

UNCLASSIFIED

AD

426 656

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA



UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

64-6

D1-82-0316

426656

426656

AG

BOEING SCIENTIFIC
RESEARCH
LABORATORIES

426656

A Note on Regular Perturbation Theories

F. E. Bisshopp

November 1963

DL-82-0316

BOEING SCIENTIFIC RESEARCH LABORATORIES
FLIGHT SCIENCES LABORATORY REPORT NO. 84

A NOTE ON REGULAR PERTURBATION THEORIES

by

F. E. Bisshopp

November 1963

A NOTE ON REGULAR PERTURBATION THEORIES

by

F. E. Bisshopp

I. INTRODUCTION

There exist in the literature several related formulations of regular perturbation theory. (c.f. Morse & Feshbach; Kumar). In deciding which formalism to apply to a specific problem we must take into account the radius of convergence of the various methods, and, if possible, the ease of application. In this note we shall develop a formalism, closely related to Fredholm's resolvent, which has an infinite radius of convergence. The present formulation shares with the formalism derived directly from the Fredholm resolvent (Morse & Feshbach) the disadvantage of being rather cumbersome in its application. Consequently, we shall go on to develop an asymptotic approximation to the result. By this means we shall obtain a relatively simple procedure for evaluating the result of a formulation which is unrestricted as regards radius of convergence. Finally, we shall discuss the close relation between the present theory and Feenberg's (1948) formulation of regular perturbation theory.

II. FORMULATION OF THE CLASS OF PROBLEMS

We shall confine our attention to eigenvalue problems which may be written in the form

$$H_{\epsilon} \psi_n = E_n \psi_n \quad (1)$$

where:

1. H_{ϵ} is a self adjoint, linear operator which depends in a continuous manner on a complex-valued parameter ϵ ;
2. $\{E_n\}$ is a complex-valued discrete spectrum;
3. ψ_n is an element of a Hilbert space, Φ ;
4. With the possible exception of a finite number of isolated singularities in the ϵ -plane, H_{ϵ} is a bounded operator.

It should be noted that the definition of the space Φ generally involves specification of certain boundary conditions on ψ_n ; we shall consider the case where the boundary conditions may depend upon ϵ , but not upon E_n .

Now, the object of a perturbation theory is to develop a technique for solving for ψ_n and E_n in terms of known solutions, u_n and $\mathcal{E}_n$, of a related eigenvalue problem,

$$H_0 u_n = \mathcal{E}_n u_n \quad (2)$$

The operator H_0 will be assumed self-adjoint, and the above problem, called the unperturbed problem, is to be related to the original in the following sense:

1. It must be possible to define the difference,

$$H_1 \equiv H_0 - H_\epsilon, \text{ on } \Phi;$$

2. The space Φ_0 spanned by the solutions $\{u_n\}$ must be a Hilbert space for which

$$\psi \in \Phi \rightarrow \begin{cases} H_0 \psi \in \Phi_0 \\ H_1 \psi \in \Phi_0 \\ E \psi \in \Phi_0 \end{cases}$$

where E is any complex number;

3. There must be a value of ϵ (which we may take to be $\epsilon = 0$) for which $H_1 = 0$.

It will be convenient to take $\{u_n\}$ to be a normal set, i.e.

$$(u_n, u_m) = \delta_{nm} \quad (3)$$

where (u,v) is a suitably defined inner product on Φ_0 . It should be noted that we have made no specific assumption of independence of H_0 on ϵ , though in practice this is often the case.

In virtue of the identity,

$$H_\epsilon \equiv H_0 - (H_0 - H_\epsilon) \equiv H_0 - H_1, \quad (4)$$

we may write

$$\psi_n = (H_0 - E_n)^{-1} H_1 \psi_n. \quad (5)$$

If we define a projection operator, P_n , according to the property,

$$(u_l, P_n u_m) = \delta_{ln} \delta_{mn}, \quad (6)$$

we have

$$\begin{aligned}
 \psi_n &= P_n \psi_n + (1 - P_n) \psi_n \\
 &= g(\epsilon) u_n + (1 - P_n) (H_0 - E_n)^{-1} H_1 \psi_n \\
 &= g(\epsilon) (1 - K_n)^{-1} u_n
 \end{aligned} \tag{7}$$

where

$$\begin{aligned}
 K_n &\equiv (1 - P_n) (H_0 - E_n)^{-1} H_1 \\
 \text{and} \quad g(\epsilon) &\equiv (u_n, P_n \psi_n) = (u_n, \psi_n)
 \end{aligned} \tag{8}$$

Note that $g(\epsilon)$ merely fixes the normalization of ψ_n ; therefore it is to a certain extent a function which may be chosen at our convenience. It will be convenient to take

$$g(0) = 1, \tag{9}$$

for then, if the eigenvalue, ϵ_n , of the unperturbed problem is non-degenerate, we have

$$K_n u_n \rightarrow 0 \quad \text{and} \quad \psi_n \rightarrow u_n \quad \text{as} \quad \epsilon \rightarrow 0. \tag{10}$$

III. BRILLOUIN-WIGNER FORMALISM

This perturbation theory, perhaps the easiest one to derive, follows upon setting

$$g(\epsilon) = 1 \quad \text{and} \quad (1 - K_n)^{-1} = \sum_{m=0}^{\infty} (K_n)^m \quad (11)$$

The eigenvalue E_n then follows by computation of the inner product of Equation (1) and u_n ; thus

$$E_n = \left(u_n, (H_0 - H_1) \sum_{m=0}^{\infty} (K_n)^m u_n \right) \quad (12)$$

By definition

$$\begin{aligned} (K_n)_{\ell m} &= \left(u_{\ell}, (1 - P_n)(H_0 - E_n)^{-1} H_1 u_m \right) \\ &= (1 - \delta_{n\ell}) \frac{H_{\ell m}}{\epsilon_{\ell} - E_n}, \end{aligned} \quad (13)$$

where $H_{\ell m} \equiv (u_{\ell}, H_1 u_m)$.

Thus

$$\begin{aligned} (H_0 K_n)_{nn} &= 0 \quad \text{for } m \neq 0 \\ (H_1 K_n)_{nn} &= \left(\sum_{\ell_1 \neq n} \right)^m \frac{H_{n\ell_1} H_{\ell_1\ell_2} \dots H_{\ell_m n}}{(\epsilon_{\ell_1} - E_n) \dots (\epsilon_{\ell_m} - E_n)} \end{aligned} \quad (14)$$

where $\left(\sum_{\ell_1 \neq n} \right)^m = \sum'_{\ell_1 \neq n} \sum'_{\ell_2 \neq n} \dots \sum'_{\ell_m \neq n}$

and

$$E_n = \varepsilon_n - H_{nn} - \sum_{m=1}^{\infty} \left(\sum_{\ell_1 \neq n} \right)^m \frac{H_{n\ell_1} H_{\ell_1 \ell_2} \dots H_{\ell_m n}}{(\varepsilon_{\ell_1} - E_n) \dots (\varepsilon_{\ell_m} - E_n)} \quad (15)$$

The above equation gives E_n implicitly; in general the right-hand side must be truncated and E_n solved for numerically.

The usefulness of this method is often impaired by the fact that the domain of the ε -plane within which it converges may be too small to include the problem under consideration. In virtue of the identity,

$$(1 - K_n) \left(\sum_{m=0}^{N'} K_n^m \right) = 1 - K_n^{N+1}, \quad (16)$$

it follows that a sufficient condition for convergence is

$$\| K_n \| \equiv \text{Max} \left((K_n u, K_n u) \right)^{\frac{1}{2}} < 1 \quad (17)$$

where the maximum is that obtained as u varies over the set $\{u_m\}$. Thus, if there is a simply connected domain of the ε -plane which contains $\varepsilon = 0$ and within which $\| K_n \| < 1$, the Brillouin-Wigner result will converge within that domain regardless of the details of the dependence of H_ε on ε . In practice it is difficult, if not impossible, to determine the domain of convergence (even when H_ε is uncomplicated) because of the appearance of $E_n(\varepsilon)$ in K_n .

It should be noted that if ε_n is a degenerate eigenvalue having associated solutions $u_n, v_n, \dots, w_n$ and if the limit as $\varepsilon \rightarrow 0$ of one or more of the quantities

$$\frac{(u_n, H_1 v_n)}{e_n - E_n}, \dots, \frac{(u_n, H_1 w_n)}{e_n - E_n}$$

is nonzero, then the present formalism is inapplicable. Since the generalization which includes such cases is well known, we shall not treat such degenerate perturbations in this or any of the following sections.

IV. THE FREDHOLM RESOLVENT

In this section we shall discuss a formalism which is similar to, but not identical with, that based on Fredholm's resolvent (cf. Morse & Feshbach, p. 1018 ft.). It will simplify matters considerably if in what is to follow we employ a device by means of which the case where H_1 depends upon ϵ in an arbitrary manner is reduced to the case where H_1 is a linear function of a new parameter, λ . First, let us reconsider what we want out of our perturbation theory: the point is, we want a continuation of the solution of Equation (1) which is given by the Brillouin-Wigner theory near $\epsilon = 0$ to a region of the ϵ -plane which includes the 'final' value, ϵ_0 (say). In general such a continuation will depend upon the path in the ϵ -plane along which the solution is followed; i.e., there will in general be branch points and singularities of H_1 . However, once a path $\bar{O}\epsilon_0$ is decided upon along which H_1 is bounded and varies continuously, we may consider at each ϵ lying on $\bar{O}\epsilon_0$ the related problem where

$$H_1 \rightarrow \lambda H_1 \quad . \quad (18)$$

The process of continuation of the solution along $\bar{O}\epsilon_0$ may now be looked upon as a continuation of the solution along a path $\bar{O}1$ in the λ -plane at each value of ϵ on $\bar{O}\epsilon_0$. Since the perturbation, λH_1 , varies linearly in the λ -plane, the solution, ψ_n , is an entire function of λ , and the continuation is independent of the particular path $\bar{O}1$. Furthermore, if ϵ be taken sufficiently close to zero, the application of the Brillouin-Wigner theory to the related problem has a radius of convergence greater than 1. Thus, we obtain a unique continuation along $\bar{O}\epsilon_0$,

satisfying the requirement of reduction to the Brillouin-Wigner result for sufficiently small ϵ .

Let us now turn to the continuation of solutions in the λ -plane. Equation (7) becomes

$$\psi_n = (1 - \lambda K_n)^{-1} u_n \quad (19)$$

where we have again set the function g which fixes the normalization of ψ_n equal to one.

The treatment of Equation (19) given by Morse & Feshbach parallels the treatment of integral equations of the Fredholm type (c.f. Mikhlin, Section 9). The essential notion involved in this treatment is that of introducing a function $\chi(\lambda)$ whose reciprocal represents the singularities, which according to Fredholm's alternative appear at values of λ for which the homogeneous equation,

$$(1 - \lambda_m K_n) \varphi_m = 0, \quad (20)$$

has a nontrivial solution. The determination of $\chi(\lambda)$ is carried out by writing Equation (19) as

$$\psi_n = \frac{1}{\chi(\lambda)} \left[\chi(\lambda) (1 - \lambda K_n)^{-1} \right] u_n, \quad (21)$$

where it is required that $\chi(\lambda)(1 - \lambda K_n)^{-1}$ shall be an entire function of λ and $\|\chi(\lambda)(1 - \lambda K_n)^{-1}\| \neq 0$. It follows that $\chi(\lambda)$ is an entire function of λ which vanishes only at the points λ_m for which Equation (20) holds; the order of the zero at λ_m is given by the degree of degeneracy of λ_m .

We shall not pursue this line of development any further since detailed accounts are available in the references cited in this section. The point we wish to make is that a somewhat simpler formalism results if we observe that the function $X(\lambda)$ need not be introduced in the present problem since we have at our disposal the normalizing function $g(\epsilon)$. We shall take the normalizing function as $g(\epsilon, \lambda)$ with $g(\epsilon) = g(\epsilon, 1)$, and, suppressing for the moment the dependence of g on ϵ , write

$$\psi_n = g(\lambda)(1 - \lambda K_n)^{-1} u_n \quad (22)$$

with the requirement that $g(\lambda)(1 - \lambda K_n)^{-1}$ shall be an entire function of λ and that $\|g(\lambda)(1 - \lambda K_n)^{-1}\| \neq 0$.

Note that the present formulation implies that if ψ_n is to be nonsingular, then its projection onto u_n vanishes at the eigenvalues, λ_m . Thus it may be seen that the singularities encountered in previous treatments follow from the unrealistic assumption that $\psi_n(\lambda)$ is never orthogonal to $\psi_n(0)$.

The function $g(\lambda)$ is not by any means uniquely determined by Equation (22) and the associated restriction; a given solution may be multiplied by any nonvanishing, entire function of λ . The formal expression for ψ_n of course depends explicitly on the choice of $g(\lambda)$; different functions, $g(\lambda)$, lead to different weighting of the various orders in the ordering of terms in the solution according to powers of λ . Consequently, the number of regular perturbation theories we may derive in this manner is limitless.

Now suppose we specify the normalization integral,

$$(\psi_n, \psi_n) = f(\lambda) \quad (23)$$

According to the properties of $g(\lambda)$, $f(\lambda)$ may be taken as any non-vanishing, entire function. The most natural choice for $f(\lambda)$ is of course $f(\lambda) = 1$; however, we shall not adopt any specific choice of normalization. Thus we have

$$[g(\lambda)]^2 \left((1 - \lambda K_n)^{-1} u_n, (1 - \lambda K_n)^{-1} u_n \right) = f(\lambda) \quad (24)$$

and

$$\begin{aligned} E_n &= \left(\psi_n, (H_0 - \lambda H_1) \psi_n \right) / (\psi_n, \psi_n) \\ &= \frac{\left((1 - \lambda K_n)^{-1} u_n, (H_0 - \lambda H_1) (1 - \lambda K_n)^{-1} u_n \right)}{\left((1 - \lambda K_n)^{-1} u_n, (1 - \lambda K_n)^{-1} u_n \right)}. \end{aligned} \quad (25)$$

It should be noted that when E_n is expressed as in Equation (25), the quantities $f(\lambda)$ and $g(\lambda)$ do not appear, i.e., all of the different perturbation theories derivable by choice of $g(\lambda)$ (or $f(\lambda)$) give precisely the same expression for E_n . Now the above equation gives E_n as an entire function of λ ; therefore if we consider sufficiently small values of λ and use

$$(1 - \lambda K_n)^{-1} = \sum_{m=0}^{\infty} (\lambda K_n)^m \quad \text{for } |\lambda| < \lambda_0 \quad (26)$$

The result, if expressed as a power series in λ , holds for all values of λ . Thus we obtain

$$E_n = \sum_{m=0}^{\infty} \lambda^m J_m / \sum_{m=0}^{\infty} \lambda^m I_m, \quad (27)$$

where

$$\left. \begin{aligned} J_m &= \sum_{\ell=0}^m (u_n, K^{+\ell} H_0 K^{m-\ell} u_n) \\ &\quad - \sum_{\ell=0}^{m-1} (u_n, K^{+\ell} H_1 K^{m-1-\ell} u_n) \\ I_m &= \sum_{\ell=0}^m (u_n, K^{+\ell} K^{m-\ell} u_n) \end{aligned} \right\} \quad (28)$$

To obtain the power series for E_n , we now treat $(\sum \lambda^m I_m)^{-1}$ as the generating function of a series $\sum \lambda^m G_m$. From Equation (13) it follows that

$$(K^{+m})_{nn} = (K^m)_{nn} = 0 \quad \text{for } m \neq 0 \quad (29)$$

and thus that

$$I_0 = 1, \quad I_1 = 0 \quad (30)$$

The recursion relation for the quantities, G_m , is now found to be

$$G_m = \sum_{k=1}^{m-1} I_k G_{m-k-1} \quad \text{for } m \geq 2 \quad (31)$$

with $G_0 = 1$, $G_1 = 0$. The power series,

$$\left. \begin{aligned} E_n &= \sum_{m=0}^{\infty} \lambda^m F_m \\ \text{with } F_m &= \sum_{k=0}^m J_k G_{m-k} \end{aligned} \right\} \quad (32)$$

is then valid for all values of λ .

Equation (32) evaluated at $\lambda = 1$ now gives $E_n(\epsilon)$ implicitly at any point in the complex ϵ -plane which may be connected to the origin by a curve along which H_1 is a bounded operator. The drawback involved in the use of this formula is that at $\lambda = 1$ it is in general necessary to retain a large number of terms in order to obtain a reasonable approximation to E_n . In the next section we shall discuss an asymptotic approximation to the characteristic equation (Equation (32)).

V. ASYMPTOTIC APPROXIMATION OF THE CHARACTERISTIC EQUATION

Let us reconsider Equations (27) and (28). What we shall do is to find asymptotic approximations of both the numerator and the denominator of the characteristic equation,

$$\frac{E_n \sum_{m=0}^{\infty} \lambda^m I_m - \sum_{m=0}^{\infty} \lambda^m J_m}{\sum_{m=0}^{\infty} \lambda^m I_m} = 0 \quad (33)$$

The reason for keeping both numerator and denominator of the above equation is that when the numerator formally diverges the denominator diverges in the same manner. As we shall see, it is possible to write the asymptotic approximations in such a manner that the divergences cancel and a well defined characteristic equation results.

Now from Equations (13) and (28) we have:

$$\begin{aligned} I_{m+1} &= \sum_{l=1}^m \left(K^{+l} K^{m+1-l} \right)_{nn} \\ &= \left(\sum_{i \neq n} \right)^m \frac{H_{nl_1} H_{l_1 l_2} \dots H_{l_m n}}{(e_{l_1} - E_n) \dots (e_{l_m} - E_n)} \left(\frac{1}{e_{l_1} - E_n} + \dots + \frac{1}{e_{l_m} - E_n} \right) \\ &= - \sum_{j \neq n} \frac{\partial}{\partial (e_j - E_n)} \left\{ \left(\sum_{i \neq n} \right)^m \frac{H_{nl_1} \dots H_{l_m n}}{(e_{l_1} - E_n) \dots (e_{l_m} - E_n)} \right\} \\ &= - \sum_{j \neq n} \frac{\partial}{\partial (e_n - E_n)} \left(u^{+L^{m-1}} u \right) \end{aligned} \quad (34)$$

for $m \geq 1$, where the vectors u^+ and u and the matrix L are defined as

$$u_l^+ = \frac{H_{nl_1}}{\sqrt{e_l - E_n}}, \quad u_l = \frac{H_{ln}}{\sqrt{e_l - E_n}},$$

$$L_{ll'} = \frac{H_{ll'}}{\sqrt{e_l - E_n} \sqrt{e_{l'} - E_n}}; \quad l, l' \neq n \quad (35)$$

(In the above definitions the branch of $\sqrt{e_l - E_n}$ is to be assigned in any consistent manner, e.g.

$$0 \leq \arg(\sqrt{e_l - E_n}) < \bar{u} \quad (36)$$

As will be seen presently the results are independent of Equation (36).)

Also,

$$J_{m+1} = \sum_{l=1}^m (K^{+l} H_0 K^{m+1-l})_{nn} - \sum_{l=0}^m (K^{+l} H_1 K^{m-l})$$

$$= \left(\sum_{i \neq n} \right)^m \frac{H_{nl_1} \dots H_{l_m n}}{(e_{l_1} - E_n) \dots (e_{l_m} - E_n)} \left(\frac{e_{l_1}}{e_{l_1} - E_n} + \dots + \frac{e_{l_m}}{e_{l_m} - E_n} \right)$$

$$- (m+1) \left(\sum_{i=n} \right)^m \frac{H_{nl_1} \dots H_{l_m n}}{(e_{l_1} - E_n) \dots (e_{l_m} - E_n)}$$

$$= E_n I_{m+1} - (u^+ L^{m-1} u) \quad (37)$$

for $m \geq 1$.

The use of Equations (35) and (37) together with

$$I_0 = 1, \quad I_1 = 0, \quad J_0 = e_n, \quad J_1 = -H_{nn} \quad (38)$$

gives the characteristic equation the form

$$\frac{(E_n - \epsilon_n + \lambda H_{nn}) + \lambda^2 \sum_{m=0}^{\infty} \lambda^m (u^+ L^m u)}{1 - \lambda^2 \sum_{m=0}^{\infty} \lambda^m \left\{ \sum_{j \neq n} \frac{\partial}{\partial (\epsilon_j - E_n)} (u^+ L^m u) \right\}} = 0 \quad (39)$$

(It is interesting to note that setting the numerator of Equation (39) equal to zero gives exactly the result which would be obtained by the Brillouin-Wigner theory.)

Now Equation (39), as it stands, still represents an ordering of terms in the characteristic equation in powers of λ . Since we wish to take $\lambda = 1$ this is clearly the wrong ordering, and consequently we propose to order terms in Equation (39) according to quantities whose smallness is independent of the value of λ . The small quantities we shall use are in fact the elements of u^+, u and L . In order to make the point clearer, let us make the transformation

$$\lambda \rightarrow \left(\frac{\lambda}{\lambda_0} \right), \quad H_1 \rightarrow \lambda_0 H_1 \quad (40)$$

where

$$\lambda_0 \equiv \text{Max}_{l, l'} H_{ll'} \quad (41)$$

(One, but not both, of l and l' may be n .)

Equation (39) is formally unchanged by the above transformation, but the matrix elements H_{nl} and $H_{ll'}$ are now bounded above by 1 and the original problem is recovered for $\lambda = \lambda_0$. Now for finite values of E_n there will be an integer, l_n , for which

$$|\sqrt{\epsilon_{l_i} - E_n}| < |\sqrt{\epsilon_{l_{i+1}} - E_n}| \quad \text{when } l_i > l_n. \quad (42)$$

Thus we may attempt to formulate an approximation based on the smallness of the matrix elements of L which follows from

$$|L_{ll'}| \leq \frac{1}{|\epsilon_l - E_n|} \quad \text{for } l' \geq l > l_n \quad . \quad (43)$$

In what follows we shall speak only of the case where $n = 0$; the generalization to cases where $n \neq 0$ involves essentially nothing new.

Now, suppose for the moment that

$$|\sqrt{\epsilon_1 - E_0}| < |\sqrt{\epsilon_2 - E_0}| \quad . \quad (44)$$

Then

$$\frac{(E_0 - \epsilon_0 + \lambda H_{00}) + \lambda^2 \left[\frac{H_{01} H_{10}}{\epsilon_1 - E_0} \sum_{m=0}^{\infty} (\lambda L_{11})^m + o\left(\frac{1}{\epsilon_2 - E_0}\right) \right]}{1 - \lambda^2 \left[\frac{\partial}{\partial(\epsilon_1 - E_0)} \left\{ \frac{H_{01} H_{10}}{\epsilon_1 - E_0} \sum_{m=0}^{\infty} (\lambda L_{11})^m \right\} + o\left(\frac{1}{\epsilon_2 - E_0}\right) \right]} = 0 \quad (45)$$

provides an asymptotic approximation (valid in the limit of large separation of unperturbed eigenvalues) to the characteristic equation. In order to evaluate the result of the asymptotic approximation at $\lambda = \lambda_0$, we find the analytic continuation of both the numerator and the denominator. In the case at hand this is a particularly simple task since we need only sum the series,

$$\sum_{m=0}^{\infty} (\lambda L_{11})^m = \frac{1}{1 - \lambda L_{11}} = \frac{1}{1 - \frac{\lambda H_{11}}{\epsilon_1 - E_0}} \quad . \quad (46)$$

Equation (45) now becomes

$$\frac{\left(E_o - \varepsilon_o + \lambda H_{oo}\right) - \lambda^2 \left[\frac{H_{o1} H_{1o}}{E_o - \varepsilon_1 + \lambda H_{11}} + O\left(\frac{1}{\varepsilon_2 - E_o}\right) \right]}{1 + \lambda^2 \left[\frac{H_{o1} H_{1o}}{(E_o - \varepsilon_1 + \lambda H_{11})^2} + O\left(\frac{1}{\varepsilon_2 - E_o}\right) \right]} = 0 \quad (47)$$

From the above equation (47) it is clear that the pole at λ_1 such that $E_o = \varepsilon_o - \lambda_1 H_{11}$ is not a singularity of the characteristic equation even though λ_1 is (within our approximation) a point at which the Brillouin-Wigner theory diverges. If we now multiply through by one troublesome factor and forget about the denominator we obtain the simpler characteristic equation,

$$\left(E_o - \varepsilon_o + \lambda H_{oo}\right) \left(E_o - \varepsilon_1 + \lambda H_{11}\right) = \lambda^2 \left[H_{o1} H_{1o} + O\left(\frac{1}{\varepsilon_2 - E_o}\right) \right] \quad (48)$$

(Note that within the accuracy implied by the above equation there is no value λ_1 for which the Brillouin-Wigner formula diverges unless $H_{o1} H_{1o} = 0$. This apparent anomaly is peculiar to the lowest approximation of this type.)

Equation (48) represents only the lowest order of approximation in this formulation. If it should turn out to be necessary to evaluate the characteristic equation at a value of λ for which Equation (44) is violated, we would naturally wish to retain more terms in our asymptotic approximation. It turns out to be possible to provide the analytic continuations necessary for the evaluation of the asymptotic approximations of arbitrarily high order. Consider, for example, the case where terms up to and including $O[(\varepsilon_N - E_o)^{-1}]$ are retained; then

Equation (39) becomes

$$\frac{(E_0 - \epsilon_0 + \lambda H_{00}) + \lambda^2 \left[\sum_{m=0}^{\infty} \lambda^m (u_N^+ L_N^m u_N) + O\left(\frac{1}{\epsilon_{N+1} - E_0}\right) \right]}{1 - \lambda^2 \left[\sum_{j=1}^N \delta(\epsilon_j - E_0) \left\{ \sum_{m=0}^{\infty} \lambda^m (u_N^+ L_N^m u_N) \right\} + O\left(\frac{1}{\epsilon_{N+1} - E_0}\right) \right]} = 0 \quad (49)$$

where

$$\begin{aligned} u_N^+ &= \left(\frac{H_{01}}{\sqrt{\epsilon_1 - E_0}}, \dots, \frac{H_{0N}}{\sqrt{\epsilon_N - E_0}} \right) \\ u_N &= \begin{pmatrix} \frac{H_{10}}{\sqrt{\epsilon_1 - E_0}} \\ \vdots \\ \frac{H_{N0}}{\sqrt{\epsilon_N - E_0}} \end{pmatrix} \end{aligned} \quad (50)$$

and

$$L_N = \begin{pmatrix} \frac{H_{11}}{\epsilon_1 - E_0} & \dots & \frac{H_{1N}}{\sqrt{\epsilon_1 - E_0} \sqrt{\epsilon_N - E_0}} \\ \vdots & \ddots & \vdots \\ \frac{H_{N1}}{\sqrt{\epsilon_N - E_0} \sqrt{\epsilon_1 - E_0}} & \dots & \frac{H_{NN}}{\epsilon_N - E_0} \end{pmatrix}$$

Now for sufficiently small values of λ we have

$$\sum_{m=0}^{\infty} (\lambda L_N)^m = (I_N - \lambda L_N)^{-1} \quad (51)$$

where I_N is the $N \times N$ unit matrix. The analytic continuation of Equation (49) is thus given by

$$\frac{(E_0 - \epsilon_0 + \lambda H_{00}) + \lambda^2 \left[u_N^+ (I_N - \lambda L_N)^{-1} u_N \right] + O\left(\frac{1}{\epsilon_{NH} - E_0}\right)}{1 - \lambda^2 \left[\sum_{j=1}^N \frac{\partial}{\partial(\epsilon_j - E_0)} \left(u_N^+ (I_N - \lambda L_N)^{-1} u_N \right) + O\left(\frac{1}{\epsilon_{N+1} - E_0}\right) \right]} = 0 \quad (52)$$

where the matrix

$$(I_N - \lambda L_N)^{-1} = \frac{M_N}{\Delta N} \quad (53)$$

is to be computed by Kramer's rule, i.e.

$$\left. \begin{aligned} \Delta N &= \det |I_N - \lambda L_N| \\ (M_N)_{ij} &= \text{Cofactor } (I_N - \lambda L_N)_{ji} \end{aligned} \right\} \quad (54)$$

The second and third approximations may now be computed with relative ease; they are to be evaluated at $\lambda = \lambda_0$, with the following results in terms of the original matrix elements derived from $H_1(\epsilon)$:

1st approximation

$$M_0 = \frac{H_{01} H_{10}}{M_1}, \quad (55)$$

2nd approximation

$$M_0 = \frac{H_{01}}{M_1 - \frac{H_{12} H_{21}}{M_1}} \left(H_{10} + \frac{H_{12}}{M_2} H_{20} \right) + \frac{H_{02}}{M_2 - \frac{H_{21} H_{12}}{M_1}} \left(H_{20} + \frac{H_{21}}{M_1} H_{10} \right) \quad (56)$$

3rd approximation

$$M_0 = \frac{H_{01}}{M_1 - \frac{H_{12}H_{21}}{M_2 - \frac{H_{23}H_{32}}{M_3}} - \frac{H_{13}H_{31}}{M_3 - \frac{H_{32}H_{23}}{M_2}}} \left[H_{10} + \frac{H_{12}}{M_2 - \frac{H_{23}H_{32}}{M_3}} (H_{20} + \frac{H_{23}}{M_3} H_{30}) - \frac{H_{13}}{M_3 - \frac{H_{32}H_{23}}{M_2}} (H_{30} + \frac{H_{32}}{M_2} H_{20}) \right] +$$

+ (two terms obtained by cyclic permutation of the indices 1 2 3),

(57)

where

$$M_j = \epsilon_j - E_0 - \lambda H_{jj} \quad (58)$$

From the foregoing formulae it may be inferred that the N^{th} approximation to the characteristic equation may be written in the form

$$M_0 = C_N \{ T_{0,1} \dots N \} \quad (59)$$

where C_N denotes a sum over cyclic permutations of the indices following the comma, and

$$\left. \begin{aligned} T_{1,j_1 \dots j_P} &= \frac{H_{1j_1}}{M_{j_1,j_2 \dots j_P}} \left(H_{j_1 0} + C_{P-1} \{ T_{j_1,j_2 \dots j_P} \} \right) \\ M_{j_1,j_2 \dots j_P} &= M_{j_1} - C_{P-1} \left\{ \frac{H_{j_1 j_2} H_{j_2 j_1}}{M_{j_2,j_3 \dots j_P}} \right\} \\ M_j &= \epsilon_j - E_0 - H_{jj} \end{aligned} \right\} \quad (60)$$

It can be shown that an equivalent way of writing the N^{th} approximation is,

$$M_0 = \sum_{j_1=1}^{N'} \frac{H_{0j_1} H_{j_1 0}}{\mu_{j_1}^{(N)}} + \sum_{j_1=1}^{N'} \sum_{j_2=1}^{N'} \frac{H_{0j_1} H_{j_1 j_2} H_{j_2 0}}{\mu_{j_1}^{(N)} \mu_{j_1 j_2}^{(N)}} + \dots +$$

$$+ \sum_{j_1=1}^{N'} \dots \sum_{j_N=1}^{N'} \frac{H_{0j_1} \dots H_{j_N 0}}{\mu_{j_1}^{(N)} \dots \mu_{j_1 \dots j_N}^{(N)}} \quad (61)$$

with

$$\mu_{j_1 \dots j_P}^{(N)} = M_{j_P} - \sum_{j_{P+1}=1}^{N'} \frac{H_{j_P j_{P+1}} H_{j_{P+1} j_P}}{\mu_{j_1 \dots j_{P+1}}^{(N)}}$$

subject to the restrictions $j_m \neq j_l$ when $m > l$. The above form is exactly the result obtained if Feenberg's N^{th} approximation is truncated by deleting every term which contains an index greater than N ; in the limit where $N \rightarrow \infty$ the two results are the same.

Equations (55) through (61) have been derived by simply neglecting the denominator of Equation (52). Consequently, they are subject to the appearance of apparent divergence whenever $\Delta_N \rightarrow 0$. However, as discussed previously, multiplication of the foregoing equations by Δ_N to give the result,

$$(u_N^+ M_N u_N) = 0 \quad \text{when } \Delta_N = 0, \quad (62)$$

is a justifiable procedure. Thus we have a means of justifying the use of Feenberg's formula beyond the points where it diverges (where

$\Delta_n = 0$, c.f. Morse & Feshbach on this point) and, if desired, a procedure which may be used in the neighborhood of such points as well.

The convergence of our asymptotic approximations is of course not insured in general. A discussion of sufficient conditions for convergence would involve specification of the behavior of the unperturbed eigenvalues, $\mathcal{E}_n$, the degeneracies, w_n , and the matrix elements, H_{mn} , as m and n become large. Such a discussion is beyond the scope of this note.

VI. AN EXAMPLE

As an example of the manner in which the formulae of the preceding section may be applied, let us consider the Mathieu equation,

$$\frac{d^2\psi}{dx^2} - \epsilon \cos^2 x \psi = E\psi, \quad \psi(0) = \psi(2\pi) \quad (63)$$

We take as the unperturbed problem,

$$\frac{d^2u}{dx^2} = \epsilon u, \quad u(0) = u(2\pi), \quad (64)$$

with even solutions,

$$\left. \begin{aligned} u_0 &= \frac{1}{\sqrt{2\pi}}, & \epsilon_0 &= 0 \\ u_n &= \frac{1}{\sqrt{\pi}} \cos nx, & \epsilon_n &= -n^2, \end{aligned} \right\} \quad (65)$$

and add solutions,

$$v_n = \frac{1}{\sqrt{\pi}} \sin n\pi x, \quad \epsilon_n = -n^2. \quad (66)$$

Since $H_1 (= \cos^2 x)$ is symmetric about $x = \pi$ the odd eigenfunctions do not affect the perturbation of the even ones and vice versa. Let us confine our attention to even eigenfunctions; then we need consider only a nondegenerate problem with

$$\left(H_{mn} \right) = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{\sqrt{8}} & 0 & 0 & - & - & - \\ 0 & \frac{3}{8} & 0 & \frac{1}{4} & 0 & - & - & - \\ \frac{1}{\sqrt{8}} & 0 & \frac{1}{2} & 0 & \frac{1}{4} & - & - & - \\ 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 & - & - & - \\ 0 & 0 & \frac{1}{4} & 0 & \frac{1}{2} & - & - & - \\ - & - & - & - & - & & & \\ - & - & - & - & - & & & \\ - & - & - & - & - & & & \end{pmatrix} \quad (67)$$

From the formulae of the last section and the 'zereth' approximation ($M_0 = 0$) we obtain the following results:

e	$- E_0^{(0)}$	$- E_0^{(1)}$	$- E_0^{(2)}$	$- E_0^{(exact)}$
0	0	0	0	0
4	2	1.5505	1.54487	1.54486
36	18		~ 5.8	~ 6
100	50		~ 10.8	~ 10

The interesting point is that the 2nd approximation, which in this case gives the relatively simple cubic characteristic equation,

$$\left(E_0^{(2)} + \frac{1}{2}\epsilon\right) \left[1 - \frac{\epsilon^2}{16 \left(E_0^{(2)} + \frac{1}{2}\epsilon + 4\right) \left(E_0^{(2)} + \frac{1}{2}\epsilon + 16\right)} \right] \sim \frac{\epsilon^2}{8 \left(E_0^{(2)} + \frac{1}{2}\epsilon + 4\right)}, \quad (68)$$

gives a result which at $\epsilon = 4$ is two orders of magnitude more accurate than the corresponding quartic approximation derived from the Fredholm resolvent (c.f. Morse & Feshbach). In addition, at $\epsilon = 100$, a value 20 times greater than the radius of convergence of the Brillouin-Wigner formulation of this problem, the approximate result is still within 20% of the exact. In fact, as has been noted by Morse & Feshbach, Feenberg's formula gives an exact continued fraction representation of the eigenvalue in this case.

It might be argued that we have not given the method a fair test in taking a problem for which it converges. This, however, would miss the point -- the usefulness of this method lies in the ease of application on the one hand and the estimate of the error $\left(O\left(\frac{1}{\epsilon_{N+1} - E_0} \right) \right)$, on the other. Needless to say, the error estimates can be refined considerably in any specific problem.

Finally it should perhaps be noted that the computation of our second approximation requires the evaluation of at most nine distinct matrix elements. This is to be contrasted with the evaluation (in principle) of infinite series required in the more cumbersome theory based on Fredholm's resolvent.

VII. REFERENCES

Morse & Feshbach - Methods of Theoretical Physics, v.2,
Section 9.1.

Mikhlin - Integral Equations, Section 9.

Kumar - Perturbation Theory and the Nuclear Many-Body
Problem, Chapter 1.

Feenberg, E., Phys. Rev., v.74 (1948) p. 206.