REMARKS ON THE PERTURBATION FORMULAE OF BRILLOUIN AND WIGNER

H. A. Kramers

Consider the eigenvalue problem of a Hermitian matrix $H_{k\ell}$

(1) $$\sum_{\ell} H_{k\ell} C_{\ell} = F\, C_k \qquad \sum_{k} C^{*}_{k} H_{k\ell} = C^{*}_{\ell} F$$

in which $H_{k\ell}$ differs only little from a diagonal matrix $E_k\, \delta_{k\ell}$:

(2) $$H_{k\ell} = E_k \delta_{k\ell} + V_{k\ell}$$

Brillouin[1] has established the following relation which -- in case of convergence of the infinite series -- is satisfied by eigenvalues

(3) $$F - E_1 = V_{11} + \sum_{k} \frac{V_{1k} V_{k\ell}}{F - E_k} + \sum_{k,\ell} \frac{V_{1k} V_{k\ell} V_{\ell 1}}{(F-E_k)(F-E_{\ell})} + \ldots \quad (k, \ell \neq 1)$$

In this series one of the indices -- in our formula the index 1 -- plays a part of its own. If $|V_{1k}| \ll |E_k - E_1|$ for all k's, the first two or three terms in (3) are often very useful to calculate the perturbed eigenvalue which are near to E_1, even if the V's and E's are such that the series does not converge. The difference between (3) and the closely analogous series of the customary perturbation theory of Rayleigh-Schrödinger lies in the appearance of the eigenvalue in the right member of (3). Thus Schrödinger's well known second approximation is given by

(4) $$F - E_1 = V_{11} + \sum_{k} \frac{V_{1k} V_{k1}}{E_1 - E_k}$$

Lennard-Jones[2] was the first to point out the importance in some

1) L. Brillouin, Journ. de Ph. 4, 1, 1933.
2) J. E. Lennard-Jones, Proc. Roy. Soc., London, A129, 598, 1930.

cases of expressions with "resonance denominators" which contain the eigenvalue itself.

Wigner[3] has proved that if the series (3) is cut off after an odd number $2\nu+1$ of terms, the relation (3) precisely yields the approximate eigenvalue(s) which follow if we assume that the coefficients C_k can be approximately represented by the following finite expression consisting of ν terms,

$$C_1 = 1, \quad C_k = \underset{(T_1)}{\frac{V_{k1}}{F-E_k}} + \underset{(T_2)}{\sum_{\ell} \frac{V_{k\ell}V_{\ell 1}}{(F-E_k)(F-E_\ell)}} + \dots + T_\nu \tag{5}$$

He showed, indeed, that the espression

$$E = \sum_{k,\ell} C_k^{*} H_{k\ell} C_\ell \Big/ \sum_k C_k^{*} C_k$$

in which the C's are represented by (5), becomes precisely equal to F if F satisfies (3) (with $2\nu+1$ terms). In their proofs Brillouin and Wigner assumed that the E_k's are identical with the diagonal terms of $H_{k\ell}$ (thus, for instance, $V_{11} = 0$), but it is easy to see that this assumption is not necessary.

We shall now give a simple proof of the formal validity of the (infinite) series (3), which fundamentally does not differ from Brillouin's own proof but which still might be of some interest. For this purpose we decompose $V_{k\ell}$ into two matrices V' and V" such that V' in its first row coincides with $V_{k\ell}$ but otherwise has nothing but zeros, whereas V" has zeros in its first row but otherwise coincides with $V_{k\ell}$:

$$V = V' + V'' \qquad \begin{array}{ll} V'_{1\ell} = V_{1\ell} & V''_{1\ell} = 0 \\ V'_{k\ell} = 0 \quad (k \neq 1) & V''_{k\ell} = V_{k\ell} \quad (k \neq 1) \end{array} \tag{6}$$

The second of the equations (1) can now symbolically be written as follows

[3] E. Wigner, Math. Anseiger Ungar. Akad. 53, 477, 1935.

$$(7) \qquad C^{*}H \equiv C^{*}(D + V' + V'') = C^{*}F,$$

where D is the diagonal matrix with elements $E_{k\ell}\ \delta_{k\ell}$. D, V' and V" are matrices operating on C^* whereas F is a real number.

From (7) follows

$$C^{*}V' = C^{*}(F - D - V'')$$

$$C^{*}V'\frac{1}{F-D} = C^{*}(1 - V''\frac{1}{F-D})$$

$$C^{*} = C^{*}V'\frac{1}{F-D}(1 - V''\frac{1}{F-D})^{-1}$$

$$= C^{*}(V'\frac{1}{F-D} + V'\frac{1}{F-D}V''\frac{1}{F-D} + V'\frac{1}{F-D}V''\frac{1}{F-D}V''\frac{1}{F-D} + \ldots)$$

Reintroducing indices we find from this equation in the first place

$$C_1^{*} = C_1^{*}(\frac{V'_{11}}{F-E_1} + \sum_k \frac{V'_{1k}V''_{kl}}{(F-E_k)(F-E_l)} + \ldots.) \quad (k \neq 1)$$

or, after dividing by C_1^* and multiplication with $F-E_1$

$$F - E_1 = V'_{11} + \sum_k \frac{V'_{1k}V''_{kl}}{F-E_k} + \ldots.$$

This is identical with (3) since the accents can be omitted.

In the second place one finds

$$C_k^{*} = C_1^{*}(\frac{V'_{1k}}{F-E_k} + \sum_k \frac{V'_{1\ell}V''_{\ell k}}{(F-E_\ell)(F-E_k)} + \ldots.)$$

but this is (for $C_1^* = 1$) precisely the same as equation (5) (with $\nu \to \infty$).

It does not seem as if the introduction of (2) and (6) would help towards simplifying the proof of Wigner's interesting theorem in a significant way.

In the approximation method of Lennard-Jones, all denominators in (4) are replaced by an appropriately chosen average denominator Δ. In this connection we will inquire what will be the result of (3) if all E_k's are chosen equal to $E_1 + \Delta$ where E_1 is defined as the matrix element

$$E_1 = H_{11} .$$

Introducing the following abbreviations:

$$\varepsilon = F - E_1 , \quad \left\{(H - E_1)^n\right\}_{11} = a_n \tag{8}$$

we find, after some calculation:

$$\varepsilon = -\frac{a_2}{\Delta - \varepsilon} + \frac{a_3 - \Delta a_2}{(\Delta - \varepsilon)^2} - \frac{(a_4 - a_2^2) - 2\Delta a_3 + \Delta^2 a_2}{(\Delta - \varepsilon)^3}$$
$$+ \frac{(a_5 - 2a_3 a_2) - 3\Delta(a_4 - a_2^2) + 3\Delta^2 a_3 - \Delta^3 a_2}{(\Delta - \varepsilon)^4} \tag{9}$$

If one cuts off after an even number 2ν of terms, the ε value which corresponds to $d\varepsilon/d\Delta = 0$ would--in view of Wigner's theorem --yield the best approximation within the frame of this method.

If H can be written in the form

$$H_{k\ell} = H_{hk}\,\delta_{k\ell} + \lambda u_{k\ell}$$

where λ is a small parameter, and if one asks only for the contribution to ε which is proportional to λ^2, one may first of all omit all ε's in the denominators of (8) and one gets

$$\varepsilon = -\frac{2a_2}{\Delta} + \frac{a_3}{\Delta^2} \qquad (\nu = 1) \tag{10}$$

$$\varepsilon = -\frac{4a_2}{\Delta} + \frac{6a_3}{\Delta^2} - \frac{4(a_4 - a_2^2)}{\Delta^3} + \frac{a_5 - 2a_3 a_2}{\Delta^4} \qquad (\nu = 2) \tag{11}$$

If we denote the 11-element of a matrix M by a bar:

$$M_{11} = \overline{M}$$

one finds for the terms proportional to λ^2 in the a_n's

$$a_n = \lambda^2\, \overline{u(H - E_1)^{n-2} u} \tag{12}$$

Putting moreover

$$M^{(0)} = \overline{M}, \quad M^{(1)} = \frac{i}{\hbar}\overline{(HM - MH)}, \quad M^{(\nu+1)} = \frac{i}{\hbar}\overline{(HM^{(\nu)} - M^{(\nu)}H)}$$

which, if one prefers so, can be written also as

$$M^{(\nu)} = \overline{\left(\frac{d}{dt}\right)^{\nu} M}$$

we find from (12)

$$a_{2\nu+2} = \lambda^2 \hbar^{2\nu} \overline{u^{(\nu)2}}$$

$$a_{2\nu+3} = \tfrac{1}{2}\lambda^2 \hbar^{2\nu+1} \overline{i(u^{(\nu+1)}u^{(\nu)} - u^{(\nu)}u^{(\nu+1)})}$$

or

$$a_n = \lambda^2 \left(\frac{\hbar}{i}\right) \overline{u\, u^{(n-2)}}$$

For (10) and (11) we can therefore write:

$$\frac{\varepsilon}{\lambda^2} = -\frac{\overline{2u^2}}{\Delta} + \frac{1}{2}\hbar i \frac{\overline{u^{(1)}u - u\, u^{(1)}}}{\Delta^2} \tag{14}$$

$$\frac{\varepsilon}{\lambda^2} = -\frac{\overline{4u^2}}{\Delta} + 3\hbar i \frac{\overline{u^{(1)}u - u\, u^{(1)}}}{\Delta^2} - 4\hbar^2 \frac{\overline{u^{(1)2}}}{\Delta^3} + \frac{1}{2}\hbar^3 i \frac{\overline{u^{(2)}u^{(1)} - u^{(1)}u^{(2)}}}{\Delta^4} \tag{15}$$

by variation of Δ one finds the extreme values for ε. Thus from (14) we have simply

$$\frac{\varepsilon}{\lambda^2} = -\frac{(\overline{u^2})^2}{\frac{1}{2}\hbar i\, \overline{(u^{(1)}u - u\, u^{(1)})}} \tag{16}$$

This is the formula of Hassé's[4] approximation method.

If one asks for the polarizability of a one-electron system we may take for u one of the cartesian coordinates, say x, and (16) gives the well-known approximative expression

---

4) H. R. Hasse, Proc. Cambr. Phil. Soc. 26, 542, 1930. H. Margenau, Rev. Mod. Phys. 11, 1, 1939.

$$\varepsilon = -\lambda^2 \frac{\overline{(x^2)}^2}{\hbar^2/2m} .$$

If, however, one tries to apply (15) to the same problem, one meets the difficulty that in the hydrogen atom the fourth term gets undetermined and that--if expressions containing 6 or more terms are used ($\nu = 3, 4, \ldots$)--all terms after the fourth diverge. This probably means that in general (8) will be of little use.

In many quantum-mechanical problems (compare the H_2-molecule) approximative wave functions can be established although one cannot distinguish between unperturbed Hamiltonian and perturbation energy (like the λu above). In such cases the minimum value which ε can take in the expression (see (9)) might sometimes represent a good second approximation. Putting

$$\varepsilon = -\frac{a_2}{\Delta - \varepsilon} + \frac{a_3 - \Delta a_2}{(\Delta - \varepsilon)^2}$$

one finds easily

$$\varepsilon = \frac{1}{2}\frac{a_3}{a_2}\left(1 - \sqrt{1 + 4a_2{}^3/a_3{}^2}\right) .$$

which expresses ε in terms of the second and third "moment" of H (comp. (8)).