Phil. Soc. 57, 341 (1961)
Fröman and Hall, J. Mol. Spc. 7, 410 (1961)
M. E. Schwartz, Proc. Phys. Soc. 90, 51 (1967)
Mazziotti, J. Chem. Phys. 50, 3330 (1969)

We will not attempt to present the results in any detail but only note that in general it seems harder (one needs a more flexible set of $\tilde{\psi}$ ) to get a good lower bound than to get a good upper bound.

## 43. The Temple-Kohn-Kato (TKK) Lower Bound

We now want to discuss the choice of $\alpha$ in more detail. If one has experimental information about $E_n$ and $E_{n+1}$ then one would clearly put $\alpha = (E_n + E_{n+1})/2$ and proceed as indicated in the previous section. However clearly in such cases one is not really interested in determining the energy, one already knows it, but in determining an optimal wave function from among a set of trial functions, or in testing a wave function found in some other way, for example from the ordinary variational method, by comparing $E_n^L$ with $E_n$ or with $\tilde{E}$ . However suppose one really wants to determine an $E_n^L$ in an abinitio fashion; then how does one choose $\alpha$ ? Well we first note that if $\mathcal{E}_n^L$ and $\mathcal{E}_{n+1}^L$ are lower bounds to $E_n$ and $E_{n+1}$ resepctively then (1) is certainly satisfied if

$$\alpha \leq \frac{\mathcal{E}_n^L + \mathcal{E}_{n+1}^L}{2} \tag{43-1}$$

Now let us suppose that we know an $\mathcal{E}_{n+1}^L$ from somewhere (experiment or one of the purely theoretical procedures which we will describe later) then we can proceed as follows. Choose an $\alpha$ and calculate the l.h.s.

of (42-2). Continuing to denote it by $E_n^L(\alpha)$, though at this stage we don't really know that it is a lower bound to $E_n$, then we will have consistency if

$$\alpha \leq \frac{E_n^L(\alpha) + \mathcal{E}_{n,n-1}^L}{2} \tag{43-2}$$

and hence $E_n^L(\alpha)$ will be a lower bound to $E_n$. Moreover it seems clear that the best $\alpha$ we could choose in this way would be the one for which the equality in (2) is realized. Denoting this value of $\alpha$ by $\mathring{\alpha}$ and $E_n^L(\mathring{\alpha})$ by $\mathring{E}_n^L$ we then have

$$\mathring{\alpha} = \frac{\mathring{E}_n^L + \mathcal{E}_{n,n-1}^L}{2} \tag{43-3}$$

$$\mathring{E}_n^L = \mathring{\alpha} - \sqrt{\tilde{E}^2 - 2\mathring{\alpha}\tilde{E} + \mathring{\alpha}^2} \tag{43-4}$$

If now we insert (3) into (4), transpose $\mathring{\alpha}$ to the left and square, we can solve for $\mathring{E}_n^L$ to find

$$\mathring{E}_n^L = \frac{\tilde{E}\,\mathcal{E}_{n,n-1}^L - \tilde{E}^2}{\mathcal{E}_{n,n-1}^L - \tilde{E}} \tag{43-5}$$

or

$$\mathring{E}_n^L = \tilde{E} - \frac{\tilde{E}^2 - \tilde{E}^2}{\mathcal{E}_{n,n-1}^L - \tilde{E}} \tag{43-6}$$

However since we got this result by squaring, we must go back and check that it is actually a solution of (4). One then readily sees that all

is well (one needs $\dot{E}^L_n - \alpha < v$ ) if and only if

$$\mathcal{E}^L_{n+1} - \tilde{E} > 0 \qquad (43\text{-}7)$$

which provides an additional restriction, this time on $\tilde{\psi}$. Note that from (2) this then implies

$$\dot{E}^L_n \leq \tilde{E} < \mathcal{E}^L_{n+1} \leq E_{n+1} \qquad (43\text{-}8)$$

For the ground state this result is due to Temple (Proc. Roy. Soc. A119, 276 (1928)). For arbitrary $n$ it was first derived by Kohn (Phys. Rev. 71, 902 (1947)) and independently by Kato (J. Phys. Soc. Japan 4, 334 (1949)). See also Maehly (N. Cimento 8, 466 (1951)). For an interesting geometrical interpretation see Washizu (Quart. J. Mech. and Appl. Math. 8, 311 (1955)).

If we know that $\mathcal{E}^L_{n+1}$ is not only a lower bound to $E_{n+1}$ but is also an upper bound to $E_n$ then the right hand side of (6) is an example of a "Bracketing function" that is we put in an upper bound to $E_n$ (an upper bound which also satisfies (5)) And we get out a lower bound. More generally as long as (7) is satisfied, $\mathcal{E}^L_{n+1}$ may be a very poor lower bound to $E_{n+1}$ and (6) still provides a lower bound to $E_n$. However we can be more precise if we know that in fact for some $K \leq n$

$$E_K \leq \mathcal{E}^L_{n+1} \leq E_{n+1} \qquad (43\text{-}9)$$

Namely the right hand side of (6) will then evidently be a lower bound to $E_K$ (and hence, of course, to $E_n$ ).

Problem: The procedures outlined are not self contained. One must supply $\mathcal{E}_{n+1}^{L}$ "from outside". Does the self contained choice $\mathcal{E}_{n+1}^{L} = \mathcal{E}_{n}^{L}(\alpha)$ lead anywhere?

Problem: Given an $\mathcal{E}_{n+1}^{L}$ show that one can use these procedures to determine lower bounds for all $E_{n'}$ with $n' \leq n$.

Problem: Prove that for any $\tilde{\Psi}$ one has $(\tilde{\Psi}, (H-\mathcal{E}_n)(H-\mathcal{E}_{n+1})\tilde{\Psi}) \geq 0$

Show that this leads to (6) when (7) holds.

To make more precise the sense in which the choice $\alpha = \mathring{\alpha}$ is a "best choice" consider what happens if one chooses $\alpha$ "better" value of $\alpha$, namely

$$\alpha' = \frac{\mathcal{E}_n^{L} + \mathcal{E}_{n+1}^{L}}{2}$$

where

$$\mathcal{E}_n^{L} > \mathring{E}_n^{L} \tag{43-10}$$

so that $\alpha' > \mathring{\alpha}$ and hence is guaranteed to give a better bound $E_n^{L'}$ where

$$E_n^{L'} = \alpha' - \sqrt{\tilde{E}^2 - 2\alpha'\tilde{E} + \alpha'^2}$$

It is now easy to show (Problem: Fill in the details) that

$$E_n^{L'} - \mathcal{E}_n^{L} = \frac{\mathcal{E}_{n+1}^{L} - \mathcal{E}_n^{L}}{2} - \sqrt{\left(\frac{\mathcal{E}_{n+1}^{L} - \mathcal{E}_n^{L}}{2}\right)^2 + (\mathcal{E}_{n+1}^{L} - \tilde{E})(\mathcal{E}_n^{L} - \mathring{E}_n^{L})}$$

whence from (7) and (10) we see that

$$E_m^{L\prime} - \mathcal{E}_n^L < 0$$

i.e. the lower bound that one gets out, although it is better than $\mathring{E}_m^L$, is worse than the bound (namely $\mathcal{E}_n^L$) that one puts in. Thus the choice $\alpha = \mathring{\alpha}$ is optimal insofar as one wants to use (43-2) to estimate $E_m$. To quote E. B. Wilson Jr. (J. Chem. Phys. 43, S 172 (1965) Footnote b) "If one knows a better value of $\mathcal{E}_n^L$ in advance there is no use in calculating a poorer value for the quantity." (See also Schmid and Schweger, Z. fur Phys. 210, 309 (1968)).

One further point "in favor" of the choice $\alpha = \mathring{\alpha}$ is that the resultant $\mathring{E}_m^L$ involves only a second order error, i.e. $\alpha = \mathring{\alpha}$ yields a variational bound. Thus writing

$$\mathring{E}_m^L = \frac{(\tilde{\Psi}, H(\mathcal{E}_{n\mathring{\alpha}}^L - H)\tilde{\Psi})}{(\tilde{\Psi}, (\mathcal{E}_{n\mathring{\alpha}}^L - H)\tilde{\Psi})} \tag{43-11}$$

one readily shows that if

$$\tilde{\Psi} = \Psi_n + \hat{\Delta}$$

then

$$\mathring{E}_m^L = E_m + O(\hat{\Delta}^2) \tag{43-12}$$

Problem: Derive (12).

Problem: Give an analogous discussion of the general bound (42-2). Note the difference in the cases $\alpha > E_n$ and $\alpha < E_n$. Reconcile this result with (12). Hint: Show that $\alpha - E_n$ is a positive number plus terms of order $\Delta^2$.

Applications of the Temple bound to the ground state of Helium may be found in Wilets and Cherry, Phys. Rev. 123, 112 (1956), Kinoshita, Phys. Rev. 105, 1490 (1957), and in Pekeris, Phys. Rev. 126, 1471 (1962). In these calculations $E_2$ was taken from experiment and $\psi$ was taken from an ordinary variational calculation. In general, except for the work of Pekeris, the quality of the results is poor - the gaps between the upper and lower bounds are large and most of the authors then supplement their calculations by various extrapolation procedures with the hope of getting a better answer. See for example Kinoshita, also Conroy, J. Chem. Phys. 41, 1336 (1964). Applications to $H_2$ and $H_2^+$ with similar results can be found in Goodisman and Secrest J. Chem. Phys. 41, 3610 (1964); 45, 1515 (1966), and in Walmsley and Coulson, Proc. Camb. Phil. Soc. 67, 769 (1966).

The Temple bound has also been used variationally. (Like the Stevenson-Crawford method a linear space of trial functions leads to a linear problem which we will discuss in more detail in the next section.) For examples of such calculations for model problems and for real one- and two-electron problems see the last four references given at the end of Sec. 42. For applications to nuclear physics see for example Tang et al. Nuc. Phys. 65, 203 (1965). The quality of results is again much as indicated before - one must usually work harder to get a good lower bound than a good upper bound.

We have been concerned mainly with these formulae as sources of lower bounds for the energy. However if they are used variationally then they also provide an optimal trial function and one might wonder about the quality of that wave function. One point in this connection is the following: Consider the TKK formulae and suppose that we take for $\mathcal{E}^{L}_{n+1}$, its best value, namely $E_{n+1}$. Then from (11) we immediately see that $\tilde{\Psi}$ and $\tilde{\Psi} + \epsilon \Psi_{n+1}$ yield the same $\dot{E}^{L}_{n}$ for any number $\epsilon$. This degeneracy suggests that in general one need not expect an optimal $\tilde{\Psi}$ to look much like $\Psi_n$. (See also J. Goodisman, Theoret. Chim. Acta 4, 343 (1966)). However, if in fact a calculation yields degeneracy or near degeneracy (several solutions with $\mathcal{E}^{L}_{n}$'s varying within tolerable limits) then one could well ask for that solution which also yields the best upper bound.

## 44. The Lehmann-Maehly Bounds

If we insert a linear trial function

$$\tilde{\Psi} = \sum_{k=1}^{M} a_k \chi_k \tag{44-1}$$

into the TKK bound and require that

$$\frac{\partial \dot{E}^{L}_{n}}{\partial a_k} = 0 \tag{44-2}$$

we are led to the set of linear equations (Problem: Fill in the details)

$$\sum_{p} \left( \chi_k, (\mathcal{E}^{L}_{n+1} - H)(\dot{E}^{L}_{n} - H)\chi_p \right) a_p = 0 \tag{44-3}$$

and hence to the secular equation

$$\left| (\chi_\mu, (\mathcal{E}_{m+1}^{L} - H)(\mathring{E}_\mu^{L} - H)\chi_\rho) \right| = 0 \qquad (44\text{-}4)$$

from which to determine $\mathring{E}_m^{L}$ . In general (4) will have $M$ solutions and, in accord with the discussion in the previous section, we would choose as the best approximation to $E_m$, the largest root which has the property that the associated $Q_{\mu\kappa}$ yield

$$\mathcal{E}_{m+1}^{L} > \tilde{E} \qquad (44\text{-}5)$$

From what we have asid so far, no interpretation can be given to the other roots except that of course the smaller ones are also lower bounds to $E_m$ .

It is now of interest to remark that there is another method for finding bounds [we will follow the discussion of H. S. Maehly, Helv. Phys. Acta 25, 547 (1952). Essentially the same results were found earlier from a somewhat different point of view by N. J. Lehmann, Z. Angew. Math. Mech. 29, 341 (1949); 30, 1 (1950). See also Washizu, Q. J. Mech. and Appl. Math. 8, 311 (1955)] which yields an interpretation for all the roots.

The basic idea is as follows. Consider the operator $(H-\epsilon)^{-1}$ where $\epsilon$ is a real number which we assume not equal to an eigenvalue of $H$ . The eigenvalues, (or $(H-\epsilon)^{-1}$) are clearly

$$\gamma_k = (E_k - \epsilon)^{-1} \qquad (44\text{-}6)$$

whence we have

$$E_k = \epsilon + \frac{1}{\lambda_k} \tag{44-7}$$

(Problem: Prove this)

Now the point is that the $\lambda_k$ are bounded above and below. Suppose now we apply the ordinary linear variational method with $M$ functions to the operator $(H-\epsilon)^{-1}$. Since the $\lambda_k$ are bounded below we know from our previous work that we will get upper bounds to $M$ successive $\lambda_k$ starting with the lowest. However since the $\lambda_k$ are also bounded from above one can also show, by a slight extension of our previous method, that we also get lower bounds to $M$ successive $\lambda_k$ starting with the highest. Suppose now that $E$ is an $E_k$ which is less than $\epsilon$ and lies within $M$ eigenvalues of $\epsilon$. Then since $(E-\epsilon)^{-1}$ is $<0$ it follows that the linear variational method will yield an upper bound to $(E-\epsilon)^{-1}$. Denoting the result of the linear variational method by $(\check{E}-\epsilon)^{-1}$ we then have

$$(\check{E}-\epsilon)^{-1} > (E-\epsilon)^{-1} \tag{44-8}$$

Now suppose that we also have

$$(\check{E}-\epsilon) < 0 \tag{44-9}$$

Then we can multiply (44-8) by $(\check{E}-\epsilon)(E-\epsilon)$ which is $>0$ to find

$$E-\epsilon > \check{E}-\epsilon$$

or

$$E > \check{E} \tag{44-10}$$

On the other hand if $E$ is within $M$ eigenvalues of $\epsilon$ above $\epsilon$ then since $(E-\epsilon)^{-1}>0$ the linear variational method will yield lower bounds to $(E-\epsilon)^{-1}$ . Thus

$$(\check{E}-\epsilon)^{-1} < (E-\epsilon)^{-1} \tag{44-11}$$

whence if

$$(\check{E}-\epsilon) >0 \tag{44-12}$$

so that $(\check{E}-\epsilon)(E-\epsilon) >0$ then we have

$$E-\epsilon < \check{E}-\epsilon$$

or

$$E < \check{E} \tag{44-13}$$

Suppose now that in fact we find that

$$\check{E}_j < \epsilon < \check{E}_{j+1} \tag{44-14}$$

Then it follows from the above that we get lower bounds for the $j$ levels below immediately $\epsilon$ and upper bounds for the $M-j$ levels immediately above $\epsilon$ .

Now we want to make contact with (4). If we simply apply the linear variational method to $(H-\epsilon)^{-1}$ then we are led to

$$|(\phi_k, [(H-\epsilon)^{-1} - \hat{\lambda}]\phi_l)| = 0 \tag{44-15}$$

which is usually not very tractable because of the presence of the inverse operator.

However suppose that we write

$$\Phi_k = (H - \epsilon)\,\chi_L$$

then (Problem: Fill in the details) (15) is replaced by

$$\left| (\chi_k, (H-\epsilon)(H-\check{E})\chi_L) \right| = 0 \tag{44-16}$$

where

$$\check{E} = \epsilon + \frac{1}{\lambda} \tag{44-17}$$

which, if we identify $\epsilon$ with $\mathcal{E}^L_{n+1}$, and $\check{E}$ with $\dot{E}^L_n$, becomes (4). But now we have an interpretation of all the roots. Namely suppose that $\mathcal{E}^L_{n+1}$ satisfies

$$\check{E}_j \leq \mathcal{E}^L_{n+1} \leq \check{E}_{j+1} \tag{44-14*}$$

And (44-18)

$$\cdots, E_{N+1} \cdots \leq E_{N+j} \leq \mathcal{E}^L_{n+1} \leq E_{N+j+1} \leq \cdots E_{N+M} \cdots$$

for the same $j$. Then it follows from the preceding discussion that

$$\check{E}_i \leq E_{N+i} \qquad i = 1, 2 \cdots j \tag{44-19}$$

$$\check{E}_k \geq E_{N+k} \qquad k = j+1, \cdots M \tag{44-20}$$

Note however that in order to make these precise statements we must know $N$, i.e. we must know precisely where $\mathcal{E}^L_{n+1}$ lies with respect to the $E_i$. If all we know is that $\mathcal{E}^L_{n+1}$ is a lower bound to $E_{n+1}$

then the most we can say is that

$$\check{E}_i \leq E_n \qquad i = 1,2\cdots j \tag{44-21}$$

So far we seem to have made no contact with (5). We will now show that condition (5) is in fact equivalent to (14*) which defines $j$. Namely we clearly have (Problem: Show this)

$$\hat{\lambda}_i = \frac{(\check{\psi}_i, (H-\epsilon)\check{\psi}_i)}{(\check{\psi}_i, (H-\epsilon)^2\check{\psi}_i)} \tag{44-22}$$

where $\check{\psi}_i$ is the appropriate linear combination of the $\chi_L$. But then from (22) and (17) one readily finds that

$$\epsilon - \check{E}_i = \frac{(\check{\psi}_i, (H-\epsilon)^2\check{\psi}_i)}{(\check{\psi}_i, (\epsilon-H)\check{\psi}_i)} \tag{44-23}$$

From which it follows that

$$\epsilon - \check{E}_i \gtrless 0 \quad \text{as} \quad \epsilon - \frac{(\check{\psi}_i, H\check{\psi}_i)}{(\check{\psi}_i, \check{\psi}_i)} \text{ is } \gtrless 0 \tag{44-24}$$

which proves the point

Further we can now recognize the result

$$\check{E}_i \leq E_{N+j} \qquad i = 1,2\cdots j$$

which follows from (19), is equivalent to our earlier result (43-9). (given following)

As outlined above, in order to find $j$ one must first find the $\tilde{E}_i$.

In fact Maehly shows that this is not the case. Namely he proves, and we refer the reader to his paper for details, that $j$ can be fixed by

$$\hat{E}_j < \epsilon < \hat{E}_{j+1} \tag{44-25}$$

where the $\hat{E}_i$ are the result of a linear variational calculation applied to $H$, using the $\chi_i$ as a basis set. More precisely he shows that

$$\tilde{E}_i \leq \hat{E}_i \qquad i = 1 \cdots j \tag{44-26}$$

$$\tilde{E}_k \geq \hat{E}_k \qquad k = j+1 \cdots M \tag{44-27}$$

which, combined with (14), yields (25)

---

One might expect that for a given basis set one would get the best lower bound for $E_\ell$ by having $\epsilon$ lie between $E_\ell$ and $E_{\ell+1}$ rather than between some higher pairs. However from an examination of Figure One in Maehly's paper, it is clear that this is not necessarily the case.

## 45. The Upper Bound Formulae

Since the formal discussion here runs completely parallel to that on lower bounds, we will not go into details. If one requires only that

$$\frac{E_{n-1} + E_n}{2} \leq \alpha \tag{45-1}$$

then we still have the upper bound formula

$$E_n < \alpha + \sqrt{\tilde{E}^2 - 2\alpha\tilde{E} + \alpha^2} \tag{45-2}$$

Since the right hand side of (2) is an increasing function of $\alpha$ one gets the best upper bound for a given $\Psi$ by making $\alpha$ as small as possible consistent with (1). In particular for the ground state (recall $E_0 = -\infty$ ) this means $\alpha = -\infty$ whence it is easy to see (Problem: Prove this), that the right hand side of (2) becomes $\tilde{E}$ , i.e. the familiar variational upper bound.

If we put $\alpha = \ddot{\alpha}$ where

$$\ddot{\alpha} = \frac{\overset{\vee}{\mathcal{E}}_{n-1} + \dot{\overset{\vee}{E}}_n}{2} \tag{45-3}$$

where $\overset{\vee}{\mathcal{E}}_{n-1}$ is an upper bound to $\mathcal{E}_{n-1}$ and where

$$\dot{\overset{\vee}{E}}_n = \ddot{\alpha} + \sqrt{\tilde{E}^2 - 2\ddot{\alpha}\tilde{E} + \ddot{\alpha}^2} \tag{45-4}$$

Then one finds

$$\dot{\overset{\vee}{E}}_n = \frac{\tilde{E}^2 - \tilde{E}\,\overset{\vee}{\mathcal{E}}_{n-1}}{\tilde{E} - \overset{\vee}{\mathcal{E}}_{n-1}} \tag{45-5}$$

or

$$\dot{\overset{\vee}{E}}_n = \tilde{E} + \frac{\tilde{E}^2 - \tilde{E}^2}{\tilde{E} - \overset{\vee}{\mathcal{E}}_{n-1}} \tag{45-6}$$

provided that

$$\tilde{E} - \mathcal{E}^{\vee}_{n-1} > 0 \tag{45-7}$$

These results are due to Kohn (1947) and Kato (1949) (op. cit.). See also Goodisman (Theoret. Chem. Acta. 4, 343 (1966)).

Problem: Fill in the details of the derivation of (5) and (7)

Problem: Show that $(\tilde{\Psi}, (H-E_n)(H-E_{n-1})\tilde{\Psi}) \geq 0$
Show that when (7) is satisfied then this implies (5).

Problem: Show that $\alpha = \overset{\circ}{\alpha}$ is a "best choice" by discussing what happens if one puts $\alpha = \alpha''$ where

$$\alpha'' = \frac{\mathcal{E}^{\upsilon}_{n-1} + \mathcal{E}^{\upsilon}_{n}}{2}$$

with

$$\mathcal{E}^{\upsilon}_{n} < \overset{\circ}{E}{}^{\upsilon}_{n}$$

Problem: Show that $\overset{\circ}{E}{}^{\upsilon}_{n}$ is a variational bound.

Equation (5), with (7), provides an alternative method, different from the linear variational method, for getting upper bounds. It has the advantage over the linear variational method that all one needs is a $\tilde{\Psi}$ and an $\mathcal{E}^{\upsilon}_{n-1}$ satisfying (7), calculates a few integrals (but one of then is $\tilde{E}$) and one is in - there are no secular equations to solve. Note also that in contrast to the corresponding lower bound, it is potentially self contained. Given $\overset{\circ}{E}{}^{\upsilon}_{1}$ (which, as we have seen is best calculated from the usual variational principle) one can use this as $\mathcal{E}^{\upsilon}_{1}$ to calculate $\overset{\circ}{E}{}^{\upsilon}_{2}$ etc. etc.

However in contrasting (5) with the linear variational method, the following remark should also be made. Suppose we take for $\tilde{\Psi}$ the function $\hat{\Psi}_n$ - the n-th solution derived from a linear variational calculation Since $\tilde{E} = \hat{E}_n$ is an upper bound to $E_n$, (7) can easily be satisfied, but then we see from (6) that

$$\overset{\circ}{E}{}^{\upsilon}_{n} > \hat{E}_n$$

i.e. in this case (5) gives a worse bound than the linear variational method.

One can, of course, also use (5) variationally, i.e. given $\mathcal{E}^{U}_{n-1}$ choose a best $\tilde{\Psi}$ from a set of trial functions by minimizing $\tilde{E}^{U}_{n}$ subject to (7). Calculations of this sort have been carried out by Goodisman, and by Goodisman and Secrest for $H_2^+$ and for $H_2$ (J. Chem. Phys. 43, 2806 (1965), J. Chem. Phys. 45, 1515 (1964) and Theoret. Chem. Acta. 4, 343 (1966) with fair results. We refer the reader to their paper for details and numbers.

An interesting question in this connection is the following: Suppose we would use a linear space of trial functions, and suppose that we would use the same set of functions in the linear variational method. Which would yield the better results? Some calculations relevant to the question are reported in the third reference above, which also gives some general examples to show that sometimes (5) may be better. Thus consider the first excited state and put $E_1^V = E_1$ . Then suppose that the basis set (unbeknowest to us) consists of $\Psi_2 + a\Psi_1$ , and $\Psi_3$ where $a$ is a constant. Then clearly (Problem: Fill in the details) the second root of the linear variational method will be $E_3$ while the lowest root arising from (5)$^V$ will be $E_2$ . Hence in this CASE (5) is superior. However one can show (M. Barnsley; Private communication) that if one chooses $\mathcal{E}^{U}_{n-1} = \hat{E}_{n-1}$ then $\tilde{E}^{V}_{n} \geq \hat{E}_n$ , i.e. under these circumstances the linear variational method will be no worse and will usually be better.

Problem: Suppose that one knows not only that $E^{U}_{n-1}$ is an upper bound to $E_{n-1}$ but that in fact it lies between $E_{k-1}$ and $E_k$ where $k \geq m$. Show that then $E^{U}_{m}$ is an upper bound to $E_k$.

Problem: Make detailed contast with the Lehmann-Maehly procedure by looking at the details of what happens when one uses a linear space of trial functions and (5).

FOR AN SCF Procedure based on the TKK bounds see a series of papers by Messmer and Bruss, Theor. Chim. Acta 14, 192-205 (1969).

## 46. Perturbation Theory - The Rebane Bounds

Partly to fill the gap which we encountered in our discussion in Sec. 17 Part E, but mainly simply to provide lower bounds to perturbation energies, we will now consider the perturbation expansion of the TKK upper and lower bounds Eqs. (43-6) and (45-5), which we will now write is

$$\tilde{E} - \frac{\tilde{E^2} - \tilde{E}^2}{\Delta} \leq E_m \leq \tilde{E} + \frac{\tilde{E^2} - \tilde{E}^2}{\Delta'} \tag{46-1}$$

where, from (43-7) and (45-7) $\Delta$ and $\Delta'$ are chosen as large as possible consistent with

$$0 < \Delta \leq E_{m+1} - \tilde{E} \tag{46-2}$$

and

$$0 < \Delta' \leq \tilde{E} - E_{m-1} \tag{46-3}$$

We now make a perturbation expansion of these results. Writing

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

$$\tilde{\psi} = \psi_n^{(0)} + \nu \tilde{\psi}^{(1)} + \cdots$$

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

where $\psi_n^{(0)}$ , which of course satisifes

$$H^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)} ,$$

Assumed
is normalized so that

$$E_n^{(1)} = (\psi_n^{(0)}, H^{(1)}\psi_n^{(0)})$$

Inserting these expansions into the three members of (1) one then finds (A) that $\widetilde{E^{\nu}} - \widehat{E}^{\nu}$ is of second order in $\nu$ (B) that the zero and first order terms are identical in the three members. Then deleting these zero and first order terms, dividing by $\nu^2$ and letting $\nu \to 0$ we arrive at bounds for $E_n^{(2)}$ . In detail they are (Problem: Derive them)

$$A - \frac{B}{\Delta^{(0)}} \leq E_n^{(2)} \leq A + \frac{B}{\Delta^{(0)}} \qquad (46\text{-}3)$$

where A is the familiar Hylleraas functional

$$A = (\tilde{\psi}^{(1)}, (H^{(0)} - E_n^{(0)})\tilde{\psi}^{(1)}) + (\tilde{\psi}^{(1)}, (H^{(1)} - E_n^{(1)})\psi_n^{(0)}) + (\psi_n^{(0)}, (H^{(1)} - E_n^{(1)})\tilde{\psi}^{(1)}) \qquad (46\text{-}4)$$

and B is given by

$$B = \left( (H^{(0)} - E_n^{(0)})\tilde{\psi}^{(1)} + (H^{(1)} - E_n^{(1)})\psi_n^{(0)}, (H^{(0)} - E_n^{(0)})\tilde{\psi}^{(1)} + (H^{(1)} - E_n^{(1)})\psi_n^{(0)} \right) \qquad (46\text{-}5)$$

Further, from (2) and (3) $\Delta^{(0)}$ and $\Delta'^{(0)}$ are restricted by

$$0 < \Delta^{(0)} \leq E_{n+1}^{(0)} - E_n^{(0)} \tag{46-6}$$

$$0 < \Delta'^{(0)} \leq E_n^{(0)} - E_{n-1}^{(0)} \tag{46-7}$$

Clearly to get the best bounds one should use the largest values of $\Delta^{(0)}$ and $\Delta'^{(0)}$ consistent with these restrictions. Therefore if one knows the positions of the zero order levels one should simply use $\Delta^{(0)} = E_{n+1}^{(0)} - E_n^{(0)}$ and $\Delta'^{(0)} = E_n^{(0)} - E_{n-1}^{(0)}$. For the ground state (7) yields $\Delta'^{(0)} \leq \infty$ whereas choosing $\Delta'^{(0)} = \infty$ (4) yields the Hylleraas bound. (Problem: Prove this)

The lower bound for the ground state was first given by Prager and Hirschfelder, J. Chem. Phys. 39, 3289 (1963). The general upper and lower bounds were first given by Rebane, Opt. Spectry. 21, 66 (1966), actually in a somewhat more general form to be described below

Problem: By writing $\tilde{\psi}^{(1)} = \psi_n^{(1)} + \delta$ verify the inequalities (4) directly. Note that this will involve using TKK bounds for $E_n^{(0)}$.

Problem: Write out the equations that one gets by maximizing (Minimizing) the lower (upper) bound (3) over a linear space of trial functions. Show that the lower bound is the same as that given by Miller, J. Chem. Phys. 50, 2758 (1969) Eq. (16).

Problem: Go one step further in the perturbation expansion. That is write $\tilde{\psi} = \psi_n^{(0)} + \nu \psi_n^{(1)} + \nu^2 \tilde{\psi}_n^{(2)} + \cdots$ and derive bounds on $E_n^{(2)}$.

A comment on notation. Evidently $E_{n+1}^{(0)}$ need not be the levels immediately above and below $E_n^{(0)}$. More precisely they are the nearest levels above and below $E_n^{(0)}$ which can mix with $\psi_n^{(0)}$ in first order.

Although one can, by suitable reidentification derive the general Rebane bounds from (3) we will not attempt to do so here (see the brief remark below) but will merely state them

Consider a quantity $\mathcal{E}_m(\epsilon)$ of the form

$$\mathcal{E}_m(\epsilon) \equiv \sum_k \frac{|(\psi_k^{(0)}, U \psi_m^{(0)})|^2}{\epsilon - E_k^{(0)}} \tag{46-7}$$

With suitable identification of $U$ and of $\epsilon$ this could be a second order energy but it could also be a part of the frequency dependent polarizibility. The Rebane result is then that

$$Q - \frac{B}{\delta^{(0)}} \leq \mathcal{E}_m(\epsilon) \leq Q + \frac{B}{\delta'^{(0)}} \tag{46-8}$$

where

$$Q = (\tilde{\psi}^{(1)}, (H^{(0)} - \epsilon)\tilde{\psi}^{(1)}) + (\tilde{\psi}^{(1)}, U\psi_m^{(0)}) + (\psi_m^{(0)}, U^\dagger \tilde{\psi}^{(1)}) \tag{46-9}$$

$$B = (\,(H^{(0)} - \epsilon)\tilde{\psi}^{(1)} + U\psi_m^{(0)}\,,\,(H^{(0)} - \epsilon)\tilde{\psi}^{(1)} + U\psi_m^{(0)}\,) \tag{46-10}$$

and where $\delta^{(0)}$ and $\delta'^{(0)}$ are restricted by

$$0 < \delta^{(0)} \leq E_+^{(0)} - \epsilon \tag{46-11}$$

$$0 < \delta'^{(0)} \leq \epsilon - E_-^{(0)} \tag{46-12}$$

$E_+^{(0)}$ being that $E_k^{(0)}$ which is immediately above $\epsilon$ and $E_-^{(0)}$ being that $E_k^{(0)}$ which is immediately below $\epsilon$. Evidently to get the best

bounds one wants to choose $\delta^{(0)}$ and $\delta^{(1)}$ as large as possible consistent with (11) and (12).

Applications of these results may be found in the following references:

(i) Formal applications to various quantities of interest

(a) Probability of two photon emission and first order corrections to matrix elements Braun and Rebane, Optic and Spectr. 22, 275 (1967).

See also Dmitriev and Yuriev, Int. J. W. Chem. 1, 321 (1967).

(ii) Numerical calcuations

(a) Static polarizibility of H atom, Rebane, Optic and Spec. 21, 66 (1966).

(b) Dynamic polarizibility of H atom, Adamov et al., Optic and Spec. 24, 353 (1968).

(Good agreement of upper and lower bounds for small $\omega$ , poorer for higher $\omega$ )

(c) $C_6$ for H-H, Adamov et al., Int. J. Q. Chem. 3, 57 (1969) Here upper and lower bound agreed to 6 figures to the right of the decimal point.

(d) Polarizibility of an F center, Adamov et al., Sov. Phys., Sol. State 8, 2541 (1967).

Problem: Verify (8) by writing $\tilde{\psi}^{(1)} = \chi + \gamma$ etc. where

$$(H^{(0)} - \epsilon_1)\chi + U\psi_1^{(0)} = 0$$

We want to indicate here a connection between (3) and (8) for the special case in which $(\psi_m^{(0)}, U\psi_m^{(0)}) = 0$ and in which $U$ is

Hermitian. Then the essential observation is that $\mathcal{E}_m(\epsilon)$ has the form of a second order energy with $V$ the perturbation for a problem in which the zero order state is $\psi_m^{(0)}$ but in which the zero order Hamiltonian is

$$\mathcal{H}^{(0)} = H^{(0)} + (\epsilon - E_n^{(0)}) |\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|$$

so that the energy of $\psi_n^{(0)}$ is $\epsilon$, but so that the energy of $\psi_k^{(0)}$, $k \neq 0$ is still $E_k^{(0)}$, i.e.

$$\mathcal{H}^{(0)}\psi_m^{(0)} = \epsilon\psi_m^{(0)} \quad , \quad \mathcal{H}^{(0)}\psi_k^{(0)} = E_k^{(0)}\psi_k^{(0)} \quad k \neq m$$

Thus depending on the value of $\epsilon$, $\psi_m^{(0)}$ may be the ground state of $\mathcal{H}^{(0)}$ the first excited state etc. etc. If one now applies (4) to this new problem, i.e. (4) with a suitable reinterpretation of the symbols, one is led to (8).

Problem: Fill in the details

Problem: Show that applied to frequency dependent polarizibilities below the first resonance the best upper bound from (8) is identical to the bounds found in Sec. 30.

## 47. Operator Comparison Methods - Introduction

The basic theorem (due to Weyl) is the following: Suppose that we write our Hamiltonian $H$ as

$$H = H^I + H' \tag{47-1}$$

where $H'$ is a positive Hermitian operator, i.e.

$$(\tilde{\Psi}, H'\tilde{\Psi}) = (H'\tilde{\Psi}, \tilde{\Psi}) \geq 0 \quad \text{all } \tilde{\Psi} \tag{47-2}$$

One then says that $H^I$, the "intermediate Hamiltonian", is less than $H$, or in symbols

$$H^I < H \tag{47-3}$$

Let us now denote the (ordered) descrete eigenvalues of $H^I$ by $E_1^I$, $E_2^I$, $E_3^I$ . . . and continue to denote those of $H$ by $E_1, E_2, \ldots$ Then one can show that

$$E_i^I \leq E_i \tag{47-4}$$

That is the $E_i^I$ in order furnish lower bounds to the $E_i$ in order.

Problem: Derive (4) from the MAX MIN principle.

If we write $H = {}^I H + {}'H$ where ${}'H$ is a negative Hermitian operator then of course, in obvious notation, we will have $E_i \leq {}^I E_i$, i.e. the ${}^I E_i$ in order furnish upper bounds to the $E_i$ in order.

As a trivial example we may take $H'$ to be the coulomb interaction between the electrons. Then the theorem says that the eigenvalues of the non-interacting electron problem yields lower bounds on the real problem. However this particular lower bound is usually quite crude. Thus for the Helium atom $E_1^I = -4$ a.u. while $E_1 \sim -3.4$ a.u.

However the example does emphasize, that in contrast to the methods discussed in the earlier sections, the present approach is quite self contained. The main problem in its implementation is to find $H^I$'s which are on the one hand soluble so that one can explicitly exhibit the $E_i^I$, or at least give a finite procedure for finding them, and which, on the other hand give interestingly close bounds.

This technique of operator comparision was introduced into atomic and molecular physics by Bazley (Phys. Rev. 120, 144 (1959)) based on earlier work in a more general setting by A. Weinstein and N. Aronszajn. Bazley and Fox and others have continued work on this and more general problems and extensive reviews may be found in the book by Gould - "Variational Methods for Eigenvalue Problems" and in two recent reviews, one by Fox and Rheinboldt in Siam Review 8, 427 (1966) and another by Bazley and Fox in J. Reine Angew. Math. 223, 142 (1966).

As we will discuss in a later section the $H^I$'s which have been used by these authors in the atomic and molecular context are in essence generalizations of the simple $H^I$ we discussed above. Recently however W. Miller (J. Chem. Phys. 50, 12758 (1969) has suggested a different $H^I$ which has very definite advantages over the previous suggestions. Hence we will discuss it first in some detail in the next section and then in subsequent sections will summarize the work of other authors.

A note on the error in operator comparison methods: If we denote the n-th eigenfunction of $H^I$ by $\psi_n^I$ then our lower bound can be written

$$E_n^I = \frac{(\psi_n^I, H^I \psi_n^I)}{(\psi_n^I, \psi_n^I)} = \frac{(\psi_n^I, H \psi_n^I)}{(\psi_n^I, \psi_n^I)} - \frac{(\psi_n^I, H' \psi_n^I)}{(\psi_n^I, \psi_n^I)}$$

Now we know that the first term on the right hand side differs from $E_n$ by terms of order $(\psi_n - \psi_n^I)^2$ so we have

$$E_n^I = E_n + O(H', (\psi_n - \psi_n^I)^2)$$

## 48. The Miller Intermediate Hamiltonian

Our work on upper bounds for frequency dependent polarizibilities in Sec. 30 can be summarized in the statement (see also Rosenberg et al. Phys. Rev. 118, 184 (1960)) that if the $\hat{\phi}_i$ with $i = 1 \cdots L$ are an orthonormal set of functions which diagonalize a Hermitian operator $A$ (in our earlier work $A$ was the zero field Hamiltonian), the diagonal elements being $\hat{a}_i$, then

$$(A-\epsilon)^{-1} > \sum_{i=1}^{L} \frac{|\hat{\phi}_i\rangle\langle\hat{\phi}_i|}{\hat{a}_i - \epsilon} \equiv Q \tag{48-1}$$

provided that there are as many $\hat{a}_i$ below $\epsilon$ as there are eigenvalues of $A$ below $\epsilon$. Multiplying (1) from the right and the left by $(A-\epsilon)$ we then have that

$$(A-\epsilon)\, Q\, (A-\epsilon) + \epsilon < A \tag{48-2}$$

We now apply this general result as follows. First we split $H$ according to

$$H = \mathcal{H}_0 + V \tag{48-3}$$

The special properties that we will require of $\mathcal{H}_0$ and $V$ will be made clear shortly. Now we apply (2) with $A = \mathcal{H}_0$ to find that

$$(\mathcal{H}_0 - \epsilon)\,\Theta\,(\mathcal{H}_0 - \epsilon) + \epsilon + V < H \qquad (48\text{-}4)$$

where

$$\Theta = \sum_{i=1}^{M} \frac{|\hat{\Phi}_i\rangle\langle\hat{\Phi}_i|}{\hat{\mathcal{E}}_i - \epsilon} \qquad (48\text{-}5)$$

if there are as many $\hat{\mathcal{E}}_i$ below $\epsilon$ as there are eigenvalues of $\mathcal{H}_0$. (In order to know this, of course, we must know something about the spectrum of $\mathcal{H}_0$. Hence this is a restriction on our choice of $\mathcal{H}_0$).

Evidently we can now use the left hand side of (4) as our intermediate Hamiltonian. We will denote it by $\mathcal{H}^M(\epsilon)$ and we will denote its eigenvalues by $\mathcal{E}_i^M(\epsilon)$. We will now show that if $V$ is simply a function, and most simply a positive function though this can be relaxed, then we can find the $\mathcal{E}_i^M$ (more precisely some of the $\mathcal{E}_i^M$) by solving an $M \times M$ matrix problem. Thus $\mathcal{H}^M$ is a suitable intermediate Hamiltonian in that it is soluble. Further it is reasonably clear that since by increasing $M$ we can approach arbitrarily close to completeness of the basis function, (1) can come arbitrarily close to equality and the $\mathcal{E}_i^M$ can come arbitrarily close to $E_i$. Indeed one can prove that $\mathcal{H}^{M'}(\epsilon)$ is $> \mathcal{H}^M(\epsilon)$ if $M' > M$. Thus $\mathcal{H}^M$ is also suitable in that it can give arbitrarily close bounds.

The eigenvalue problem for $\mathcal{H}^M$ is evidently

$$\sum_{i=1}^{M} \frac{(\mathcal{H}_0 - \epsilon)\,\hat{\Phi}_i\,\langle\hat{\Phi}_i|(\mathcal{H}_0 - \epsilon)\Psi^M\rangle}{\hat{\mathcal{E}}_i - \epsilon} + (\epsilon - \mathcal{E}^M + V)\Psi^M = 0 \qquad (48\text{-}6)$$

Now if $V$ is a positive function then even if it ranges from 0 to $\infty$ still $(\epsilon - \mathcal{E}^M + V)^{-1}$ will be well defined if

$$\mathcal{E}^M \leq \epsilon \tag{48-7}$$

We now restrict ourselves to this case, (Thus we are finding only those eigenvalues of $\mathcal{H}^M$ which lie below $\epsilon$) so that we can replace (6) by

$$\psi^M = - \sum_{i=1}^{M} (\epsilon - \mathcal{E}^M + V)^{-1} (\mathcal{H}_0 - \epsilon) \hat{\Phi}_i \frac{(\hat{\Phi}_i, (\mathcal{H}_0 - \epsilon)\psi^M)}{(\hat{\mathcal{E}}_i - \epsilon)} \tag{48-8}$$

Further since $V$ is simply a function, the inverse operator is not difficult to deal with since it too is simply a function. If now we take the scalar product of this equation with $(\mathcal{H}_0 - \epsilon)\hat{\Phi}_j$ we find a set of linear homogeneous equations for

$$\zeta_i \equiv \frac{(\Phi_i, (\mathcal{H}_0 - \epsilon)\psi^M)}{\hat{\mathcal{E}}_i - \epsilon} \tag{48-9}$$

In detail these equations are

$$(\hat{\mathcal{E}}_j - \epsilon)\zeta_j = - \sum_{i=1}^{M} (\hat{\Phi}_j, (\mathcal{H}_0 - \epsilon)(\epsilon - \mathcal{E}^M + V)^{-1} (\mathcal{H}_0 - \epsilon)\hat{\Phi}_i)\zeta_i \tag{48-10}$$

Whence we have as the equation to determine the $\mathcal{E}_i^M$

$$\left| (\hat{\mathcal{E}}_j - \epsilon)\delta_{ij} + (\hat{\Phi}_j, (\mathcal{H}_0 - \epsilon)(\epsilon - \mathcal{E}^M + V)^{-1} (\mathcal{H}_0 - \epsilon)\hat{\Phi}_i) \right| = 0 \tag{48-11}$$

or

$$\left| (\hat{\Phi}_j, [(\mathcal{H}_0 - \epsilon) + (\mathcal{H}_0 - \epsilon)(\epsilon - \mathcal{E}^M + V)^{-1} (\mathcal{H}_0 - \epsilon)]\hat{\Phi}_i) \right| = 0 \tag{48-12}$$

Equation (11) is special to the $\hat{\Phi}_i$ basis. However in the form (12) it is immediately applicable to any basis $\chi_i$ derived by a non-singular linear transformation of the $\hat{\Phi}_j$, that is in an arbitrary basis we have

$$\left| (\chi_j, \Sigma (\mathcal{H}_0 - \epsilon) + (\mathcal{H}_0 - \epsilon)(\epsilon - \mathcal{E}^M + V)^{-1} (\mathcal{H}_0 - \epsilon) | \chi_i) \right| = 0 \quad (48\text{-}13)$$

Problem: Prove this. Hint: Recall that the determinant of a product of matrices equals the product of the determinants.

Evidently even if $V$ is not positive, still it may be possible that $(\epsilon - \mathcal{E}^M + V)^{-1}$ exists and hence that the procedure can be carried through. See Wilson, J. Chem. Phys. 43, 5172 (1965) for some examples of a similar situation.

Thus we have a finite procedure for producing arbitrarily good lower bounds to the $E_i$ below $\epsilon$ provided that there are as many $\hat{\mathcal{E}}_i$ below $\epsilon$ as there are eigenvalues of $\mathcal{H}_0$. Further it is not hard to show that for a given basis set one will get the best bounds by taking $\epsilon$ as large as possible consistent with this requirement. Namely from the Hellmann-Feynman theory applied to $\mathcal{H}^M$ we have

$$\frac{\partial \mathcal{E}_i^M}{\partial \epsilon} = (\Psi_i^M, \frac{\partial \mathcal{H}^M}{\partial \epsilon} \Psi_i^M) \Big/ (\Psi_i^M, \Psi_i^M)$$

A short calculation then yields (Problem: Derive this).

$$\frac{\partial \mathcal{H}^M}{\partial \epsilon} = 1 - \sum_{i=1}^{M} |\hat{\Phi}_i \rangle \langle \hat{\Phi}_i| + \sum_{i=1}^{M} \frac{(\mathcal{H}_0 - \hat{\mathcal{E}}_i)|\hat{\Phi}_i\rangle\langle\hat{\Phi}_i|(\mathcal{H}_0 - \hat{\mathcal{E}}_i)}{(\hat{\mathcal{E}}_i - \epsilon)^2}$$

whence clearly (Problem: Fill in the details).

$$\frac{\partial \mathcal{E}_i^M}{\partial \epsilon} > 0$$

We refer the reader to Miller's paper (J. Chem. Phys. 50, 2758 (1969) for more details of the theory and for a successful application to Helium.

Miller also discusses the use of (2) with $A=H$, i.e. $V=0$. Show that the analog of (12) is our old friend (44-16) and that Miller's condition on $\epsilon$ implies that $N=0$ in (44-18). Show further from (44-6) that although the roots which are $>\epsilon$, roots which the Miller method does not interpret (for $V=0$ one cannot use (47-4) to conclude anything about the $\varepsilon_{M}^{i}$ for $\varepsilon_{M}^{i}>\epsilon$ because $\varepsilon^{M}=\epsilon$ is an infinitely degenerate eigenvalue of $\mathcal{H}^{M}$), can now be seen as upper bounds, they are in fact worse than the corresponding $E_i$.

The results of applying perturbation theory to $\mathcal{H}^{M}$ have already been mentioned in a problem following Eq. (46-7). In essence one finds the Rebane bound.

## 49. Summary of Work with Other Comparison Hamiltonians

Assuming that $V$ is positive (though it need not be simply a function) one can introduce $V^N$, the N-th order truncation of $V$, according to

$$V^N = \sqrt{V} \sum_{k=1}^{N} |\zeta_k\rangle\langle\zeta_k| \sqrt{V} \tag{49-1}$$

where the $\zeta_k$ are some discrete orthonormal basis set. Evidently (Problem: Prove this).

$$\cdots V^N < V^{N-1} \cdots < V \tag{49-2}$$

From (2) then an obvious choice for an intermediate Hamiltonian is $H_0 + V^N$. However in general one cannot find its eigenvalues in a finite way. An exception occurs if $\sqrt{V}\,\zeta_k$ can be written as a finite sum of the

discrete eigenfunction of $\mathcal{H}_0$ . Under these circumstances, as shown by Bazley (Phys. Rev. 120, 144 (1960) and we refer the reader to his papers for details of the general theory and for the results of an Application to Helium) ↙ one can find the eigenvalues of $\mathcal{H}_0 + V^N$ (*) The weakness of the method is that usually the discrete eigenfunctions ↙ of $\mathcal{H}_0$ are not complete whence the requirement on the $\sqrt{V}\,\zeta_k$ will usually mean that one can't let $N \to \infty$ so that the $\gamma_k$ can become complete, and hence ↙ one can't make the intermediate Hamiltonian be arbitrarily close to $H$ .

Incompleteness aside the above requirement on $\sqrt{V}\,\zeta_k$ is anyway very restrictive and to overcome it Bazley and Fox (Phys. Rev. 124, 483 (1961)) suggested a further truncation, this time of $\mathcal{H}_0$ . Namely denoting the eigenfunctions and eigenvalues of $\mathcal{H}_0$ by $\psi_i^0$ and $E_i^0$, we have

$$\mathcal{H}_0 = \sum_i E_i^0 \, |\psi_i^0\rangle\langle\psi_i^0| \tag{49-3}$$

whence we can introduce a truncated operator $\mathcal{H}_0^L$ according to

$$\mathcal{H}_0^L = \sum_{i=1}^{L} E_i^0 \, |\psi_i^0\rangle\langle\psi_i^0| + E_{L+1}^0 \sum_{k=L+1}^{\infty} |\psi_k^0\rangle\langle\psi_k^0| \tag{49-4}$$

$$= \sum_{i=1}^{L} (E_i^0 - E_{L+1}^0) \, |\psi_i^0\rangle\langle\psi_i^0| + E_{L+1}^0 \tag{49-5}$$

Evidently (Problem: Prove this)

$$\cdots \mathcal{H}_0^L < \mathcal{H}_0^{L+1} < \cdots < \mathcal{H}_0 \tag{48-6}$$

(*) Note: Bazley gives an equivalent but rather different looking formula for $V^N$

Bzley and Fox then use

$$H^{L,N} \equiv \mathcal{H}_0^L + V^N \tag{48-7}$$

as an intermediate Hamiltonian and show that its eigenvalues can be found by finite procedures. The difficulty is that although we can now let $N \to \infty$ we cannot in general let $L \to \infty$, the point being that while in writing (3) we took the customary license of using as discrete notation, in (5) it is essential that the sum be a sum and not an integral since if it involves an integral we are led to integral equations instead of algebraic equations.

Miller (J. Chem. Phys. 42, 4305 (1965) then pointed out that if $V^N$ is simply a function then one can eliminate the truncation of V and still find those eigenvalues of the $H^{L,\infty} = \mathcal{H}_0^L + V$ problem which are less than $E^0_{L+1}$ by solving an $L \times L$ algebraic problem. Indeed the method is just a special case of that of the previous section if one puts $\hat{\Phi}_i = \psi_i^0$ (and $N \to \infty \times \infty$) and $\epsilon = E^0_{L+1}$. However again lack of completeness of the discrete $\psi_i^0$ prevents complete convergence. For applications see the paper by Miller, and also Jenning, J. Chem. Phys. 46, 2442 (1967).

If one applies perturbation theory to $H^{L,\infty}$ with $V$ as the perturbation (Problem: Do this) then one finds the simple and obvious results

$$-\sum_{k=1}^{L} \frac{|(\psi_i^0, V\psi_k^0)|^2}{E_k^0 - E_i^0} - \frac{\sum_{j=L+1}^{\infty} |(\psi_i^0, V\psi_j^0)|^2}{E^0_{L+1} - E_i^0} \leq E_i^{(2)}$$

for $i \leq L+1$. For some discussion of such bounds see Goodisman, J. Chem. Phys. 47, 2707 (1967).

Bazley and Fox have suggested writing

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

and using $H^{L,0}$ as $H_0$, and $H - H^{L,0}$ as $V$. Formally this overcomes the convergence problem - the eigenfunctions of $H^{L,0}$ can be taken to form a discrete set, the eigenvalue $E^0 = E^0_{L,1}$ being infinitely degenerate. However so far no applications seem to have been made.

For one-electron problems it is possible to overcome the convergence problem in another way (See Walmsley and Coulson Proc. Camb. Phil. Soc. 62, 769 (1966), Walmsley, Proc. Camb. Phil. Soc. 63, 451 (1967)), namely one treats the (negative) energy as given and lets the potential strength be the eigenvalue. For a given $E$, the potential strength eigenvalues are then all discrete for a one-electron problem and one is in business. (Problem: What happens with more than one electron?)

The procedures discussed in this section all require that one know the $\psi_i^0$. However if $H$ is a molecular problem and $H_0$ involves several centers of potential then we can't meet the requirement. In these circumstances Bazley and Fox (J. Math. Phys. 4, 1147 (1963)) suggested a further truncation which we will illustrate using $H_2^+$ as an example. Then

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{r_a} - \frac{1}{r_b}$$

$$= \left(-\frac{1}{2}\nabla_a^2 - \frac{1}{r_a}\right) + \left(-\frac{1}{2}\nabla_b^2 - \frac{1}{r_b}\right) \equiv H_A + H_B$$

Now $H_A$ and $H_B$ are hydrogenic on centers A and B respectively and one knows all about them. One now uses as an intermediate Hamiltonian

$$H^L_A + H^N_B$$

For applications see the Walmsley papers referenced above and also Johnson and Coulson (Proc. Phys. Soc. 84, 263 (1964)).

## 50. Energy Dependent Truncation - The Method of Gay and Löwdin

In the previous sections we mentioned one set of circumstances under which $\mathcal{H}_0 + V^N$ becomes soluble, namely if we can choose the $\zeta_k$ so that $\sqrt{V}\,\zeta_k$ can be written as a finite linear combination of the discrete $\Psi_i^0$ (the Bazley "special choice"). Löwdin has pointed out another, more flexible possibility and it has been investigated in detail by Gay (Phys. Rev. 135, 4 1220 (1964)).

Namely suppose we introduce functions $\xi_k$ according to

$$\zeta_k = \frac{1}{\sqrt{V}} (\mathcal{H}_0 - E^{(\nu)}) \xi_k \Big/ N_k \tag{50-1}$$

where $E^{(\nu)}$ is the eigenvalue which we are seeking and where (remember the $\zeta_k$ are normalized)

$$N_k = (\xi_k, (\mathcal{H}_0 - E^{(\nu)}) V^{-1} (\mathcal{H}_0 - E^{(\nu)}) \xi_k) \tag{50-2}$$

Then one finds that for $E^{(\nu)}$ not equal to an $E_i^0$, so that $(\mathcal{H}_0 - E^{(\nu)})^{-1}$ exists, one can write the eigenvalue equation for $\mathcal{H}_0 + V^N$ as (Problem: Derive this result)

$$\Psi^{(\nu)} = -\sum_{k=1}^{N} \xi_k (\xi_k, (\mathcal{H}_0 - E^{(\nu)}) \Psi^{(\nu)}) / N_k \tag{50-3}$$

If now we take the scalar product of (3) with $(\mathcal{H}_0 - E^{(\nu)}) \xi_L$ we are led to a set of $N$ linear homogeneous equations for the quantities

$$(\xi_k, (\mathcal{H}_0 - E^{(\nu)}) \Psi^{(\nu)}) / N_k$$

and finally to the following equation for $E^{(\nu)}$ (Problem: Fill in the details)

$$0 = \left| \left(\xi_\lambda, (\mathcal{H}_0 - E^{(\nu)}) V^{-1} (\mathcal{H}_0 - E^{(\nu)}) \xi_\lambda\right) \delta_{\kappa\lambda} + \left(\xi_\lambda, (\mathcal{H}_0 - E^{(\nu)}) \xi_\kappa\right) \right| \quad (50\text{-}4)$$

We can now write this more compactly and at the same time transcribe it into any basis $\chi_\kappa$ derived from the $\xi_\lambda$ by a non-singular linear transformation if we note that since the $\xi_\kappa$ were to be orthogonal, (1) implies that

$$\left(\xi_\kappa, (\mathcal{H}_0 - E^{(\nu)}) V^{-1} (\mathcal{H}_0 - E^{(\nu)}) \xi_\lambda\right) = 0 \qquad \kappa \neq \lambda$$

From which it follows that (4) can be written (Problem: Derive this result)

$$\left| \left(\chi_\kappa, \left[ (\mathcal{H}_0 - E^{(\nu)}) V^{-1} (\mathcal{H}_0 - E^{(\nu)}) + (\mathcal{H}_0 - E^{(\nu)}) \right] \chi_\lambda\right) \right| = 0 \quad (50\text{-}5)$$

The solutions of (5) will give us those $E^{(\nu)}$ which are not equal to any $\varepsilon_i^0$. In addition one readily sees that with the substitution (3), $E^{(\nu)} = \varepsilon_i^0$ and $\psi^{(\nu)} = \psi_i^0$ for any $i$ solve the eigenvalue problem for $\mathcal{H}_0 + V^{(0)}$.

The main point now is that if $V$ is simply a function then (5) contains no difficult-inverse operators and the game is simply to choose a set of $\chi_\kappa$ and calculate away. Further, there is no obvious problem of convergence - the set of $\chi_\kappa$ can be as complete as we are able to handle computationally. However now several remarks must be made:

(i) (5) has the same form as (48-13) if in the latter we put $\epsilon = E^{(\nu)}$. Now if this is at all a permissible value for $\epsilon$ (if

there are as many $\hat{\varepsilon}_i$ below $\epsilon$ as there are $\varepsilon_i^0$ ) then it is almost certainly not the best choice and probably the worst. Thus as it stands this procedure results in quite similar equations but in general is inferior to that of Sec. 48 as long as $E^{HL}$ is less than $\epsilon$. Nevertheless we are considering it in some detail because of its theoretical interest and also because when used in a different way (see Sec. 54) the question of superiority and inferiority for energies below $\epsilon$ remains unsettled.

(ii) The relation of the solutions of (5) to the $E_i$ is unclear as it stands (this is often called the "ordering problem"), the point being that (47-4) does not apply since our $H^I$ now depends on the particular eigenvalue in question, i.e. we have a different Hamiltonian for each eigenvalue! One way to partially solve the ordering problem is of course to solve (48-13) instead of (5). The solutions of (5) are then the intersections of the solution curves of (48-13) as a function of $\epsilon$ with the straight line $\varepsilon^M = \epsilon$. Then as long as these intersections lie below $\epsilon$ we would know how to interpret them. This is of course an a postiori approach and certainly not very interesting from the point of view of the Gay and Löwdin method itself.

However the situation is not hopeless. If we denote the Gay and Löwdin $H^I$ by $H^{HL}(\epsilon^I)$ then we could completely solve the ordering problem a postiori if we could find the eigenvalues $\varepsilon_i^{HL}(\delta)$ of $H^{HL}(\delta)$ when $\delta$ is a parameter, since then (47-4) would apply to $H^{HL}(\delta)$ for all $\delta$ and the eigenvalues of $E_i^{HL}$ or $H^{HL}(\epsilon^I)$ would be the interaction of the solution curves $\varepsilon_i^{HL}(\delta)$ as a function of $\delta$ with the straight line $\varepsilon^{HL} = \delta$. Now in fact we

can't find the $\mathcal{E}_i^{GL}(\delta)$ but we can use their existence to make the following a priori argument: Suppose that

$$E_j^{GL} < E_{m+1}^{0} \tag{50-6}$$

Then since $E_{m+1}^{0} \leq \mathcal{E}_{m+1}^{GL}(\delta)$ for all $\delta$ it follows from (6) that $E_j^{GL}$ must be an intersection point on a curve below the $\mathcal{E}_{m+1}^{GL}(\delta)$ curve. Hence we have

$$E_j^{GL} \leq \mathcal{E}_n^{GL}(\delta) \leq E_n \tag{50-7}$$

Thus in particular if $E_j^{GL}$ is $< E_1^0$ then it is a lower bound to $E_1$. If in addition one knows that $E_j^{GL}$ lies above certain lower levels then we can sharpen (7) in the other direction. For further discussion see Gay's paper; also Miller, J. Chem. Phys. 48, 530 (1968); Bazley and Fox, Phys. Rev. 148, 90 (1966); and T. M. Wilson, J. Chem. Phys. 47, 4706 (1967).

TKK as a Special Case of GL

Eq. (5) can also be written (Problem: Show this)

$$\left| \left( \chi_k, \left[ (H - E^{GL}) V^{-1} (H - E^{GL}) - (H - E^{GL}) \right] \chi_L \right) \right| = 0$$

Suppose now that the set $\chi_k$ consists of a single function $\tilde{\Psi}$. Then we have

$$\left( \tilde{\Psi}, \left[ (H - E^{GL}) V^{-1} (H - E^{GL}) - (H - E^{GL}) \right] \tilde{\Psi} \right) = 0$$

We now take $V$ to be constant whence it follows that

$$E_{m+1} = E_{m+1}^{0} + V \tag{50-9}$$

Suppose now we can choose $V$ to be (remember is supposed to be positive)

$$V = \mathcal{E}^{L}_{n+1} - E^{(L)} \tag{50-10}$$

where, as usual, $\mathcal{E}^{L}_{n+1}$ is a lower bound to $E_{n+1}$

Then from (2) we have

$$E^{0}_{n+1} = E_{n+1} - V = (E_{n+1} - \mathcal{E}^{L}_{n+1}) + E^{(L)} > E^{(L)} \tag{50-11}$$

Whence it follows from (6) and (7) that $E^{(L)}$ will be a lower bound to $E_n$ .

If now we put (3) into (1) we readily find the result (Problem: Fill in the details)

$$E^{(L)} = \frac{(\hat{\psi}, (\mathcal{E}^{L}_{n+1} - H) H \hat{\psi})}{(\hat{\psi}, (\mathcal{E}^{L}_{n+1} - H) \hat{\psi})} \equiv E^{(L)}_{n} \tag{50-12}$$

we now note that from (12)

$$\mathcal{E}^{L}_{n+1} - E^{(L)} = \frac{(\hat{\psi}, (H - \mathcal{E}^{L}_{n+1}) \hat{\psi})}{(\hat{\psi}, (\mathcal{E}^{L}_{n+1} - H) \hat{\psi})} \tag{50-13}$$

Whence $V$ will be positive provided that

$$(\hat{\psi}, (\mathcal{E}^{L}_{n+1} - H) \hat{\psi}) > 0 \tag{50-14}$$

and we have rederived the TKK lower bound.

## 51. The Löwdin Bracketing Function - One-Dimensional Form

If one introduces a complete discrete basis set $\phi_k$, then the problem of solving the Schrödinger equation

$$H\psi = E\psi$$

is "reduced" to the problem of solving the infinite set of linear equations

$$\sum_{L=1}^{\infty} (\phi_K, (H-E)\phi_L)\, c_L = 0 \qquad (51\text{-}1)$$

Whence one has as the "equation" for the eigenvalues

$$\left| (\phi_K, (H-E)\phi_L) \right| = 0 \qquad (51\text{-}2)$$

Now if one is dealing with a finite set of equations another, and often more efficient procedure to get at the eigenvalue equation (i.e. without having to multiply out the determinant (2)) is by successive elimination; namely, one solves all but one of the equations (1) as inhomogeneous equations for all the $(\phi_L, \psi)$ in terms of a single one, say $(\phi_1, \psi)$ . Then the last equation becomes a homogeneous equation for $(\phi_1, \psi)$ alone, whence the coefficient must vanish and this yields the eigenvalue equation.

Now with an infinite set of equations one can't of course carry out such a procedure sequentially, but one can partially formalize it in the so called partitioning method. Namely we can break up the infinite matrix $(\phi_K, (H-E)\phi_L)$ as follows, and now we will single out $\phi_1$ (sometimes we will call it the "reference function") and we will assume that $\phi_1$ is normalized and is orthogonal to the remaining $\phi_K$ :

$$\left[ (\phi_K, (H-E)\phi_L) \right] = \begin{bmatrix} H_{11}-E & H_{1b} \\ H_{b1} & H_{bb}-E\mathbb{1}_{bb} \end{bmatrix}$$

where $H_{bb}$ is an $(\omega-1)\times(\omega-1)$ dimensional matrix etc. Similarly we write the column vector $(\Phi_a,\Psi)$ as

$$\begin{bmatrix} (\Phi_1,\Psi) \\ (\Phi_b,\Psi) \end{bmatrix}$$

where $(\Phi_b,\Psi)$ is an $(\omega-1)$ element column vector. Then (1) becomes

$$(H_{11}-E)(\Phi_1,\Psi) + H_{1b}(\Phi_b,\Psi) = 0 \tag{51-3}$$

$$H_{b1}(\Phi_1,\Psi) + (H_{bb}-E\,I_{bb})(\Phi_b,\Psi) = 0 \tag{51-4}$$

The elimination procedure now consists in "solving" (4) for $(\Phi_b,\Psi)$ in terms of $(\Phi_1,\Psi)$ namely

$$(\Phi_b,\Psi) = -(H_{bb}-E\,I_{bb})^{-1} H_{b1}(\Phi_1,\Psi) \tag{51-5}$$

If we now insert this into (3) the result is

$$\{(H_{11}-E) - H_{1b}(H_{bb}-E\,I_{bb})^{-1} H_{b1}\}(\Phi_1,\Psi) \tag{51-6}$$

which <u>if</u> $(\Phi_1,\Psi) \neq 0$ , yields the eigenvalue equation (equivalent to (2))

$$E = H_{11} - H_{1b}(H_{bb}-E\,I_{bb})^{-1} H_{b1} \tag{51-7}$$

Following Löwdin (for a review see his article in "Perturbation Theory and its Applications in Quantum Mechanics" C. H. Wilcox ed.

Wiley 1966) we now define the function $f(\mathcal{E})$ by

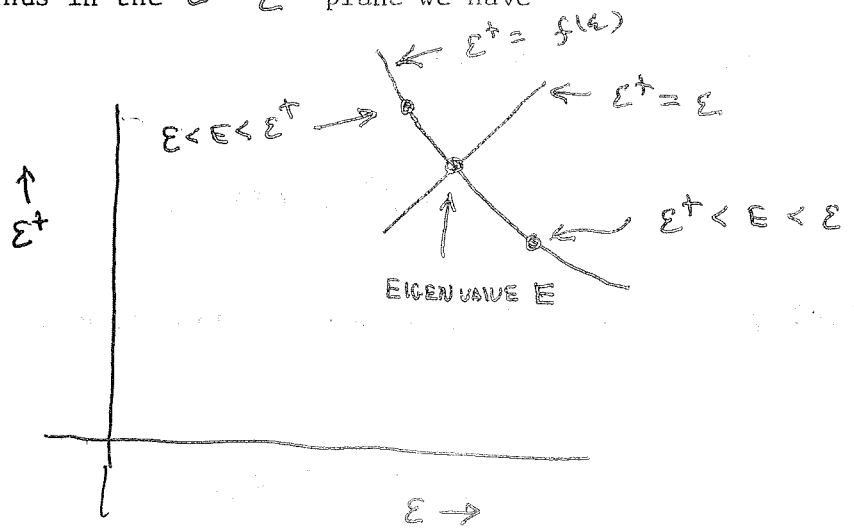
$$f(\mathcal{E}) \equiv H_{11} - H_{1b}\,(H_{bb} - \mathcal{E}\,1_{bb})^{-1} H_{b1} \tag{51-8}$$

that is $f(\mathcal{E})$ is the right hand side of (7) but with some "trial energy" $\mathcal{E}$ instead of $\epsilon$. One then has a solution of (7) when equals $f(\mathcal{E})$ again. More graphically if we introduce an "$\mathcal{E}^+ - \mathcal{E}$ plane" then we can say that the solutions to (7) are the intersections of the graph of $\mathcal{E}^+ = f(\mathcal{E})$ with the straight line $\mathcal{E}^+ = \mathcal{E}$.

Löwdin now observed that $f(\mathcal{E})$ is in fact a bracketing function in that if $\mathcal{E}$ is an upper bound to some eigenvalue then $f(\mathcal{E})$ is a lower bound and conversely. The proof follows from the observation that

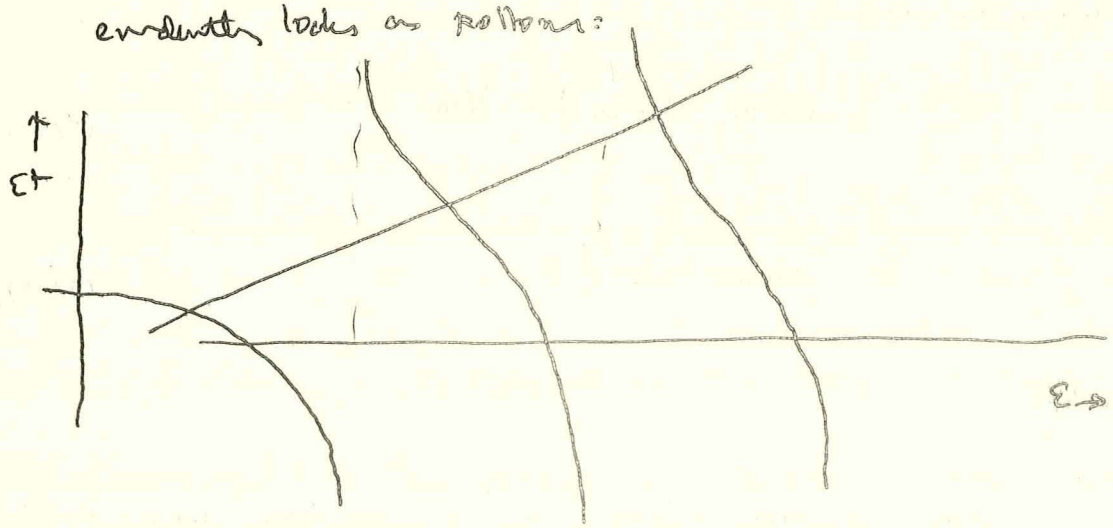
$$\frac{\partial f(\mathcal{E})}{\partial \mathcal{E}} = -H_{1b}\,(H_{bb} - \mathcal{E}\,1_{bb})^{-2} H_{b1} < 0$$

Thus in the $\mathcal{E}^+ - \mathcal{E}$ plane we have



which proves the point. Note however that we again have an ordering problem in that we in general don't know which eigenvalue is being bracketed. However clearly if we know that $\mathcal{E}$ is a lower bound to $E_2$ and if we find that $f(\mathcal{E})$ is $< \mathcal{E}$ then it follows that $f(\mathcal{E})$ must be a lower bound to $E_1$. Similarily if $\mathcal{E}$ is a lower bound to $E_1$, and we find $f(\mathcal{E}) > \mathcal{E}$ then we can say that $f(\mathcal{E})$ is an upper bound to $E_1$. We will discuss the situation for excited states in later sections.

In the figure we have shown only a small section of the graph of $\mathcal{E}^+ = f(\mathcal{E})$. Actually one can quite easily see the qualitative structure of the entire graph. Namely suppose, as we may, that the functions $\phi_2, \ldots$ diagonalize $H$ within the $(\omega-1)$ dimensional basis set. Then $(H_{bb} - \mathcal{E} I_{bb})^{-1}$ will be a diagonal matrix with diagonal elements $(\tilde{E}_k - \mathcal{E})^{-1}$, the $\tilde{E}_k$ being the diagonal elements of the diagonal matrix $H_{bb}$. Thus we see that $f(\mathcal{E})$ will have poles at those $\tilde{E}_k$ for which the corresponding $H_{1k}$ don't vanish. Further we note that as $|\mathcal{E}| \to \infty$, $f(\mathcal{E}) \to H_{11}$ so that an overall view



Problem: Using the above rederive the result that $H_{11}$ is an upper bound to $E_1$ .

Problem: Using the above rederive the result that $\tilde{E}_i$ is upper bound to $E_i$ .

As we noted in the derivation, (7) does not determine the $E_i$'s whose associated wave functions are orthogonal to $\Phi_1$. Now in our graphical approach, if we imagine varying one $H_{1k}$, then clearly we lose an intersection when $H_{1k}=0$ To connect with the first remark show that $H_{1k}=0$ implies that $\Phi_k$ (which is anyway defined to be orthogonal to $\Phi_1$ ) is an eigenfunction of $H$ . (Hint: Show first that in general $H\Phi_k = E_k\Phi_k + \lambda\Phi_1$ where $\lambda$ is some number. Then show that $H_{1k}=0$ implies $\lambda = 0$ .)

## 52. Some Rewritings of the Bracketing Function

In our previous work we often found it useful to split $H$ into

$$H = H_0 + V \tag{52-1}$$

two parts and to treat the two parts rather differently. It will prove convenient to do similar things here. We are interested in

$$T_{bb}^{-1} \equiv E\,|_{bb} - H_{bb} \tag{52-2}$$

which we now write as

$$T_{bb}^{-1} = (E\,|_{bb} - H_{0\,bb}) - V_{bb} \tag{52-3}$$

$$\equiv T_{0\,bb}^{-1} - V_{bb} \tag{52-4}$$

so that

$$T_{bb} = (T_{0bb}^{-1} - V_{bb})^{-1} = (1 - T_{0bb}V_{bb})^{-1} T_{0bb} \tag{52-5}$$

or alternatively

$$T_{bb} = T_{0bb}\,(1 - V_{bb}\,T_{0bb})^{-1} \tag{52-6}$$

Eqs. (5) and (6), which yield two equivalent forms for $T_{bb}$, can often profitably be considered as the solutions of the equations

$$T_{bb} = T_{0bb} + T_{0bb}\,V_{bb}\,T_{bb} \tag{52-7}$$

and

$$T_{bb} = T_{0bb} + T_{bb}\,V\,T_{0bb} \tag{52-8}$$

$T_{0bb}$ and $T_{bb}$ are $(\infty - 1) \times (\infty - 1)$ matrices. It is very convenient to introduce corresponding operators $T_0$ and $T$ defined over <u>all</u> of Hilbert space. To do this we merely define $T_{01}$, $T_{0b}$, $T_{0b1}$, $T_{11}$, $T_{1b}$, $T_{b1}$ to be zero. Then (7) and (8) can be replaced by (<u>Problem</u>: Derive (9) and (10))

$$T = T_0 + T_0\,V\,T \tag{52-9}$$

$$T = T_0 + T\,V\,T_0 \tag{52-10}$$

whose "solutions" then yield two equivalent forms for $T$ :

$$T = (1 - T_0 V)^{-1}\,T_0 \tag{52-11}$$

or

$$T = T_0 (1 - V T_0)^{-1} \tag{52-12}$$

Following Löwdin we now introduce the "generalized reaction operator" $t$ according to

$$T = T_0 + T_0 t T_0 \tag{52-13}$$

whence from (10) and (11) we can derive the two equivalent forms

$$t = V (1 - T_0 V)^{-1} \tag{52-14}$$

or

$$t = (1 - V T_0)^{-1} V \tag{52-15}$$

(Problem: Derive (14) and (15). Hint: Note that even if A is an operator still $(1-A)^{-1} = (1-A)^{-1}(1 - A + A) = 1 + (1-A)^{-1} A$ ) If further $V^{-1}$ exists then recalling that $(AB)^{-1} = B^{-1} A^{-1}$ we see that $t^{-1}$ takes a very simply form, namely

$$t^{-1} = V^{-1} (1 - V T_0) .$$

whence

$$t^{-1} = V^{-1} - T_0 \tag{52-16}$$

Some other useful relations are the following: From (14) we have

$$t = V + t T_0 V \tag{52-17}$$

while from (15) we have

$$t = V + V T_0 t \tag{52-18}$$

By multiplying (9) by V from the right, and (18) by $T_0$ from the left and comparing the results we then find that (Problem: Derive this result)

$$TV = T_0 t \tag{52-19}$$

while multiplying in the other order we find that (Problem: Derive this result)

$$VT = t T_0 \tag{52-20}$$

We now introduce the bracketing operator $F$, $F_H$ being the bracketing function $f$,

$$F = H + H T H \tag{52-21}$$

Using the preceding relations we can now write it in terms of $H_0$ and $t$ as follows (Problem: Derive (22))

$$F = (H_0 + H_0 T_0 H_0) + (1 + H_0 T_0) t (1 + H_0 T_0) \tag{52-22}$$

Whence the bracketing function becomes

$$f = (\Phi_1, [(H_0 + H_0 T_0 H_0) + (1 + H_0 T_0)\, t\, (1 + H_0 T_0)]\Phi_1) \tag{52-23}$$

$$\equiv (\Phi_1, H_0 \bar{\Phi}_1) + (\bar{\Phi}_1, t\, \bar{\Phi}_1) \tag{52-24}$$

$$= (\bar{\Phi}_1, H_0 \Phi_1) + (\bar{\Phi}_1, t \bar{\Phi}_1) \tag{52-25}$$

where

$$\bar{\Phi}_1 = (1 + T_0 H_0)\Phi_1 \tag{52-26}$$

## 53. Bracketing Functions for Intermediate Hamiltonians

Having developed all this formalism we must now point out that as practical matter if $H$ is a realistic atomic or molecular Hamiltonian then in fact we can't exhibit $P$ nor can we calculate $f$. However we will now show that if $H$ is one of the intermediate Hamiltonians discussed earlier then we can calculate $f$'s. Now as a way of dealing with intermediate Hamiltonians the bracketing function is probably not of enormous interest - except for the case $H^I = H^{GL}$ (see below) there may be circumstances in which one would be satisfied with lower bounds to the $E_i^I$ (which then are still lower bounds to the $E_i$), also the graphical approach may on some occasions be a useful one for finding the $E_i^I$ [See the paper by T. M. Wilson (J. Chem. Phys. 47, 4706 (1967)) for some related comments]. More interesting is the fact that, as we will discuss in the next section, almost all the techniques which have been used to approximate $f$ for a realistic

can be seen as the use of an exact $f$ for some intermediate Hamiltonian $H^I$ (a point which is not made explicitly in the literature but which, as I have learned via private communications from T. M. Wilson and from O. Goscinskα, has been realized by some of the practitioners).

Thus we will now apply the preceding formulae to $H^I$ instead of to $H$. For $\mathcal{H}_0$ we will now write $\mathcal{H}_0^I$ and for $V$ we will write $V^I$. Thus we can take over all the preceding formulae by simply introducting the superscript $I$ at appropriate places. Further we will suppose that the reference function $\Phi_1$ is an eigenfunction of $\mathcal{H}_0^I$, for example $\psi_1^{0I}$:

$$\mathcal{H}_0^I \psi_1^{0I} = E_1^{0I} \psi_1^{0I} \tag{53-1}$$

This has the effect of simplifying things considerably since then

$$T_0^I \mathcal{H}_0^I \Phi_1 = E_1^{0I} T_0^I \Phi_1 \equiv 0 \tag{53-2}$$

and therefore

$$\overline{\Phi}_1 = \Phi_1$$

whence from (52-24) we have

$$f^I = E_1^{0I} + (\psi_1^{0I}, t^I \psi_1^{0I}) \tag{53-3}$$

Note that if one is using $f^I$ to determine exact eigenvalues of $H^I$ then the choice of $\Phi_1$ is irrelevant (except that one gets no information about eigenvalues associated with eigenfunctions orthogonal to $\Phi_1$). On the other hand if one is going to

use $f^I$ to yield bounds, then the quality of the bounds can depend crucially on the choice of $\phi_1$.

We now consider several of the special intermediate Hamiltonians which we discussed earlier.

(i) $H_0^I = H_0^L$, $V^I = V$, $\psi_1^{0I} = \psi_1^0$, $E_1^{0I} = E_1^0$ with $V$ simply a function. Thus this is the intermediate Hamiltonian $H^{L,\infty}$, considered by Miller in his 1965 paper and discussed in Sec. 48.

To find $t^{L,\infty}$ we must solve

$$t^{L,\infty} = V + V T_0^L t^{L,\infty} \tag{53-4}$$

where

$$T_0^L = \sum_{i=2}^{L} \frac{|\psi_i^0\rangle\langle\psi_i^0|}{\mathcal{E} - E_i^0} + \frac{1}{\mathcal{E} - E_{L+1}^0}\left[1 - \sum_{i=1}^{L} |\psi_i^0\rangle\langle\psi_i^0|\right] \tag{53-5}$$

Problem: Derive (5). Hint: Note that from (52-2) it follows that if $\phi_1, \phi_2, \ldots$ are actually $\psi_1, \psi_2, \ldots$ the eigenfunctions of $H$, then

$$T = \sum_{i=2}^{\infty} \frac{|\psi_i\rangle\langle\psi_i|}{\mathcal{E} - E_i} \tag{53-6}$$

By a little rearranging, (4) with (5) can be written in the form

$$(\mathcal{E} - E_{L+1}^0 - V)\, t^{L,\infty} = A + \sum_{i=1}^{L} B_i |\psi_i^0\rangle\langle\psi_i^0|\, t^{L,\infty} \tag{53-7}$$

where A and the $B_i$ are known functions. If now $(\mathcal{E} - E_{L+1}^0 - V)^{-1}$ exists, which if $V$ is positive will be the case if

$$\mathcal{E} < E_{L+1}^0 \tag{53-8}$$

then (7) can replaced by

$$t^{L,\infty} = a + \sum_{i=1}^{L} b_i |\psi_i^0\rangle\langle\psi_i^0| t^{L,\infty}$$

where now $a$ and the $b_i$ are known functions. If now we take the scalar product with $\psi_j^0$ we will have a set of inhomogeneous equations for the $(\psi_i^0, t^{L,\infty}\psi_j^0)$ from which, in particular we can determine $(\psi_1^0, t^{L,\infty}\psi_1^0)$, the quantity of interest. (Problem: Fill in the details.)

Note that if $V$ is positive then from (52-12) we can write

$$f^{L,\infty} = E_1^0 + (\psi_1^0, (V^{-1} - T_0^L)^{-1}\psi_1^0)$$

(ii) $V^I = V^N$ but we leave $\mathcal{H}_0^I$ unspecified for the moment except to assume that $\psi^{0I} = \psi_1^0$ with eigenvalue $E_1^0$, i.e. we assume that $\mathcal{H}_0^I \psi_1^0 = E_1^0 \psi_1^0$.

To find $t^I$ we must solve

$$t^I = V^N + V^N T_0^I t^I \tag{53-9}$$

If one now makes the ansatz

$$t^I = \sum_{k=1}^{N}\sum_{l=1}^{N} \sqrt{V}|s_k\rangle t_{kl}^I \langle s_l|\sqrt{V} \tag{53-10}$$

where the $t^I_{KL}$ are numbers, then one finds that if $V^{-1}$ exists (Problem: Fill in the details) this will be the solution of (9) provided that

$$t^I_{KL} = \delta_{KL} + \sum_{K'=1}^{N} (\zeta_K, \sqrt{V}\, T_0^I \sqrt{V}\, \zeta_{K'})\, t^I_{K'L} \qquad (53\text{-}11)$$

which is a set of inhomogeneous equations which we can solve for the $t^I_{KL}$ if we can calculate the coefficients. Three obvious cases in which we can do this are (a) $\mathcal{H}_0^I = \mathcal{H}_0^L$ (b) $\mathcal{H}_0^I = \mathcal{H}_0$ so that

$$T_0^I = T_0 = \sum_{i=2}^{\infty} \frac{|\psi_i^0 \rangle\langle \psi_i^0|}{\mathcal{E} - E_i^0} \qquad (53\text{-}12)$$

but where the conditions of Bazley's special choice are satisfied, i.e. where $\sqrt{V}\,\zeta_K$ is a finite linear combination of the (really) discrete $\psi_i^0$. (c) $\mathcal{H}_0^I = \mathcal{H}_0$ and the Gay-Löwdin choice $\zeta_K = \frac{1}{\sqrt{V}}(\mathcal{H}_0 - \mathcal{E})\zeta_K / \sqrt{n_K}$ (Problem: Carry through the details in each case. Hint: For case (c) note that $(\mathcal{E} - \mathcal{H}_0)\, T_0 = 1 - |\psi_i^0\rangle\langle\psi_i^0|$.

To facilitate comparison with the work of Lowdin and his students we note that since $(\zeta_K, \zeta_L) = \delta_{KL}$ it follows from (11) that $t^I_{KL}$ is the $KL$'th element of the $N \times N$ matrix which is the inverse of the $N \times N$ matrix with elements $(\zeta_K, \mathbb{1} - \sqrt{V}\, T_0^2 \sqrt{V}\, \zeta_{K'})$. With this observation our results can now be written very compactly as follows: Introduce the functions $h_K$ according to

$$h_K = \sqrt{V}\, \zeta_K$$

Then we have

$$t^I = \sum_{K=1}^{N} \sum_{L=1}^{N} |h_K\rangle\, t^I_{KL}\, \langle h_L|$$

where $t^I_{KL}$ is the $KL$ 'th element of the $N\times N$ matrix which is the inverse of the $N\times N$ matrix with elements

$$(h_K, [V^{-1} - T_0^{-1}] h_{K'})$$

Problem: We derived these results assuming that the $\xi_K$ are orthonormal. Show that in the present form the $h_K$ can be an arbitrary basis by showing that the form of $t^I$ is invariant to an arbitrary non-singular linear transformation of the $h_K$ .

Problem: Show that $t^{GL}$ can be written

$$t^{GL} = \sum_{K=1}^{N}\sum_{L=1}^{N} (\mathcal{E}-H_0)|\xi_K\rangle\, t^{GL}_{KL}\, \langle\xi_L|\,(\mathcal{E}-H_0)$$

where $t^{GL}_{KL}$ is the $KL$ 'th element of the $N\times N$ matrix which is the inverse of the $N\times N$ matrix with elements

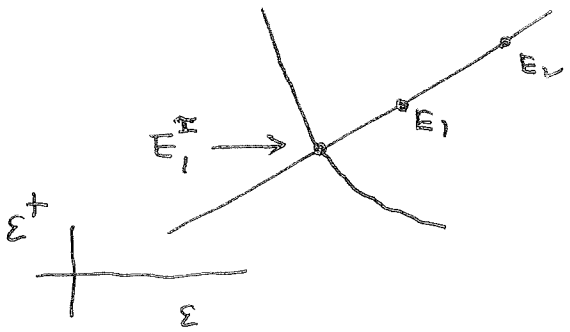
$$(\xi_K, [(\mathcal{E}-H_0)V^{-1}(\mathcal{E}-H_0) - (\mathcal{E}-H_0)(1-|\psi_0^0\rangle\langle\psi_0^0|)]\xi_{K'})$$

Show that in this form the $\xi_K$ can be an arbitrary basis.

As we described in Sec. 49, the Gay-Löwdin method, involving as it does an energy dependent truncation, has certain peculiar features from the point of view of intermediate Hamiltonians. Not surprisingly then its bracketing functions also has peculiar features. Namely since $H^{GL}$ depends on $\mathcal{E}$ , one cannot prove that $f^{GL}$ is a decreasing function of $\mathcal{E}$ , and in fact it is not. This has the following consequence. If we introduce again the operators $H^{GL}(b)$ of Sec. 49 then we know that $\mathcal{E}$ and $f^{GL}(b)$ will bracket an eigenvalue of $H^{GL}(b)$ and therefore the smaller of them, being a lower bound

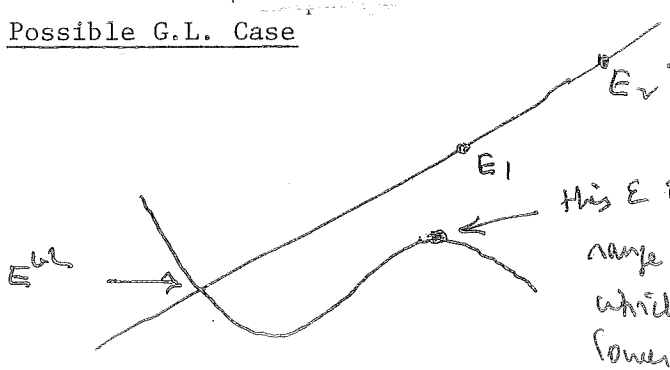
to some eigenvalue of $H^{(2)}(\delta)$ , is also a lower bound to the corresponding $E_i$ . Now since this is true for all $\delta$ it is in particular true for $\delta = \varepsilon$ . So thus far there is nothing peculiar. The peculiar feature is that $f^{(2)} = \varepsilon$ , which determines the eigenvalues of $H^{(2)}$ , does not in general yield the best lower bound. What we mean is best seen graphically: Suppose we are dealing with the ground state and that we know that $\varepsilon < E_2$ and we find that the bracketing function from our intermediate Hamiltonian is less than $\varepsilon$ then:

Usual Case




$\varepsilon$ must lie in the range $f(\varepsilon) < \varepsilon < E_2$ and clearly we get the best lower bound to $E_1$ when $\varepsilon = \varepsilon_1^I$ , that is $E_1^I$ is the best lower bound

Possible G.L. Case



Thus by use of the bracketing functions one can improve on simply using $H^{GL}$ (for some numbers see the references in the next section). How this approach compares with the use of $\mathcal{H}^M$ is not clear at the moment, hence our remark at the end of comment (i) following eq.(49-5).

(iii) $H^I = \mathcal{H}^M$ , $V^I = V$ Thus we have for our $\mathcal{H}_0^I$

$$\mathcal{H}_0^M = \sum_{i=1}^{L} \frac{(\mathcal{H}_0 - \epsilon)|\hat{\phi}_i\rangle\langle\hat{\phi}_i|(\mathcal{H}_0 - \epsilon)}{\hat{\varepsilon}_i - \epsilon} + \epsilon \qquad (53\text{-}13)$$

We now need to find the eigenfunctions of $\mathcal{H}_0^M$ . First we note that any function orthogonal to all the $(\mathcal{H}_0 - \epsilon)\hat{\phi}_i$ is an eigenfunction, with eigenvalue $\epsilon$ . To find the others we must then solve (47-8) but with $V = 0$ . If we denote the resultant eigenfunctions by $\psi_i^M$ and the eigenvalues by $\varepsilon_i^M$ then evidently if we choose $\psi_1^I$ to be $\psi_1^M$ we have (Problem: Derive this)

$$T_0^M = \sum_{i=2}^{L} \frac{|\psi_i^M\rangle\langle\psi_i^M|}{\varepsilon - \varepsilon_i^M} + \frac{1}{\varepsilon - \epsilon}\left[1 - \sum_{i=1}^{L} |\psi_i^M\rangle\langle\psi_i^M|\right] \qquad (53\text{-}14)$$

from which one can readily find $t^M$ and $f^M$ (Problem: Fill in the details.)

In a certain sense there is no ordering problem when dealing with the bracketing function for a soluble $H^I$ - one can simply solve $H^I$ . However this is really going outside the bracketing function approach. We have already indicated for the general case (soluble or not) an internal procedure for dealing with the ground state. This has been generalized to excited states in the following way (we will merely sketch the results and refer to T. M. Wilson, J. Chem. Phys. 47, 3912 (1967)! Instead of partioning into $1 \oplus (\infty - 1)$ one partitions into $g \oplus (\infty - g)$ Then instead of (51-7) one finds

$$|E I_{aa} - H_{aa} + H_{ab}(H_{bb} - E I_{bb})^{-1} H_{ba}| = 0 \qquad (53\text{-}15)$$

and correspondingly we define an analogue to the functional

and correspondingly we define an analogue to the functional
relation $\varepsilon^+ - f(\varepsilon) = 0$ by

$$| \varepsilon^+ |_{aa} - H_{ac} + H_{ab}(H_{bb} - \varepsilon)^{-1} H_{bc} | = 0 \qquad (53\text{-}16)$$

This is called "the Multivalued Bracketing Function". For each $\varepsilon$ there are $g$ roots, call them $\varepsilon_i^+$, of (16). One can then show the following. If $H = H_0 + V$ where $V$ is positive and $H$ one uses as reference functions $\psi_1^0, \psi_2^0 \cdots \psi_g^0$ then

$$E_i < \varepsilon < E_{g+1}^0$$

then one can show that $\varepsilon_i^+$ will be a lower bound to $E_i$

$$\varepsilon_i^+ < E_i \qquad (53\text{-}17)$$

(See also Miller J. Chem. Phys. 48, 530 (1968)).

---

Problem: Show that one can exhibit the multivalued bracketing function ~~obtained as above~~ in explicit form for one or all of the $H^I$ descussed earlier. Not that one can still introduce a bracketing operation $\cdot H \cdot = H T H$ and all the formalism goes through except that now $T_0 = \sum_{g+1}^{\infty} \frac{\psi_i^0 \times \psi_i^0}{\varepsilon - \varepsilon_i^0}$

## 54. Applications of the Bracketing Fucntion

As we have mentioned, in order to implement the bracketing approach for atoms and molecules one is forced to approximations. In the last section we discussed the approximation by means of intermediate Hamiltonians. Löwdin and co-workers following a rather different approach were in fact led to essentially the same approximations to the bracketing functions are in fact exact bracketing functions for various $H^I$. (As remarked earlier this is not noted explicitly in the literature though several authors (see for example T. M. Wilson, J. Chem. Phys. 47, 3912, 4706 (1964), Wilson and Reid, J. Chem. Phys. 47, 3920 (1967) and references there) did point out that in various cases the intersection points yielded the eigenvalues of certain $H^I$ ). Thus

rather than detail their approach we will simply give a list of references accompanied by brief comments. (Note that when we say "using" we will mean "in effect using".) At the end we will mention a different approach due to Pearl Wang.

1. Löwdin, Phys. Rev. 139 A 357 (1965): The example is He. Lower bounds were found using $H^{GL}$.

2. Reid, J. Chem. Phys. 43, 5186 (1965): The example is an anharmonic oscillator. Lower bounds to lower bounds were found by using a Bazley special choice $H^{I}$. The results of course are not as good as solving $H^{I}$ directly but there may be some numerical advantages (see footnote 4a of the paper).

3. Choi and Smith, J. Chem. Phys. 43, S 189 (1965): The example is a rigid rotator. Comments similar to those given in 2 apply.

4. Bunge and Bunge, J. Chem. Phys. 43, S 194 (1965): The example is a double minimum potential. Comments similar to those given in 2 apply. All the preceding calculations employed a single reference function.

5. Wilson and Reid, J. Chem. Phys. 47, 3920 (1967): Here the multi-valued bracketing function was used and applied to Helium. Both $H^{L,N}$ and $H^{GL}$ were used. Complete regions of the $\varepsilon_t^i(\varepsilon)$ curves are plotted out and intersections found. In particular the determination of optimal lower bounds in the $H^{GL}$ case is considered in detail.

We now wish to briefly describe the method used by Pearl Wang (J. Chem. Phys. 48, 4131 (1968). With $\phi_1 = \psi_1^0$ we have, similarly to (53-3)

$$f = E_1^0 + (\psi_1^0, \varepsilon \psi_1^0)$$

and the Löwdin approach was to get lower bounds to $f$ by introducing various operator lower bounds to $t$. These approximations then, as we have now said several times, turn out to be various $f^I$.

What Wang did was to go back to

$$f = (\phi_1, [H + H\tau H]\phi_1)$$

and get a lower bound to $f$ by using lower bounds to $\tau$, and in fact she replaced $\tau$ by $\tau^I$. However her reference functions was not such that

$$(H^I - H)\phi_1 = 0$$

so that her $f$

$$f^W = (\phi_1, [H + H\tau^I H]\phi_1)$$

is different from

$$f^I = (\phi_1, [H^I + H^I \tau^I H^I]\phi_1)$$

(Problem: Prove this.) The problems she considered were $H_1$ and $H_2$, and for $H^I$ she used the Hamiltonian of Bazley and Fox mentioned at the end of Sec. 48. Her results were quite good and in particular much better than those furnished by the eigenvalues of $H^I$.