CHAPTER 2

TIME-INDEPENDENT PERTURBATION THEORY

Synopsis:

In this chapter, we discuss certain aspects of the Rayleigh-Schrödinger (RS) and the Brillouin-Wigner (BW) perturbation theories (PT). In Sec.2.1., we start out with a brief outline of the convergence behaviour of RSPT pointing out some current problems followed by a discussion of our own work related to linearisation of polynomial perturbations. Section 2.2. presents our contribution to the problem: 'what is $H_o$?'. Here, we find out a systematic scheme of repartitioning the total Hamiltonian by exploiting a flexibility in the RS method and demonstrate the calculational advantages of some such repartitions. Special attention is paid to a particular repartition which corresponds to scaling of $H_o(H_o \rightarrow \mu H_o)$, and Sec.2.3. is devoted to this $\mu$ transformation topic where we test its workability to ground- and excited-states of bound-and quasi-bound Hamiltonians, try to analyse why it works and examine some recent strategies of evaluating the parameter $\mu$. Then, in Sec.2.4., we develop the BW theory and summarise some remarks on convergence properties, showing also that the finite-dimensional BW theory can always be made convergent for the ground state by a simple change of scale. Finally, Sec.2.5. is concerned with possible variants of BW's method based on ways of looking at the perturbed energy appearing in the denominators of correction-terms expressed in sum-over-states forms; here we discuss the Morse-Feshbach formulation of the
method, the upper-bound properties of odd-order BW energy-sums for the ground state for the variants and some useful inequalities which provide bracketing and error-estimates of the ground-state energy. How far the asymptotic RS sums are useful can also be verified through the use of such inequalities, as we have indicated.
2.1. Aspects of Convergence Of RSPT:

In PT, the first task is to split $H$ into two parts:

$$H = H_0 + \lambda V$$

($H_0$: 'simple', i.e. solvable; $V$: 'small'; $\lambda$: a natural or artificial parameter); the working scheme is then developed through power-series expansions of $\Psi$ and $\mathcal{R}$ ($H\Psi = \mathcal{R}\Psi$) in $\lambda$. Whether such expansions would converge depends on the 'smallness' of $V$. So, before computations, one should pay attention to criteria of 'regularity' of $H^1$ and lower bounds ($\lambda^2$) to the radius of convergence ($\mathcal{R}_c$) of RSPT in order to be convinced beforehand whether a perturbative approach would be proper. Usually two types of $V$s are considered:

1) $\mathcal{D}(V) \supseteq \mathcal{D}(H_0)$, $\|\psi\| \leq a + b \|H_0\psi\|$, $\|\phi\| = 1$, $\phi \in \mathcal{D}(H_0)$;

$$a, b > 0 \quad (D: \text{domain})$$

ii) $|\langle \phi | V | \phi \rangle| \leq a + b |\langle \phi | H_0 | \phi \rangle|$, $\|\phi\| = 1$, $\phi \in \overline{\mathcal{D}}$ (common domain of $H_0$ & $V$), $b > 0$, $\epsilon_0$ (lowest eigenvalue of $H_0$) $\rightarrow -\infty$.

Although, for some problems, one can check at the onset whether such conditions for the applicability of RSPT are satisfied (e.g. see Kato's work on $Z^{-1}$ atomic PT), it is admittedly a formidable task for most perturbations of practical interest.

More often, a rather naive approach* is followed: we proceed with assumed expansions (in $\lambda$) for $\Psi$ and $\mathcal{R}$ to evaluate the coefficients of $\lambda^n$ ($n \geq 1$), checking at the end whether such expansions are permitted (i.e. convergent) to some desired orders (finite). This practical approach, we shall see, causes troubles for non-linear problems, unless proper care is taken.

* It is based rightly on 'an act of faith rather than the result of knowledge'.
A simplification in the context of determining the RC of RSP2 (note: determination of RC is important, because perturbative results are meaningful only within the RC) has been achieved through the use of variation-perturbation (VP) schemes, of course leading, understandably, to an approximate value \( \lambda_0 \) of the RC. The idea involved here is as follows: if we have good variational approximations \( \phi \) and \( \omega \) to \( \psi \) and \( \delta \), respectively, then the power-series expansion of, say, \( \omega(\lambda) \) in \( \lambda \) would give us a radius \( \lambda_0 \) of its convergence which should approximate well the value \( \lambda_0 \), the true RC of \( \psi(\lambda) \). We shall not discuss elaborately this scheme (see ref.7 for a recent exhaustive account) but simply emphasize here that one should not be very hopeful about it. For example, the VP equations for \( \lambda_0 \) may lead to finite radii for, say, the anharmonic oscillator problem, the Stark problem, etc.; whereas, we know that for these problems \( \lambda_0 = 0 \).

The above discussion clearly indicates that we should pay more attention to find out simple criteria of divergence, rather than to compute \( \lambda_0 \) or approximate it—a much harder task. Noting that perturbative expansions are power-series expansions and that power series converge in circles, it is apparent that if \( \psi(\lambda) \) admits of a power expansion in \( \lambda \), so would \( \psi(-\lambda) \) also, implying absolute convergence. But, then, both \( \psi(\pm \lambda) \) should be well-defined, single-valued and finite. Hence, if it so happens that for some \( |\lambda| \gg \lambda_c \), either of \( \psi_o(\pm \lambda) \equiv \langle \phi_o^+ | H(\pm \lambda) | \phi_o \rangle \) (\( \| \phi_o \| = 1 \), \( \phi_o \): absolutely flexible trial function) is unbounded from below, PT would lose a variational meaning which means, in turn, the region of applicability of
RSPT is also defined by $|\lambda| < \lambda_0$, $\lambda_0 \ll \lambda_c$. This argument, used first by Dyson to justify the appearance of divergence in quantum electrodynamics, is now well-known as the sign-change argument (SCA); as a simple indicator (see also Sec.1.4.), it can be employed fruitfully except for polynomial perturbations (PP) where some difficulties, to be discussed below, arise.

Although the question of (absolute) convergence of RSPT is one important aspect of the study, we may note that very often we have to handle series like

$$Y(\lambda) = \sum_{i} y^{(i)} \lambda^i , \quad (2.1)$$

where it has been found that the sum to the numerically smallest term of the RHS of (2.1) the asymptotic sum $(Y_a)$, reasonably approximates $Y(\lambda)$ for small $|\lambda| - \lambda_0$; then say that (2.1) is an asymptotically convergent series (ACS), the error decreasing as $(|\lambda| - \lambda_0) \to 0^+$. A second pertinent question of no less importance also arises; it is concerned with a theoretical development leading to an asymptotic formulation of RSPT (ARSPT). An asymptotic series at the origin, i.e. when $\lambda_0 = 0$, is traditionally given after the Poincare notion:

$$\lim_{|\lambda| \to 0} |\lambda|^N \left| Y(\lambda) - \sum_{i=0}^N y^{(i)} \lambda^i \right| = 0 .$$

Let us note that there exist various approaches, other than the usual power-series procedure, to formulate RSPT, e.g. through the Weierstrass polynomial approximation theorem, repeated differentiations of $H(\lambda) \psi(\lambda) = B(\lambda) \psi(\lambda)$, etc. To frame RSPT in a form that includes ACS also within the theory, Krieger...
adopted the second strategy (for a criticism, see ref. 13). A particularly transparent formulation of ARSPT, put forward recently by Wilson and Sharma 4, is outlined below:

Consider the perturbed energy-eigenvalue equation

\[(H_0 + \lambda V)\psi(\lambda) = E(\lambda)\psi(\lambda),\] (2.2)

where \(H_0 \phi = \epsilon \phi\), and write

\[\psi(\lambda) = \phi + \psi^{(1)}\] (2.3)

to obtain

\[(H_0 + \lambda V)(\phi + \psi^{(1)}) = E(\phi + \psi^{(1)}).\] (2.4)

We choose the normalizations as

\[\langle \phi | \psi \rangle = \langle \psi | \psi \rangle = 1.\] (2.5)

Then, take the limit \(\lambda \rightarrow 0\) in (2.4), remembering that we wish to have in this limit the unperturbed eigenvalue equation \(H_0 \phi = \epsilon \phi\), to find from the left hand side (LHS)

\[\psi^{(1)} = R(0)\] (2.6)

(note: by \(R(0)\) we mean a function which obeys \(R(0) \rightarrow 0\) as \(\lambda \rightarrow 0\); \(R(n)\) means \(R(n) \lambda^n \rightarrow 0\) as \(\lambda \rightarrow 0\); \(\lambda\) is used throughout to imply a linearisation parameter introduced artificially) unless \(\psi^{(1)} =\) constant \(\times \phi\), which cannot be owing to (2.5). Now, from the RHS we get

\[\text{limit}_{\lambda \rightarrow 0} B(\phi + \psi^{(1)}) = \epsilon \phi\] (2.7)

which shows, for finite \(B\) (obvious physically), writing \(B = \epsilon + B^{(1)}\), that

\[B^{(1)} = R(0).\] (2.8)

So, (2.4) takes the form
\[(\hat{H}_0 + \lambda \nu) (\phi + \psi(1)) = (\epsilon + \mathcal{B}(1)) (\phi + \psi(1)) \quad (2.9)\]
or,
\[(\hat{H}_0 - \epsilon) \psi(1) = (\mathcal{B}(1) - \lambda \nu)(\phi + \psi(1)). \quad (2.10)\]

Taking inner product with \(\langle \phi |\), (2.10) gives
\[\mathcal{B}(1) = \lambda \langle \phi | \nu | \phi \rangle + \lambda \langle \phi | \nu | \psi(1) \rangle \quad (2.11)\]
\[= \lambda \epsilon (1) + \mathcal{B}(2), \text{ say}, \quad (2.12)\]
where, by (2.6), we find from (2.11) that
\[\mathcal{B}(2) = \mathcal{R}(1). \quad (2.13)\]

Putting (2.12) in (2.9), we then obtain
\[(\hat{H}_0 - \epsilon) \psi(1) = \lambda (\epsilon (1) - \nu)(\phi + \psi(1)) + \mathcal{B}(2)(\phi + \psi(1)); \quad (2.14)\]
writing again \(\psi(1) = \lambda \phi(1) + \psi(2)\) in (2.14) and defining \(\phi(1)\) by
\[(\hat{H}_0 - \epsilon) \phi(1) = (\epsilon (1) - \nu) \phi \quad (2.15)\]
we then find that \(\psi(2) = \mathcal{R}(1)\). Actually here we could have also taken some amount of \(\phi\) in \(\psi(2)\), say \(\mathcal{C} \phi\), but with no real difficulty because a suitable normalization requirement would then show\(^4\) that \(\mathcal{C} = \mathcal{R}(1)\). In this way we can go forward and stop the procedure at any desired step. For a detailed account, we refer to the original work\(^4\).

Selected problems and discussion

We shall now focus attention on some current problems of convergence in RSPT. Before summarizing the problems, some specific examples and observations may be mentioned; so we consi-
under the following Hamiltonians:

\[
H^1(\lambda) = (-1/2 \nu^2 - r^{-1}) + \lambda r^{-1} \tag{2.16}
\]

\[
H^2(\lambda) = (-\nu^2 + x^2 - 1) + (\lambda^2 x^4 + \lambda x^3 - 2\lambda x) \tag{2.17}
\]

\[
H^3(\lambda) = (-\nu^2 + x^2 - 1) + (\lambda^4 x^6 + \lambda^2 x^4 - 3\lambda^2 x^2) \tag{2.18}
\]

\[
H^4(\lambda) = (-1/2 \nu^2 - r^{-1}) + (2\lambda r + 2\lambda^2 r^2) \tag{2.19}
\]

The following observations are worth-noting:

(i) For \(H^1(\lambda)\), the \(E^1_0(\lambda)\) series is finite and hence converges for any \(\lambda\), but to some 'anti-bound' state if \(\lambda \gg 1\) (see ref.13).

(ii) For \(H^2(\lambda)\), \(E^2_0(\lambda) > 0\); but, RSPT gives \(E^2_0(\text{RSPT}) = \sum_n \epsilon^{(n)}_o \lambda^{2n}\), \(\epsilon^{(n)}_o(n \gg 0) = 0\).

Hence, the RS energy-series converges, but to a wrong answer (see ref.14).

(iii) For \(H^3(\lambda)\), \(\psi^3_0(\lambda) = \exp[-(1/2 x^2 + 1/4 \lambda^2 x^4)]\), \(E^3_0(\lambda) = 0\). Here, though the \(E^3_0(\text{RSPT})\) series in \(\lambda^2\) converges to the right answer, the \(\psi^3_0(\text{RSPT})\) series diverges (see again ref.14).

(iv) For \(H^4(\lambda)\), \(E^4_0(\lambda) = -\frac{1}{2} + 3\lambda\) and hence is convergent for any \(\lambda\), but leads to the correct answer only for \(+ |\lambda|\); for \(- |\lambda|\), it converges to a wrong answer (see ref.15).

(v) For \(H^1(\lambda)\), the SCA indicates divergence for \(|\lambda| \gg 1\), but the \(E^1_0(\lambda)\)-series in finite; for \(H^4(\lambda)\), \(i = 2,3,4\), the SCA applied to \(\lambda\) does not indicate divergence, and hence inapplicability of RSPT, but observations (ii) to (iv) go against this conclusion; for \(H^4(\lambda)\), \(i = 2,4\), application of the SCA to \(\lambda^2\) implies
divergence, but it is not at all clear whether the argument is
to be applied to $\lambda$, $\lambda^2$ or something else.

Now, we are in a position to point out the pertinent problems:

1. Is it possible that the RS series-expansions for $\Psi(\lambda)$ and
$E(\lambda)$ (and possibly for $\bar{W}_1(\lambda) = \langle \Psi(\lambda) | \bar{W}_1 | \Psi(\lambda) \rangle / \langle \Psi(\lambda) | \Psi(\lambda) \rangle$)
where $\bar{W}_1$ is some hermitian operator) can have different RC? (This
question seems to have attracted considerable attention$^{16,17}$; see
(iii) also)

2. Should a perturbation series always converge to something
meaningful? (See (i); ref.17; in fact, this question in the
context of atomic $Z^{-1}$ PT has led to the speculation of BSG$^{18}$ (see
also Sec.1.3.) for the He-atom problem)

3. Can a perturbation series converge to some wrong answer?
(See (ii) and (iv))

4. Whether or not is the SCA a faithful indicator of divergence?
(See (v))

5. Do the PP cause any additional difficulties? (Note that $H_i(\lambda)$,
i = 2,3,4, are all PP)

We shall see shortly that in PP problems the trouble is actually
of a different nature compared to those encountered in the linear
problems (e.g. in cases of degeneracy$^{19}$, singular perturbations$^{20}$,
etc.) due to non-analyticity or discontinuity$^{21}$. Here, the primary
difficulty is associated with the choice of the perturbation
parameter; polynomial problems should better be linearized first.
Let us first be concerned with the linear problem of $H^1(\lambda)$. We remark, to perturbatively show that the $E_o^1(\lambda)$-series is finite, one should determine the wavefunction-correction terms and these are unambiguously evaluated only when a proper normalization condition is chosen. But, choosing $\langle \psi^1_o(\lambda) | \psi^1_o(\lambda) \rangle = 1$, we see that for $\lambda \geq 1$, $\psi^1_o(\lambda)$ becomes unnormalizable and hence one must be within $\lambda = 1^{22}$. We emphasize here that the finiteness is not the real problem: the problem is to perturbatively obtain such a series. Keeping this in mind, the observation of ref.16 is also understandable. Moreover, it is the convergence of the $\psi(\lambda)$-series that should be given more importance, for any expectation-value-series can be derived from a knowledge of $\psi(\lambda)$. In fact, Silverman and van Leuven also emphasized this, but on a different ground. They opined that since the expectation values involve integrations in calculations, one may find the apparent RC for them very different from the RC of the $\psi(\lambda)$-series, because integrations 'may cancel some or all of the singularities' in $\psi(\lambda)$.

Now, we shall proceed to analyze the peculiarities of the PP problems. The observations (ii), (iii) and (iv) suggest that definitely we are beyond the purview of RSPT, because convergence to wrong answers, divergences or non-circular convergences can arise only then. Following Killingbeck, we note that the SCA dictates divergence for all these $H^i(\lambda)$ ($i = 2,3,4$) if we linearize the problems, i.e., consider the perturbation part $V(\lambda)$ as $\mu V(\lambda)$ and apply the SCA to $\mu$, where $\mu$ is an artificial parameter. Since this is in conformity with the above suggestion, it may be interesting to go more through this linearization procedure.
Actually, using the 'P-function1 technique of Dalgarno and Lewis, we find that the energy-series are of the following form (note: for \( H^4(\lambda) \), we take the results from Killingbeck):

\[
E^2_{0}(\text{RSPT}) = 0 + \mu \left( \frac{3}{4} \lambda^2 \right) - \mu^2 \left( \frac{3}{4} \lambda^2 + \frac{21}{16} \lambda^4 \right) + \ldots
\]

\[
E^3_{0}(\text{RSPT}) = 0 + \mu \left( \frac{15}{8} \lambda^4 \right) - \mu^2 \left( \frac{15}{8} \lambda^4 + \frac{225}{16} \lambda^6 + \frac{3425}{128} \lambda^8 \right) + \ldots
\]

\[
E^4_{0}(\text{RSPT}) = -\frac{1}{2} + \mu \left( 3 + 6 \lambda^2 \right) - \mu^2 \left( 6 \lambda^2 + 54 \lambda^4 + 129 \lambda^6 \right) + \ldots
\]

Let us note that the choice \( \mu = 1 \) results in a peculiar behaviour of these series. But, remembering that if a series converges absolutely at \( \mu = 1 \), it should also converge for \( \mu = -1 \), we see easily that none of such series are absolutely convergent at \( |\mu| = 1 \).

We also find, at \( \mu = 1 \), if we agree to omit the brackets, a severe cancellation of terms occurs and rearrangement in powers of \( \lambda \) gives us back the results quoted in the observations (ii), (iii) and (iv). So, the \( \lambda \)-series we get for \( E^4_{0}(\lambda) \) are not absolutely convergent. This is because a divergent series can never be transformed to an absolutely convergent series by simple rearrangement. In fact, if we are permitted to omit the brackets, what we obtain are conditionally convergent series. Thus, the argument that some sort of "destructive interference" results in "isolated changes" from divergent energy-series (in \( \mu \)) to "convergent" series (in \( \lambda \)) at \( \mu = 1 \) does not seem to be very meaningful because this "convergent" nature should not be considered in an absolute sense.

Accepting such apparently convergent nature of the above series, the observations are now understandable (note: RSPT gives correct answers only if the series concerned converges absolutely; for conditionally convergent series, rearrangements can give almost
any answer we like to have\textsuperscript{25}).

There remains, however, another problem. Note that, in spite of the divergent nature of all these series in terms of $\mu$, in cases the rearranged (at $\mu = 1$) $\lambda$-series (e.g., see observations (iii), and (iv) for $+\lambda$) work. But, do they mean anything? Regrettably, the answer is in the negative, especially keeping in mind the contradictions faced with $B^2_o(\lambda)$ and $B^4_o(-\lambda)$; the workability is just accidental. Actually, linearization demonstrates such an assertion conclusively, we have seen. Unfortunately, however, such a strategy is not generally adopted, e.g., in studies with screened Coulomb potentials\textsuperscript{27}, charmonium Hamiltonians\textsuperscript{28}, etc. and it is not clear what the computed data in such cases signify. Another aspect of linearization is that, it shows the full effect of the perturbation $V(\lambda)$ at each order; but, if we proceed with $\lambda$-series, partial effects are incorporated at different orders (i.e., if $V(\lambda) = \sum_{i=1}^{\infty} \lambda^i V_i$, the first-order equation in $\lambda$-series incorporates the effect of $V_1$ only, the second-order takes effects of $V_1$ and $V_2$ only, etc.).

Summarizing, our discussion leads to the following conclusions: (a) RC of the RS $\psi(\mu)$-series\textsuperscript{29} should be given more attention to define the region of applicability of RSPT (where $\mu$ is the linearization parameter), (b) a perturbative series should always converge, in absolute sense, to something meaningful (i.e., if the series concerned is not absolutely convergent, then only it can apparently "converge" to something meaningless), (c) convergence to wrong answers\textsuperscript{30} of perturbative series imply only that one has uncritically extended the scope of RSPT; for polynomial problems, a due linearization and subsequent checking of the
absolutely convergent nature of the series so obtained can prevent us from confronting such disastrous situations, (d) PP problems should first be linearized to avoid any confusion, and (e) the SCA is to be applied to the linearization parameter.

2.2. Variants Of RSPT:

To avoid the appearance of slowly convergent, asymptotic or divergent series in course of a perturbative (RS) development, we can, in cases, adopt the strategy of repartitioning. This implies that instead of the conventional choice

\[ H = H_0 + \lambda V, \quad (2.21) \]

we choose an alternative partition:

\[ H = h_0 + v, \quad (2.22) \]

fruitfully (for 'internal' perturbations) with the hope of extracting important higher-order effects of (2.21) within lower orders of the repartitioned scheme given by (2.22) so that a low-order computation, as is usually feasible, would offer us good-quality results. Some of the advantages of this approach (over the other strategy of utilizing series-handling schemes, to be discussed in Ch.3) are: (i) requirements of only low-order calculations, (ii) direct correspondence with remodulations of the more fundamental wavefunction-series, and (iii) provision of a direct feel for the starting premise of approximation. However, the efficiency of a repartitioned scheme is understandably dependent on the choice of \( h_0 \). We shall see here how a wise choice of \( h_0 \) can be made.

It is apparently a problem with (2.22) that there is no
natural parameter attached to $v$. We may, of course, incorporate an artificial linearization parameter, to be set ultimately equal to unity, in (2.22) to keep the track of 'orders'. As regards $h_0$, sometimes, it has been found useful$^{32,33}$ to choose

$$h_0 = H_0 + A$$  \hspace{1cm} (2.23)

where $A$ is chosen, for $m$th-state perturbations, such that

$$A \Phi_m = 0$$  \hspace{1cm} (2.24)

is satisfied. A particularly familiar form of 'A' — the Epstein-Nesbet (BN) choice$^{34}$ — has been studied at length in ref.$^{35}$ (especially relative to the original Moller-Plesset (MP) variety). Kutzelnigg$^{36}$ considered the choice (2.33), with (2.24), as a generalized BN approach to RSPT in terms of (2.22). Wilson$^{37}$ has shown the effectiveness of MP and BN partitions by choosing a linear combination of them as the zero-order Hamiltonian.

Our work$^{38}$ in this issue of finding out useful $h_0$ is based on a flexibility — a freedom of choice for the set of orthogonality integrals between zeroth and higher-order perturbation functions — in RSPT which has been used elegantly by Amos$^{39}$ to interpret the formation of $[N+1/N]$ PA. We wish to present, in what follows, a systematic procedure of deriving a number of useful $h_0$, some of which are, however, known. An advantage of our approach is that we do not primarily require any intuitive ground, a posteriori calculational demonstration or a variational premise to justify the use of some particular repartitioned RSPT.
2.2.1. **On a flexibility in RSPT:**

Using (2.2), (2.21) and the expansions, say for the ground state,

\[
\Psi_o(\lambda) = \phi_o + \lambda \phi_o^{(1)} + \lambda^2 \phi_o^{(2)} + \ldots, \tag{2.25}
\]

\[
E_o(\lambda) = \epsilon_o + \lambda \epsilon_o^{(1)} + \lambda^2 \epsilon_o^{(2)} + \ldots, \tag{2.26}
\]

we seek in RSPT a solution of \( H(\lambda) \Psi_o(\lambda) = E_o(\lambda) \Psi_o(\lambda) \). Two features may have been noted: (i) a lack of uniqueness of the orthogonality integrals \( S_i = \langle \phi_o^{(i)} | \phi_o \rangle \), and (ii) invariance of \( \epsilon_o^{(i)} \) on the choice of the former. If we choose \( \| \phi_o \| = 1 \) and \( \langle \phi_o^{(i)} | \psi_o(\lambda) \rangle = 1 \) (intermediate normalization), a restriction on \( \phi_o^{(i)} \), denoted by \( \tilde{\phi}_o^{(i)} \), is imposed: \( \langle \phi_o^{(i)} | \tilde{\phi}_o^{(i)} \rangle = 0 \). So, the choice, e.g.,

\[
\phi_o^{(1)} = \tilde{\phi}_o^{(1)} + S_1 \phi_o \tag{2.27}
\]

in (2.25) implies effectively a different norm of \( \psi_o(\lambda) \). Finite-order sums of (2.26) also shows dependence of \( E_o(\lambda) \) on \{\( S_i \}\}. For example, if we consider that the \( E_o(\lambda) \)-series is obtained from

\[
H \psi_o(\lambda) = E_o(\lambda) \psi_o(\lambda) \tag{2.28}
\]

by premultiplying (2.28) with \( \phi_o^* \) and integrating \( 39 \), we get

\[
E_o(\lambda) = \epsilon_o + \lambda \epsilon_o^{(1)} + \lambda^2 \epsilon_o^{(2)} (1 + \lambda S_1)^{-1} \tag{2.29}
\]

where (2.25) has been used to first order, with (2.27). Now, (2.29) shows the desired dependence. It may be seen that (2.29) agrees with (2.26) to second order for any \( S_1 \), but agreement to third order can be achieved if we choose \( S_1 = -\epsilon_o^{(3)} / \epsilon_o^{(2)} \) and
in that case (2.29) would denote the \([2/1]\) Padé approximant to (2.26). Indeed, Amos\(^39\) justified the formation of \([N+1/N]\) Padé approximants in such a way. Here, however, a more general standpoint would be taken.

2.2.2. A generalized strategy:

We first note from (2.28) that

\[
\bar{E}_o(\lambda) = \frac{\langle \phi | H | \psi_0(\lambda) \rangle}{\langle \phi | \psi_0(\lambda) \rangle} \tag{2.30}
\]

holds for any suitable \(\phi\). But, if some approximation to \(\psi_0(\lambda)\), say

\[
\psi_o(n) = \phi_o + \sum_{i=1}^{n} \lambda^n \phi_i(n), \tag{2.31}
\]

is used in (2.30), with \(n\) embedded parameters \(\{S_i\}\) \((i = 1, \ldots, n)\), we find

\[
\bar{E}_o(\phi) = \frac{\langle \phi | H | \psi_o(n) \rangle}{\langle \phi | \psi_o(n) \rangle} \tag{2.32}
\]

i.e., \(\bar{E}_o\) evaluated through (2.32) will depend on \(\phi\) and will naturally agree with (2.26) up to some finite order for a fixed \(\phi\). So, we can adopt the following strategy: (i) choose various \(\phi\) s to obtain expressions for \(\bar{E}_o(\phi)\), (ii) focus attention on that equation which has a direct correspondence with the starting zero-order Hamiltonian, (iii) choose \(\{S_i\}\) such that the \(\bar{E}_o(\phi)\) for other \(\phi\)'s agree with (2.26) to at least \(n\) orders better than the prevalent ones, (iv) use these \(\{S_i\}\) in the equation obtained in (ii) and interpret the results in terms of a new partitioning.
Applications

Let us be concerned here only with the flexibility associated with $S_1$ in

$$\Psi_o(1) = \Phi_o + \lambda \left( \Phi_o^{(1)} + S_1 \Phi_o \right). \quad (2.33)$$

Then, from (2.32), we get the following equations:

$$\overline{E}_o(\Phi_1) = \epsilon_o + \lambda \left( \epsilon_o^{(1)} + \epsilon_o^{(2)} + \epsilon_o^{(3)} \right) \left( S_1 \epsilon_o^{(2)} + \epsilon_o^{(3)} \right), \quad (2.34)$$

$$\overline{E}_o(\Phi_o^{(1)}) = \epsilon_o + \lambda \epsilon_o^{(1)} + \lambda S_{11}^{-1} (S_1 \epsilon_o^{(2)} + \epsilon_o^{(3)}), \quad \epsilon_o^{(1)} = \Phi_o^{(1)}, \quad (2.35)$$

$$\overline{E}_o(\Psi_o(1)) = \epsilon_o + \lambda \epsilon_o^{(1)} + \lambda^2 \epsilon_o^{(2)} \left( \frac{1 + \lambda S_1}{1 + \lambda S_1} \right) - 1, \quad \epsilon_o^{(1)} = \left( 1 + \lambda S_1 \right)^{-1}. \quad (2.36)$$

and we already have from (2.29), with slight changes in notations,

$$\overline{E}_o(\Phi_o) = \epsilon_o + \lambda \epsilon_o^{(1)} + \lambda^2 \epsilon_o^{(2)} \left( 1 + \lambda S_1 \right)^{-1}. \quad (2.37)$$

We may note, it is eq.(2.37) that bears a direct correspondence with partitioning for this equation only has a modified second-order correction term (note : zero-plus first-order correction is invariant to partitioning of $H$), with no remainder left. So, we shall pay attention to (2.37) for various values of $S_1$ to be evaluated as follows:

A. Eq.(2.34) : It agrees with (2.26) only for the zeroth term.

Case (i) Neglect the third term completely and choose

$$S_1 = \frac{\epsilon_o^{(1)}}{\epsilon_o^{(1)} - \epsilon_o^{(3)}}$$

so that (2.26) and (2.34) agree up to first order. Using this $S_1$ in (2.37), it may be easily checked that for a two-level problem the result is equivalent to working with
\[ h_0 = H_0 - \lambda \epsilon_0^{(1)} |\phi_1\rangle \langle \phi_1| \]  \hspace{1cm} (2.38)

This particular partitioning stands as an intermediate choice between a pure RSPT \((h_0 = H_0)\) and a BW scheme (where \(h_0(\lambda) - \epsilon_0\) replaces \(\lambda \epsilon_0^{(1)}\) at the HHS). In the formulation of GPT (see Sec.1.2.1) such a scheme is quite relevant.

**Case (ii)** Neglect only the off-diagonal contributions from the third term in (2.34), choose \(S_1\) such that (2.34) then agrees with (2.26) up to first order, put this \(S_1 = \left[ \epsilon_0^{(1)} - \epsilon_1^{(1)} / (\epsilon_0 - \epsilon_1) \right]\) in (2.37) to find that the resulting energy-series corresponds to formally working, again for a two-level problem, with

\[ h_0 = H_0 + \lambda \sum_{i} (\langle \phi_i | V | \phi_i \rangle - \langle \phi_i | V | \phi_0 \rangle) |\phi_0\rangle \langle \phi_i| \] \hspace{1cm} (2.39)

This is a familiar zero-order Hamiltonian and is known to be more advantageous (see, e.g., ref.42 in the context of Lowdin's partitioning); referring to many-body problems, (2.39) is quite reminiscent of the BN partitioning.

**B. Eq.(2.35)**: This agrees with (2.26) again only to zero order. Let us also note that \(S_{11}\) satisfies the inequality:

\[ S_{11} \gg \epsilon_0^{(2)} / (\epsilon_0 - \epsilon_1). \]

**Case (i)** If \(\epsilon_0^{(3)} = 0\), or if \(\epsilon_0^{(3)} / S_{11} \approx 0\), choose \(S_1 = \lambda S_{11}\) so that (2.35) agrees with (2.26) to second order; with this choice, we find from (2.37) that the following zero-order Hamiltonian follows:

\[ h_0 = (1 + \lambda^2 S_{11}) H_0 \] \hspace{1cm} (2.40)

Admittedly, this \(h_0\) is rather unfamiliar and needs calculational support; we shall demonstrate its adequacy below.
Case (ii) When the above condition is not satisfied, choose $S_1 = -e^{(3)}_0/e^{(2)}_o$ to remove the undesirable coefficient $e^{(3)}_o$ associated with the linear $\lambda$ term (whereas, actually, it should have been with $\lambda^3$) so that agreement with (2.26) to first order is ensured. This choice, however, corresponds to the $[2/1]$ Padé approximant to (2.26) when used in (2.37) which, in turn, implies

$$h_0 = (1 - \lambda e^{(3)}_0/e^{(2)}_o) h_0;$$

(2.41)

thus, an additional justification of the formation of the $[2/1]$ Padé is obtained (see also the discussion below eq.(2.29)). We shall discuss about it later (in Sec.2.3) in detail where (2.41) will be seen to follow from the so-called $\mu$ transformation procedure.43-45.

C. Eq.(2.36) : This eq. agrees with (2.26) up to first order if $S_1 = S_1(\lambda)$, and to third order, if $S_1 \neq S_1(\lambda)$, by Wigner's $(2n+1)$ rule. Here, if we try to match (2.36) with (2.26) to fourth order by finding out a suitable value of $S_1$, it may be checked that one will then obtain two values of $S_1$, and this nonuniqueness forbids us from profitably exploiting this equation.

2.2.3. Results and discussion :

It goes without saying that the success of perturbation theoretic approach to some problem is assessed, from a practical point of view, through the comparison of a lower order computation with the true value. Out of necessity, then, the choice of $h_0$ becomes crucial. But, without a good knowledge of the more important contribution from some specific terms in the perturbation
problem being studied, it is difficult to construct some good starting $h_0$. As we have seen, our approach affords directly an advantage from this viewpoint; each of the $\{h_0\}$ we have obtained clearly gives more importance to certain situations which has been reflected from the way of derivation. In what follows, hence, we consider some two-level model problems of the general form

$$H(\lambda) = \begin{pmatrix} H_{11} & \lambda \\ \lambda & H_{00} \end{pmatrix}, \quad H_{00} = \epsilon_0 + \lambda \epsilon_0^{(1)}, \quad H_{11} = \epsilon_1 + \lambda \epsilon_1^{(1)};$$

to assess the workability of the $\{h_0\}$ derived above. The parameters $\epsilon_0, \epsilon_0^{(1)}, \epsilon_1, \epsilon_1^{(1)}$ will be adjusted suitably depending on the $h_0$ to be used.

We first focus on the $h_0$ in (2.40) and choose $\epsilon_0^{(1)} = \epsilon_1^{(1)} = 0$ (note: this choice is in accordance with the requirement $\epsilon_0^{(3)} = 0$ for (2.40) to be applicable). Table 1 shows the results; the superiority of $h_0$ over $H_0$ is remarkable. Let us note that here $E_0(2)_{RS} = E_0(3)_{RS}$; $\lambda_0 = |\lambda_b|, \lambda_b = \mathcal{I}(\epsilon_1 - \epsilon_0)/\mathcal{I}$. The success of $h_0$ may at least partly be attributed to the increased spacing in the redefined zero-order spectrum (recall the 'rule of thumb' and note $(\epsilon'_1 - \epsilon'_0)$ in Table 1; compare also with the true perturbed level spacing). Table 2 is concerned with more general cases where allowance is given for varying degrees of coupling between the states; but, again the calculational success up to the convergence limit at least may be seen. Notably, for either of these two problems the $h_0$'s defined by (2.39) or (2.41) would result in no betterment relative to the
Table 1. Superiority of $h_o$ in (2.40) over $H_o$ for ground- and excited-state energy estimates, to second order:

$\epsilon_0 = 1, \epsilon_1 = 3, \epsilon_0^{(1)} = 0 = \epsilon_1^{(1)}, \lambda_0 = 1$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$E_{o(2)}^{RS}$</th>
<th>$E_{o(2)}^1$ (exact)</th>
<th>$E_{1(2)}^{RS}$</th>
<th>$E_{1(2)}^1$ (exact)</th>
<th>$(\epsilon_1 - \epsilon_0)$</th>
<th>$(\omega_1 - \omega_0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.98</td>
<td>0.9802</td>
<td>3.02</td>
<td>3.0198</td>
<td>2.02</td>
<td>2.04</td>
</tr>
<tr>
<td>0.4</td>
<td>0.92</td>
<td>0.9231</td>
<td>3.08</td>
<td>3.0769</td>
<td>2.08</td>
<td>2.15</td>
</tr>
<tr>
<td>0.6</td>
<td>0.82</td>
<td>0.8349</td>
<td>3.18</td>
<td>3.1651</td>
<td>2.16</td>
<td>2.33</td>
</tr>
<tr>
<td>0.8</td>
<td>0.68</td>
<td>0.7241</td>
<td>3.32</td>
<td>3.2759</td>
<td>2.32</td>
<td>2.56</td>
</tr>
<tr>
<td>1.0</td>
<td>0.5</td>
<td>0.5858</td>
<td>3.5</td>
<td>3.4142</td>
<td>2.5</td>
<td>2.83</td>
</tr>
<tr>
<td>1.2</td>
<td>0.28</td>
<td>0.4706</td>
<td>3.72</td>
<td>3.5294</td>
<td>2.72</td>
<td>3.12</td>
</tr>
<tr>
<td>1.4</td>
<td>0.02</td>
<td>0.3423</td>
<td>3.98</td>
<td>3.6577</td>
<td>2.98</td>
<td>3.44</td>
</tr>
<tr>
<td>1.5</td>
<td>0.125</td>
<td>0.26</td>
<td>4.125</td>
<td>3.8028</td>
<td>3.125</td>
<td>3.61</td>
</tr>
</tbody>
</table>

* Primed quantities correspond to $h_o$. 
Table 2*. Superiority of $h_o$ in (2.40) over $H_o$ for the ground state with a set of 2x2 matrix problems:

$$\epsilon_1^{(1)} = \epsilon^{(1)}, \lambda_o = \frac{1}{2}(\epsilon_1 - \epsilon_o)$$

<table>
<thead>
<tr>
<th>$\epsilon_o$</th>
<th>$\epsilon_1$</th>
<th>$\epsilon_o(1)$</th>
<th>$\lambda$</th>
<th>$B_o^{(2)}$</th>
<th>$B_o^{(2)*}$</th>
<th>$B_o^{(exact)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>6</td>
<td>2</td>
<td>+0.5</td>
<td>4.875</td>
<td>4.8824</td>
<td>4.8820</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.5</td>
<td>2.875</td>
<td>2.8824</td>
<td>2.8820</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+1</td>
<td>5.5</td>
<td>5.6</td>
<td>5.5858</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1</td>
<td>1.5</td>
<td>1.6</td>
<td>1.5858</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>3</td>
<td>+1</td>
<td>6.5</td>
<td>6.6</td>
<td>6.5858</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5858</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>2</td>
<td>+0.5</td>
<td>4.75</td>
<td>4.8</td>
<td>4.7929</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.5</td>
<td>2.75</td>
<td>2.8</td>
<td>2.7929</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>3</td>
<td>+0.25</td>
<td>4.6875</td>
<td>4.6912</td>
<td>4.6910</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.25</td>
<td>3.1875</td>
<td>3.1912</td>
<td>3.1910</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+0.5</td>
<td>5.25</td>
<td>5.3</td>
<td>5.2929</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.5</td>
<td>2.25</td>
<td>2.3</td>
<td>2.2929</td>
</tr>
</tbody>
</table>

* Primed quantities correspond to $h_o$. 
second-order RSPS with $H_0$. As regards the $H_0$ in eq. (2.38), it is again of no advantage over RSPS for the case with Table 1, and would, understandably, lead to an imbalance if applied to the problems in Table 2 because its derivation assumes vanishing contribution from all the $\varepsilon_1^{(1)}$'s $(i > 0)$, as may be verified from the condition under which the transition from (2.39) to (2.38) of occurs. It is interest to note that a second-order BW calculation gives the exact answer for the problem in Table 1, whereas the same for the problems in Table 2 is not quite promising to draw special attention.

Let us now be concerned with the $h_0$ given by (2.38) by choosing $\varepsilon_1^{(1)} = 0$. In Table 3, the results are displayed along with second- and third-order RS estimates with $H_0$. For this problem, the exact energy is

$$E_o = \frac{1}{2}(6 + \lambda) - \frac{1}{2}(5\lambda^2 - 4\lambda + 4)^{1/2}$$

so that the corresponding perturbative energy series would have a RC depending on the 'crossing' point of the states; here the complex pair of branch points

$$\lambda_b = \frac{2}{5} \pm \frac{4}{5}i$$

indeed show a rather strong coupling (note: $\text{Im } \lambda_b \neq \text{Re } \lambda_b$) of the ground state to the intruder state and it turns out that $\lambda_0 = 2/\sqrt{5}$. Table 3 immediately demonstrates the workability of the $h_0$ in (2.38). In fact, here $E'_o(2)$ represents the second-order energy estimate corresponding to either of the $h_0$'s defined by (2.38), (2.39) or (2.41), and thus also denotes the $[2/1]$ Fade approximant result. The following observations are worth-noting: (i) for $-|\lambda|$, $E'_o(2)$ is always better than $E_o(2)$,
Table 3*. Workability of the $h'_o$ in (2.38) for the ground state:

$\epsilon'_o = 2$, $\epsilon'_o^{(1)} = 1$, $\epsilon'_1 = 4$, $\epsilon'_1^{(1)} = 0$, $\lambda'_o = 2/\sqrt{5}$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$E'<em>o(2)</em>{RS}$</th>
<th>$E'_o^{(2)}$</th>
<th>$E'<em>o(3)</em>{RS}$</th>
<th>$E'_o(exact)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>2.18</td>
<td>2.1778</td>
<td>2.178</td>
<td>2.1780</td>
</tr>
<tr>
<td>-0.2</td>
<td>1.78</td>
<td>1.7818</td>
<td>1.782</td>
<td>1.7820</td>
</tr>
<tr>
<td>0.4</td>
<td>2.32</td>
<td>2.3</td>
<td>2.304</td>
<td>2.3056</td>
</tr>
<tr>
<td>-0.4</td>
<td>1.52</td>
<td>1.5333</td>
<td>1.536</td>
<td>1.5351</td>
</tr>
<tr>
<td>0.6</td>
<td>2.42</td>
<td>2.3429</td>
<td>2.371</td>
<td>2.3780</td>
</tr>
<tr>
<td>-0.6</td>
<td>1.22</td>
<td>1.2615</td>
<td>1.269</td>
<td>1.2682</td>
</tr>
<tr>
<td>0.8</td>
<td>2.48</td>
<td>2.2667</td>
<td>2.352</td>
<td>2.4</td>
</tr>
<tr>
<td>-0.8</td>
<td>0.88</td>
<td>0.9714</td>
<td>1.008</td>
<td>0.9875</td>
</tr>
<tr>
<td>1.0</td>
<td>2.5</td>
<td>2.0</td>
<td>2.25</td>
<td>2.3820</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.5</td>
<td>0.6667</td>
<td>0.75</td>
<td>0.6972</td>
</tr>
</tbody>
</table>

* Primed quantities correspond to $h'_o$. 
(ii) beyond $\lambda_0$, actually $E'_0(2)$ is even better than $E'_0(3)$;
(iii) for $+|\lambda|$, $E'_0(2)$ offer numbers lying between $E'_0(2)$ and $E'_0(3)$ for small $\lambda$, (iv) but, for moderately large $\lambda$, it fails to provide even reasonable estimates. A closer scrutiny reveals that this peculiar behaviour is due to the existence of extrema at $\lambda = (2\pi \sqrt{2})$ of $E'_0(2)$; thus, something more subtle happens near $\lambda \approx 0.586$ (maximum), and hence beyond. We have also seen that here again the second-order SW result is exact.

From the above discussion concerned with some simple 2x2 matrix problems, we see that a wiser choice of the starting $h_0$ may well offer as good-quality values within lower orders and the advantage may continue even beyond $\lambda_0$, the RG. Although, we may mention, this latter advantage is also provided by several types of summation techniques (see, e.g., Ch.3), we feel at the same time that since they do not have any direct connection with $H_0$, it becomes difficult to assess the status of the approximate starting premise.

Finally, the approach presented here can be applied to excited states as well without any additional difficulty; higher $S_i$'s ($i \geq 2$) can also be employed to achieve that end (of deriving profitable $h_0$). It may also seem comforting, if not tricky at the same time, to view the question of search for the 'very best' $h_0$ as echoing the most balanced choice of $\{S_i\}$.

2.3. $\mu$ Transformation In RSPT:

A special variant of RSPT, the so-called $\mu$ transformation procedure (leading to 'geometric approximation' (GA); also known as the Feenberg Scaling), requires, however, greater attention.
because of its wide popularity and hence we shall discuss it here separately.

We first note that (2.37), with \( S_1 \neq S_1(\lambda) \), implies only a scaling of the zero-order spectrum and letting \((1 + \lambda S_1) = \mu\), we find that this energy-expression corresponds to second-order result of

\[
h_\sigma = H_\sigma + (\mu - 1)(H_\sigma - \epsilon_0)
\]

(2.42)

or, in a projected form

\[
h_\sigma = H_\sigma + \sum \frac{1}{(1-\mu)(\epsilon - \epsilon_1)} |\phi_1\rangle \langle \phi_1| .
\]

(2.43)

A variational evaluation of \( \mu \) leading to upper bounds to energy-eigenvalues is available. But, here we like to discuss other ways of determining an optimum value of \( \mu \) such that some \( m \)th partial-sum of any scaled series (for energy, wavefunction or any property; obtained through the zero-order Hamiltonian

\[
h_\sigma (= \mu H_\sigma)
\]

agrees better with the function that the series tries to represent, than with the unscaled series (in terms of \( H_\sigma \)). As we shall see, such a scaling (see Sec. 1.4.) argument leads to (2.41) for energy-calculations as a special case.

2.3.1. The \( \mu \)-transformation:

For convenience, we shall drop here the factor \( \lambda \) attached to \( V \) (i.e. absorb \( \lambda \) within \( V \)) so that from the parent partition

\[
H = H_\sigma + V
\]

(2.44)

we can pass on to

\[
H = \mu H_\sigma + [V + (1-\mu)H_\sigma]
\]

\[
= h_\sigma + v
\]

(2.45)
to obtain the following results:

\[ E_i = (\varepsilon_i + \varepsilon_i^{(1)}) + \sum_{r=2}^{\infty} \varepsilon_i^{(r)} = (\varepsilon_i + \varepsilon_i^{(1)}) + \sum_{r=2}^{\infty} \varepsilon_i^{(r)}, \quad (2.46) \]

\[ \varepsilon_i^{(r)} = \mu^{-(r-1)} \sum_{j=2}^{r} \frac{(r-2)!}{(j-1)!} \mu^{-1} x^{-j} \epsilon_i^{(j)}; \quad (2.47) \]

\[ \langle \psi \rangle_i = w_i + \sum_{r=1}^{\infty} \frac{w_i^{(r)}}{w_i} = \sum_{r=1}^{\infty} w_i^{(r)} + w_i, \quad (2.48) \]

\[ w_i^{(r)} = \mu^{-r} \sum_{j=1}^{r-1} \frac{r!(r-1)!}{j!(r-j)!} \mu^{-1} x^{-j} w_i^{(j)}; \quad (2.49) \]

\[ \langle q \rangle_i = \sum_{r=0}^{\infty} \frac{q_i^{(r)}}{q_i} = \sum_{r=0}^{\infty} q_i^{(r)} \]

\[ q_i^{(r)} = \mu^{-(r+1)} \sum_{j=0}^{r} \frac{r!(r-1)!}{j!(r-j)!} \mu^{-1} x^{-j} q_i^{(j)}. \quad (2.51) \]

Here, \( \langle \psi \rangle_i \) and \( \langle q \rangle_i \) refer respectively to some first- and second-order properties for the \( i \)th state. The series for the wavefunction \( \psi_i \) has the form similar to (2.48) with scaled and unscaled functions related in a manner similar to (2.49).

2.3.2. Usual strategy:

Let us focus attention on the scaled energy-series (2.46) and implement the idea that a scaled series can be optimally modulated by choosing the scale parameter such that a specific term of the series concerned vanishes. We then find, a unique value of \( \mu \) can be obtained only if we choose \( \varepsilon_i^{(3)} = 0 \). This choice leads to

\[ \mu_i = 1 - \frac{\varepsilon_i^{(3)}}{\varepsilon_i^{(2)}}, \quad (2.52) \]
and considering then (2.46) as an AGS we neglect the remainder to obtain

\[ E_i \sim (\epsilon_i + \epsilon_i^{(1)}) + \epsilon_i^{(2)} (1 - \frac{\epsilon_i^{(3)}}{\epsilon_i^{(2)}})^{-1} = E_i(GA), \tag{2.53} \]

Various other interpretations of (2.53) are, however, available. Let us note that \( E_i(GA) \) is actually the \([2/1]\) Padé approximant to the unscaled \( E_i \)-series (see also around (2.41) for some other justifications). In what follows, we shall see the suitability of this particular GA by choosing especially problems involving excited and quasi-bound states (for which tests have not been done so far).

Some applications of (2.53)

Here, we choose the following cases (for detailed account, see ref. 43):

(i) \( H = -\frac{1}{2} \nabla^2 - \frac{1}{r} e^{-\lambda r} \tag{2.54} \)

(ii) \( H = -\frac{1}{2} \nabla^2 - \frac{1}{r} e^{-a_0 z^{1/3} r} \tag{2.55} \)

(iii) \( H = -\frac{1}{2} \nabla^2 - \frac{1}{r} + \lambda r \tag{2.56} \)

(iv) \( H = -\nabla^2 + \frac{1}{4} x^2 + \frac{1}{4} \lambda x^4 \tag{2.57} \)

Since (2.53) requires knowledge up to \( \epsilon_i^{(3)} \) of the conventional partitioning, results have always been compared with the third-order partial sums \( E_i(3) \) of unscaled series (note: for \( \mu \)th-order partial sums of scaled series, we use \( E_i(\mu) \) as the symbol). When exact results are not available, results are compared with higher-order perturbative values.

* A comment: neglect the off-diagonal terms in \( \epsilon_i^{(3)} \) and assume the ratio \( \langle \Phi_i \mid V \mid \Phi_i \rangle / \langle \Phi_i \mid \Phi_i \rangle \) to be independent of \( i \) to get the BNN h\(_0\), (2.80). This way of looking may be interesting from an interpretative standpoint.
Table 4 summarizes the results for $H$ in (2.54) - the Yukawa potential case. Here the energy-series is

$$E_n = \frac{1}{2} n^{-2} + \lambda - 2 \frac{n^2}{4} \lambda^2 + \frac{1}{12} n^2(5n^2 + 1) \lambda^3 - \cdots \quad (2.58)$$

when the Coulomb problem is considered as the unperturbed one (note: if we start from the Hulthen (quite close to the Yukawa) potential, we get $E_n(H)$, as tabulated).

Table 5 shows the results for $H$ in (2.55) which is used as a realistic model for 'screening' in atomic binding energy calculations $^{52,55}$. Here the energy-series is in terms of $a_0^2 z^{-2/3}$:

$$E_n = \frac{1}{2} n^2 - 2 - z^4/3 a_o - \frac{3}{4} n^2 z^{2/3} a_o^2 + \frac{1}{12} n^2(5n^2 + 1) a_o^3 - \cdots \quad (2.59)$$

which implies divergence for low $Z$, large $n$ or $a_o$ (note: as $Z \to \infty$, $j_n(E) \to 1$ and hence $h_o \to H_o$).

Cases (iii) and (iv) are primarily chosen to show the workability of (2.53) for QBS calculations. We see that (2.56) and (2.57) are somewhat less harmful Hamiltonians compared to, e.g., the Stark $H$, unbounded on either side of $\lambda = 0$ (for recent discussions, see refs. 56,57; for the use of LSM (see Ch.1) in this context, see ref.58). The energy-series for these cases, being respectively

$$E_o = - \frac{1}{2} + \frac{3}{2} \lambda - \frac{3}{2} \lambda^2 + \frac{27}{4} \lambda^3 - \frac{735}{16} \lambda^4 + \cdots \quad , \quad (2.60)$$

$$E_o = \frac{1}{2} + \frac{3}{4} \lambda - \frac{21}{8} \lambda^2 + \frac{333}{16} \lambda^3 - \frac{30885}{128} \lambda^4 + \cdots \quad , \quad (2.61)$$

are dangerously divergent $^{59}$. Tables 6 and 7 display the results of interest.
Table 4. Energy-estimates for (2.54).

<table>
<thead>
<tr>
<th>λ</th>
<th>μ_n(Σ)</th>
<th>E_n(3)</th>
<th>E_n(3A)</th>
<th>E_n(H)</th>
<th>Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) n = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>1.0013</td>
<td>-0.49800</td>
<td>-0.49800</td>
<td>-0.49800</td>
<td>-0.49800</td>
</tr>
<tr>
<td>0.01</td>
<td>1.0067</td>
<td>-0.49007</td>
<td>-0.49007</td>
<td>-0.49007</td>
<td>-0.49007</td>
</tr>
<tr>
<td>0.02</td>
<td>1.0133</td>
<td>-0.48030</td>
<td>-0.48030</td>
<td>-0.48030</td>
<td>-0.48030</td>
</tr>
<tr>
<td>0.05</td>
<td>1.0333</td>
<td>-0.45181</td>
<td>-0.45181</td>
<td>-0.45180</td>
<td>-0.45180</td>
</tr>
<tr>
<td>0.10</td>
<td>1.0667</td>
<td>-0.40700</td>
<td>-0.40703</td>
<td>-0.40705</td>
<td>-0.40705</td>
</tr>
<tr>
<td>0.20</td>
<td>1.1333</td>
<td>-0.32600</td>
<td>-0.32647</td>
<td>-0.32660</td>
<td>-0.32680</td>
</tr>
<tr>
<td>0.50</td>
<td>1.3333</td>
<td>-0.12500</td>
<td>-0.14062</td>
<td>-0.14360</td>
<td>-0.14810</td>
</tr>
<tr>
<td>(b) n = 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>1.0046</td>
<td>-0.12301</td>
<td>-0.12301</td>
<td>-0.12300</td>
<td>-0.12300</td>
</tr>
<tr>
<td>0.01</td>
<td>1.0233</td>
<td>-0.11529</td>
<td>-0.11530</td>
<td>-0.11530</td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>1.0467</td>
<td>-0.10615</td>
<td>-0.10615</td>
<td>-0.10615</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>1.1167</td>
<td>-0.08163</td>
<td>-0.08171</td>
<td>-0.08175</td>
<td></td>
</tr>
<tr>
<td>0.10</td>
<td>1.2333</td>
<td>-0.04800</td>
<td>-0.04933</td>
<td>-0.04993</td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>1.4667</td>
<td>-0.01100</td>
<td>-0.00682</td>
<td>-0.00851</td>
<td></td>
</tr>
<tr>
<td>(c) n = 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>1.0102</td>
<td>-0.05358</td>
<td>-0.05358</td>
<td>-0.05360</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>1.0511</td>
<td>-0.04620</td>
<td>-0.04620</td>
<td>-0.04620</td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>1.1022</td>
<td>-0.03798</td>
<td>-0.03801</td>
<td>-0.03802</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>1.2556</td>
<td>-0.01812</td>
<td>-0.01900</td>
<td>-0.01907</td>
<td></td>
</tr>
</tbody>
</table>

* See ref.54.
Table 5. Energy-estimates for (2.55).

<table>
<thead>
<tr>
<th>Z</th>
<th>$\mu_n(E)$</th>
<th>$R_n(3)$</th>
<th>$R_n(GA)$</th>
<th>$R_n(12)^*$</th>
<th>$R_n(20)^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a) $n = 1$, $a_0 = 0.85$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.1938</td>
<td>-177.1</td>
<td>-178.5</td>
<td>-179.7</td>
<td>-181.6</td>
</tr>
<tr>
<td>7</td>
<td>1.1549</td>
<td>-402.6</td>
<td>-403.7</td>
<td>-404.5</td>
<td>-404.5</td>
</tr>
<tr>
<td>9</td>
<td>1.1310</td>
<td>-724.5</td>
<td>-725.4</td>
<td>-726.2</td>
<td>-726.2</td>
</tr>
<tr>
<td>14</td>
<td>1.0976</td>
<td>-1963.6</td>
<td>-1964.2</td>
<td>-1964.9</td>
<td>-1964.9</td>
</tr>
<tr>
<td>20</td>
<td>1.0769</td>
<td>-4286.9</td>
<td>-4287.3</td>
<td>-4288.0</td>
<td>-4288.0</td>
</tr>
<tr>
<td></td>
<td>(b) $n = 2$, $a_0 = 0.70$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>1.1910</td>
<td>-1010.1</td>
<td>-1020.6</td>
<td>-1026.8</td>
<td>-1027.3</td>
</tr>
<tr>
<td>30</td>
<td>1.1692</td>
<td>-1606.6</td>
<td>-1615.9</td>
<td>-1621.3</td>
<td>-1621.2</td>
</tr>
<tr>
<td>40</td>
<td>1.1396</td>
<td>-3239.1</td>
<td>-3246.9</td>
<td>-3251.7</td>
<td>-3251.6</td>
</tr>
<tr>
<td></td>
<td>(c) $n = 3$, $a_0 = 0.50$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>1.2388</td>
<td>-668.2</td>
<td>-690.8</td>
<td>-708.1</td>
<td>-731.7</td>
</tr>
<tr>
<td>40</td>
<td>1.2185</td>
<td>-977.3</td>
<td>-998.3</td>
<td>-1010.9</td>
<td>-1014.2</td>
</tr>
<tr>
<td>50</td>
<td>1.1883</td>
<td>-1779.1</td>
<td>-1797.5</td>
<td>-1807.2</td>
<td>-1806.9</td>
</tr>
</tbody>
</table>

* See ref. 52.
Table 6. Energy-estimates for (2.56).

<table>
<thead>
<tr>
<th>λ</th>
<th>μ₀(3)</th>
<th>Third order</th>
<th>Least squares</th>
<th>Complex coordinate rotation</th>
<th>BW (Pade)</th>
<th>GA</th>
<th>Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>+0.05</td>
<td>1.225</td>
<td>0.4279</td>
<td>0.4266</td>
<td>0.4295</td>
<td>-0.4281</td>
<td>-0.4282</td>
<td></td>
</tr>
<tr>
<td>-0.02</td>
<td>0.910</td>
<td>0.5306</td>
<td>0.5303</td>
<td>0.5306</td>
<td>-0.5307</td>
<td>-0.5307</td>
<td></td>
</tr>
<tr>
<td>-0.04</td>
<td>0.820</td>
<td>0.5628</td>
<td>0.5610</td>
<td>0.5628</td>
<td>-0.5629</td>
<td>-0.5631</td>
<td></td>
</tr>
<tr>
<td>-0.06</td>
<td>0.730</td>
<td>0.5969</td>
<td>0.5923</td>
<td>0.5949</td>
<td>-0.5974</td>
<td>-0.5983</td>
<td></td>
</tr>
</tbody>
</table>

1. See ref. 60.
2. See ref. 61.
3. See ref. 62.

Table 7. Energy-estimates for (2.57).

<table>
<thead>
<tr>
<th>λ</th>
<th>μ₀(E)</th>
<th>Third order</th>
<th>Least squares</th>
<th>GA</th>
<th>Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>+0.05</td>
<td>1.396</td>
<td>0.5335</td>
<td>0.5331</td>
<td>0.5328</td>
<td>0.5327</td>
</tr>
<tr>
<td>-0.01</td>
<td>0.921</td>
<td>0.4922</td>
<td>0.4923</td>
<td>0.4922</td>
<td>0.4922</td>
</tr>
<tr>
<td>-0.03</td>
<td>0.762</td>
<td>0.4746</td>
<td>0.4750</td>
<td>0.4744</td>
<td>0.4742</td>
</tr>
<tr>
<td>-0.05</td>
<td>0.604</td>
<td>0.4533</td>
<td>0.4548</td>
<td>0.4516</td>
<td>0.4507</td>
</tr>
</tbody>
</table>

1. See ref. 60.

Table 8. Results for (2.62).

<table>
<thead>
<tr>
<th>λ</th>
<th>E₁-E₀</th>
<th>μ₀(E)</th>
<th>μ₀(μ)(E₁-E₀)</th>
<th>E₀(3)</th>
<th>E₀(GA)</th>
<th>E₀(exact)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+0.9</td>
<td>1.45</td>
<td>2.90</td>
<td>1.3943</td>
<td>1.3802</td>
<td>1.3784</td>
<td></td>
</tr>
<tr>
<td>-0.9</td>
<td>2.00</td>
<td>0.55</td>
<td>1.10</td>
<td>0.4032</td>
<td>0.3659</td>
<td>0.3162</td>
</tr>
<tr>
<td>+1.0</td>
<td>1.50</td>
<td>3.00</td>
<td>1.4375</td>
<td>1.4167</td>
<td>1.4142</td>
<td></td>
</tr>
<tr>
<td>+2.0</td>
<td>2.00</td>
<td>4.00</td>
<td>2.0000</td>
<td>1.7500</td>
<td>1.7321</td>
<td></td>
</tr>
<tr>
<td>+3.0</td>
<td>2.50</td>
<td>5.00</td>
<td>3.0625</td>
<td>2.0500</td>
<td>2.0000</td>
<td></td>
</tr>
</tbody>
</table>
From these results, we can say that the GA is quite useful for practical purposes. But the question is, why it is so successful. Let us note an important point that $E_0(GA)$ is closer to $E_0$ for $\mu_0(B) > 1$ than for $\mu_0(B) < 1$ (see Tables 6 and 7). We further verify it in Table 8 for the Hamiltonian

$$H = -\nabla^2 + x^2 + \lambda x^2$$

which gives

$$E_0 = (1 + \lambda)^{1/2}, \quad \lambda_0(BC) = 1. \quad (2.63)$$

Now, we may conclude that (i) results of the $\mu$-scaling procedure are usually always better, i.e., $E_n(GA)$ is better than $E_n(3)$; (ii) this success may partly be due to increased level-spacings (recall the 'rule of thumb', when $\mu_n(B) > 1$; otherwise, if $\mu_n(B) < 1$, solely the optimal remodulation effect is operative; (iii) $E_n(GA)$ is suitable also for excited states; (iv) for QBS calculations, the GA is better than other available methods. Thus, for $\mu_n(B) > 1$, we should always expect $E_n(GA)$ to offer much better-quality results than $E_n(3)$.

Other properties and $\mu$

It might seem quite logical to demand that

$$\Psi_i \sim \phi_i + \mu_i^{-1}(B) \Phi_i^{(1)}$$

when one applies the GA to the $B_i$-series, with $\mu_i(B)$ given by (2.52). Then, of course, $\langle W \rangle_i = \langle \Psi_i \mid W \mid \Psi_i \rangle / \langle \Psi_i \mid \Psi_i \rangle$ may be approximated accordingly. But, probably a more balanced result may be obtained by directly optimising the $\langle W \rangle_i$ - series to obtain
\[ \mu_i(W) = 1 - \frac{w_i(2)}{w_i(1)} , \quad \langle W \rangle_i \, (GA) = w_i + \mu_i^{-1}(W) \, w_i(1). \quad (2.65) \]

A similar remark applies to \( \langle Q \rangle_i \), leading to

\[ \mu_i(Q) = 1 - \frac{q_i(1)}{q_i(0)} , \quad \langle Q \rangle_i \, (GA) = \mu_i^{-1}(Q) \, q_i(0). \quad (2.66) \]

In Table 9, the usefulness of (2.66), among others, may be found.

**Transferability and uniqueness of \( \mu \)**

We may note that the GA adopted calculation for some \( \langle X \rangle_i \) involves two steps: (a) evaluation of \( \mu_i(X) \) from

\[ \mu_i(X) = 1 - \frac{x_i^{(m)}}{x_i^{(m-1)}} \quad (2.67) \]

and (b) use of

\[ \langle X \rangle_i \sim \langle X \rangle_i \, (GA) = \sum_{n=0}^{m-2} x_i^{(n)} + \mu_i^{-1}(X) \, x_i^{(m-1)} \quad (2.68) \]

where \( m = 1, 2 \) or 3 depending on whether \( X = Q, W \) or \( H \), respectively.

Because of computational intractability in higher orders with a large basis set, recently \(^{66,67} \), two simplifications have been advocated:

(i) \( \mu_i^A(X) = \mu_i^B(X) \)

(ii) \( \mu_i^C(X) = \mu_i^G(H) \)

where by \( \mu_i^A(X) \) we mean \( \mu_i(X) \) has been calculated by using (2.67) with \( x_i^{(m)} \)'s evaluated through a restricted basis set \( A \) (set \( A \) is smaller than set \( B \)). Some results are displayed in Table 9; here \( \omega_o^W(W) \) actually denotes \( (\mu_o^A(\alpha))^{-1} \omega_o^{(0)}(0) \) (in this case refers to computations using \( B \)) and \( \omega_o^W(\omega) \) means \( (\mu_o^G(\beta))^{-1} \omega_o^{(0)}(0) \) (in this case corresponds to the same basis set \( C \)).
Table 9. Polarisability of the $\text{H}_2$ molecule.

<table>
<thead>
<tr>
<th>Basis set</th>
<th>$\mu_o(\mathcal{E})^{A}=0.8299$</th>
<th>$\mu_o(\mathcal{E})^{B}=0.8372$</th>
<th>Exact*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td>$\perp$</td>
<td>$\perp$</td>
<td>$\perp$</td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.7212 0.4594</td>
<td>0.7209 0.4883</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.1704 0.1233</td>
<td>0.1704 0.1373</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.0339 0.0285</td>
<td>0.0336 0.0335</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td>0.94549 0.67834</td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(0)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(1)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
<tr>
<td>$\alpha_o^{(2)}$</td>
<td>0.9255 0.6112</td>
<td>0.9249 0.6590</td>
<td></td>
</tr>
</tbody>
</table>
To be precise, however, the simplification (i) implies that the property $\langle X \rangle_1$ is not much sensitive to the change $A \rightarrow B$ of basis; in other words, it states that the smaller basis $A$ is adequate for the property $\langle X \rangle_1$ concerned. Thus, it is better to use (i) as a test of whether a given basis set is adequate for some property. Essentially, as may be seen from our calculated values of $\mu_0(\alpha)$ in Table 9, basis $A$ is almost equivalent to $B$ for $\alpha_0(\parallel)$ but is not so for $\alpha_0(\perp)$ and that is why $\varpi_0(\perp)$ for the perpendicular component is not very promising. Simplification (ii) is, in fact, more severe; even the employment of a complete basis may not ensure its validity. The characteristic modulations of the two series depend actually on their sensitivity to the perturbation and it is quite likely that they would be different, unless $X$ and $Y$ happen to be a pair of properties of very similar type; otherwise, the workability of (ii) should be considered as just accidental, like the case with $\varpi_0^W(2)$ in Table 9 where the agreement follows from the closeness of $\mu_0(B)$ and $\mu_0(\alpha)$. We note also that $\mu_0(3A)$ is quite successful compared with other prescriptions.

Another problem is, if we demand vanishing of any other (than those mentioned) scaled term, e.g., any $\xi_i^1(k)(k > 3)$ = 0 to obtain $\mu_1(B)$, multivaluedness of $\mu$ causes a problem. For example, referring to Table 9, we evaluate $\mu_0(\alpha)$ from

$$\xi_0^{-3} \alpha_0(2) + 2 \mu_0^{-3}(\mu_0-1) \alpha_0(1) + \mu_0^{-3}(\mu_0-1)^2 \alpha_0(0) = 0$$

for the parallel component of polarisability employing the basis $A$ and find $\mu_0^A(\alpha)_1 = 0.8576$, $\mu_0^A(\alpha)_2 = 0.6698$. Using these values, we then compute $\varpi_0(1)$ to get $\varpi_0(1) = 0.9330 & 0.9106$, which are rather poor estimates. To be wise after the event, we should avoid any non-uniqueness of $\mu$. 
2.3.2. An alternative strategy:

In order to avoid the difficulty of nonuniqueness with the previous strategy, the following scheme has been advocated:

(i) construct energy-expressions invariant to scaling of \( H_0 \) (i.e. \( H_0 \to \mu H_0 + kI \), \( \{I_i(m)\} \)), which agree to some \( m \)th order of the series, (ii) evaluate \( \mu_{im} \) by requiring \( \bar{E}_i(2) = I_i(m) \). Let us note here that \( \mu_{im} \) will be unique if, and only if, \( I_i(m) \) is unique for a fixed \( m \) (especially for \( m > 3 \)). But, since this latter question has not been settled elsewhere, we feel obliged to pay attention to it (see, e.g., our work in ref. for a detailed account); such a scrutiny will also help us in estimating the status of the \( I_i(m) \)s.

By inspection, in fact, we have found that both \( I_i(4) \) and \( I_i'(4) \), given by

\[
I_i(4) = \epsilon_i + \epsilon_i^{(1)} + \frac{\epsilon_i^{(2)}}{1 - \frac{\epsilon_i^{(3)}}{\epsilon_i^{(2)}}} + \frac{\epsilon_i^{(4)} - \epsilon_i^{(3)}^{2}/\epsilon_i^{(2)}}{(1 - \frac{\epsilon_i^{(3)}}{\epsilon_i^{(2)}})^3},
\]

\[
I_i'(4) = \epsilon_i + \epsilon_i^{(1)} + \epsilon_i^{(2)}(1 - \frac{\epsilon_i^{(3)}}{\epsilon_i^{(2)}}) - \frac{\epsilon_i^{(4)} - \epsilon_i^{(3)}^{2}/\epsilon_i^{(2)}}{\epsilon_i^{(2)}}(1 - \frac{\epsilon_i^{(3)}}{\epsilon_i^{(2)}})^{-1},
\]

are fourth-order invariants. Moreover, the relative efficiency of \( I_i(4) \) and \( I_i'(4) \) is problem-dependent, as we have checked for the harmonic oscillator case perturbed by quadratic and quartic potentials. Also, we may emphasize that actually we can always construct a family of such invariants because of the flexibility present in the remainder \( O(m+1) \) where

\[
I_i(m) = \bar{E}_i(m) + O(m+1).
\]
Thus, we find

\[ \{ T_i^P(3) \} = \epsilon_i + \epsilon_i^{(1)} + \sum_{n=1}^{P} \frac{k_n \epsilon_i^{(2)}_n}{(1 - \epsilon_i^{(3)} / \epsilon_i^{(2)})_n} ; k_i = 1, k_n \neq k_n(\mu), \]

the usual third-order invariant corresponds to \( p = 1 \). Similarly, a family of fourth-order invariants is possible to construct. As a consequence, an infinite number of \( \{ \mu_{1m}^P \} \) are possible for fixed \( i \) and \( m \). So, we conclude that (i) it is difficult to expect that the invariant forms, \( I_i(m) \), would 'allow a sort of extrapolation between the given ordinary approximants and the series limit', (ii) for any fixed \( i \) and \( m \), \( I_i(m) \) is not a very significant quantity in PT due to nonuniqueness and hence approximate energy-calculations with any \( I_i(m) \) should not have a sound theoretical status, (iii) the alternative procedure for the evaluation of \( \mu_{1m}^P \) from the equality \( \bar{E}_i(2) = I_i(m) \) is far less satisfactory relative to the conventional strategy, for here nonuniqueness arises even for \( m=3 \), (iv) not only for the energy, but for other first- or second-order properties also, the above arguments and demonstrations would apply (only the terms would involve different orders, the expressions would remain similar in forms), and lastly, (v) considering the odd-order invariants \( \{ I_i^P(2m+1) \} \), there exists a **unique** one \( I_i^{PA}(2m+1)(= [m+1/m]) \) which is the \([m+1/m]\) Padé approximant to \( E_i \) and is known to be scale-invariant72 of \( H_o \). Probably, only these \( I_i^{PA} \) 's have special appeal, for a specific Padé approximant to a given series is **unique**73.
2.4. **Aspects Of Convergence Of BWPT**

Whereas the convergence properties of RSPT have been studied by several authors from time to time (see e.g. ref.2, to quote a few), the BWPT has attracted rather scanty attention in this respect, probably because of its more intricate character than the RS formulation. But, recent interests in the specific context of the 'intruder' problems ('crossing' in complex-\(\lambda\) plane) have urged studies in this particular field. So, we shall here consider this convergence question briefly, pointing out some subtle characteristics of BWPT and the salient features of finite-dimensional BW development. Here, however, we proceed, for convenience, in a different manner to develop the theory.

As usual, we have

\[ H = H_o + \lambda V, \quad \mathcal{D}(V) \supseteq \mathcal{D}(H_o) \]  

(2.69)

where \( H_o \phi_n = \varepsilon_n \phi_n \), and we define

\[ P = |\phi_n\rangle \langle \phi_n|, \quad Q = I - P. \]  

(2.70)

Rearranging (2.2) for the \(n\)th state,

\[ (H_o + \lambda V) \psi_n(\lambda) = E_n(\lambda) \psi_n(\lambda), \]  

(2.71)

in the form

\[ (E_n - H_o) \psi_n = \lambda V \psi_n \]  

(2.72)

and writing

\[ \psi_n = \phi_n + \kappa_n, \quad \| \phi_n \| = 1, \langle \phi_n | \kappa_n \rangle = 0 \quad (\text{intermediate normalization}), \]  

(2.73)

we find from (2.72)

\[ (E_n - H_o) \phi_n = \lambda V \psi_n \]  

(2.74)
\[(B_n - H_0) \chi_n = \lambda Q \psi_n \quad (2.75)\]

respectively for the P- and Q-projection parts. From (2.74), it then follows that

\[
E_n = \langle \phi_n | H_0 | \phi_n \rangle + \lambda \langle \phi_n | V \psi_n \rangle \quad (2.76)
\]

which may be used to evaluate \(E_n\) if \(\psi_n\) is known.

To achieve this end, we assume \(E_n(\lambda) \neq \epsilon_i\) \((i = 0, 1, 2, \ldots, n-1, n+1, \ldots)\), define the reduced resolvent \(T_n = Q(E_n - H_0)^{-1}Q\) of \(H_0\) and rearrange (2.75) to obtain

\[
\chi_n = \lambda T_n \psi_n = \lambda T_n V \phi_n + \lambda T_n V \psi_n, \quad \overline{V} = Q V Q, \quad (2.77)
\]

which gives, using (2.73),

\[
\psi_n = (I + \lambda T_n V) \phi_n + \lambda T_n \overline{V} \psi_n
\]

\[
= \theta_n + \lambda T_n \overline{V} \psi_n. \quad (2.78)
\]

This equation is the key equation of a BW development, and it has to be solved iteratively, leading to

\[
\psi_n = (I + \lambda T_n \overline{V} + \lambda^2 T_n \overline{V} T_n \overline{V} + \ldots) \theta_n. \quad (2.79)
\]

Obviously, this iteration process would converge only if the condition

\[
||T_n \overline{V}|| < \frac{1}{|\lambda|} \quad (2.80)
\]

is satisfied, and then one can write

\[
\psi_n = (1 - \lambda T_n \overline{V})^{-1} \theta_n, \quad (2.81)
\]

which, in turn, can be employed in (2.76) to compute \(E_n\). Thus, crucial in the issue of convergence of BWPT is the satisfaction of (2.80). At this point, we note the following:
(a) Since $E_n = E_n(\lambda)$, $T_n = T_n(\lambda)$ and hence the LHS of (2.80) is $\lambda$-dependent. So, we cannot have here a criterion of convergence in the usual sense, i.e., a radius $\lambda_0$ (the RC) which would dictate for any $|\lambda| < \lambda_0$ absolute convergence and for $|\lambda| > \lambda_0$ divergence. Instead, we see from (2.80) that what is important here is the 'point convergence' - convergence at some particular value of $\lambda$ (by checking whether (2.80) is obeyed).

(b) It is natural, in respect of what we have remarked just above, to conclude that one has to know the exact $E_n(\lambda)$ beforehand to ensure convergence of BWPT, and hence the criterion (2.80) becomes useless. This is because, if $E_n(\lambda_c)$ for some particular value of $\lambda = \lambda_c$ is to be known exactly beforehand there is no point of talking about a perturbative development! However, we shall see that, at least for the ground state, it is possible to ensure the validity of (2.80) without requiring the exact knowledge of $E_n(\lambda_c)$.

(c) The usual practice (see the last two works of ref.74) to bypass the situation mentioned above is to consider $E_n(\lambda)$ as some fixed number 'f' and then to evaluate a RC $\lambda_0^x$ which limits the success of the BW iteration scheme as $\lambda$ is varied (from $\lambda = 0$). We may emphasize here that, except for the point $\lambda_f$ that satisfies $E_n(\lambda_f) = f$ (and hopefully $|\lambda_f| < \lambda_0^x$), a definite remainder will always be left implying that one is not dealing with a true BW development.

(d) Although we remarked in (a) that in BWPT one has to talk of point convergence and not a RC, for finite-dimensional problems
we shall see that it is possible to evaluate something akin to the concept of a RC.

The finite-dimensional case

For a finite-dimensional case, \( \| V \| \) is finite and since
\[
\| T_n V \| \leq \| T_n \| \| V \| ,
\]
(2.82)
it follows that the requirement (2.80) can be replaced by the sufficiency condition
\[
|\lambda| < \frac{1}{\| T_n \| \| V \|}
\]
(2.83)
which gives
\[
|\lambda| < \frac{|E_n(\lambda) - \varepsilon_x|}{\| V \|} = F_n(\lambda),
\]
(2.84)
\( (\varepsilon_x = \varepsilon_{n+1}, \text{nearest to } E_n(\lambda)).\)

The HHS of (2.84), considered as some function of \( \lambda \), shows that at \( \lambda = 0 \), it has a finite positive value. This means, for a finite-dimensional problem, the BW expansion for any \( n \)th state must converge around a finite region of the \( \lambda = 0 \) point. Thus, here we can talk about a RC which conveys a meaning somewhat different from the usual notion. For example, if for any positive \( \lambda \) within \( 0 \leq \lambda \leq \lambda_n^+ \), the inequality \( |\lambda| \langle F_n(\lambda) \rangle \) is satisfied and similarly for any negative \( \lambda \) satisfying \( 0 \leq \lambda \leq \lambda_n^- \), \( |\lambda| < F_n(\lambda) \) holds, then \( \min \{ \lambda_n^+, \lambda_n^- \} \) would give us a radius of definite convergence, a lower bound to the RC. But such a RC does not mean that for any \( +\lambda \text{ or } -\lambda \) having numerical value larger than this radius, the BWPT must diverge. Another point is (as (2.49) shows), to be within such a RC, a necessary condition to be obeyed is the 'rule of thumb' \(^6\) (usually stated in
the context of RSPT) — the energy-shift must be numerically less than the unperturbed energy-gap (considering the nearest neighbour).

We shall now show that, for the ground state, the BWPT can always be made convergent. To proceed, we note following Ahlrichs

\[ \| T_n V \| = \| V T_n \| \quad (2.85) \]

(such a conclusion follows also from the requirement of convergence of \( \langle \Psi_n \rangle \) from (2.78) — the iteration equation) which implies that (2.80) can be written as

\[ T_0 V^2 T_0 < \frac{1}{|\lambda_0|^2} = \frac{1}{\lambda_c^2} \quad (\lambda_c \text{ real}) \quad (2.86) \]

(recall the definition of norm) for some particular value of \( \lambda = \lambda_c \). But, this is equivalent to the requirement

\[ \langle s | V^2 | s \rangle < \frac{1}{\lambda_c^2} \langle s | E_0(\lambda_c) - H_0 \rangle^2 | s \rangle, s \in Q \mathcal{D}(H_0). \]

... (2.87)

Let us accept that this inequality is not obeyed, in general, for the parent partition \( H = H_0 + \lambda_c V \). So, we consider a repartition:

\[ H = h_0 + \lambda_c v, \quad h_0 = H_0 + k \lambda_c, \quad v = V - k. \quad (2.88) \]

Then, we would have the condition

\[ \langle s | V^2 | s \rangle < \frac{1}{\lambda_c^2} \langle s | (E_0 - h_0)^2 | s \rangle. \quad (2.89) \]

Using (2.88), this becomes

\[ \langle s | V^2 | s \rangle - \frac{1}{\lambda_c^2} \langle s | (E_0 - H_0)^2 | s \rangle \langle \frac{2k}{\lambda_c} \langle s | H - E_0 | s \rangle. \]

... (2.90)
Note that \((H-E_0)\) is positive definite in the domain \(QD(H_0)\) concerned. Now, in this domain, functions \(\{h_i\}\) which obey
\[
\langle h_i | V^2 | h_i \rangle \leq \frac{1}{\lambda_c^2} \langle h_i | (E_0 - H_0)^2 | h_i \rangle
\]  
(2.91)

would automatically obey (2.90) if we agree to (and we do) choose \(k\) of the same sign as that of \(\lambda_c\) and we need not worry for such functions. But, there would also be functions \(\{\phi_i\}\) which do not obey an equation of the form (2.91). For these functions \(\phi_i\), which lead to
\[
\langle \phi_i | V^2 | \phi_i \rangle \geq \frac{1}{\lambda_c^2} \langle \phi_i | (E_0 - H_0)^2 | \phi_i \rangle
\]  
(2.92)

we choose

(i) for positive \(\lambda_c\),
\[
k > \frac{\lambda_c}{2} \max_{\phi_i} \frac{\langle \phi_i | V^2 - \frac{1}{\lambda_c^2} (E_0 - H_0)^2 | \phi_i \rangle}{\langle \phi_i | H - E_0 | \phi_i \rangle}
\]  
(2.93)
o
(ii) for negative \(\lambda_c\),
\[
k < \frac{\lambda_c}{2} \max_{\phi_i} \frac{\langle \phi_i | V^2 - \frac{1}{\lambda_c^2} (E_0 - H_0)^2 | \phi_i \rangle}{\langle \phi_i | H - E_0 | \phi_i \rangle}
\]  
(2.94)

to satisfy (2.90). Thus, we have seen that \(k\) can be suitably chosen to make BW expansions in terms of the partitioning (2.88) always convergent for the ground state. It may be mentioned that Leinaas-Kuo also arrived at the same result, but proceeding through a different and more complicated route.

2.5. Variants of BWPT And Ground-State Energy-Bounds:

Here, by 'variants' we do not wish to imply a discussion on various kinds of partitions and their usefulness in a BW
context, rather we shall be concerned with the various possible ways of viewing the exact energy appearing in the BW expansions, for a clarification of this particular feature seems of interest from a practical point of view. Indeed, we analyse the Morse-Feshbach formulation\textsuperscript{79} of BWPT in detail. Special attention is paid to the odd-order energy-sums which provide upper bounds to the exact ground-state energy.

Possible schemes

Here, for convenience, we follow the straightforward\textsuperscript{44} derivation of BWPT by writing the TISE \( H(\lambda) \psi_i(\lambda) = E_i(\lambda) \psi_i(\lambda) \) in the form

\[
(H_0 + \lambda V)(\phi_i + \sum_{k=1}^{\infty} \lambda^k \phi^{(k)}_i) = (E_i + \sum_{k=1}^{\infty} \lambda^k E^{(k)}_i) \psi_i.
\]

(2.95)

Let us now note two important features of this kind of decomposition:

a) Although we wish to express \( E_i(\lambda) \) and \( \psi_i(\lambda) \) as

\[
E_i(\lambda) = \epsilon_i + \sum_{k=1}^{\infty} \lambda^k \epsilon_i^{(k)}, \quad \psi_i(\lambda) = \phi_i + \sum_{k=1}^{\infty} \lambda^k \phi_i^{(k)},
\]

(2.96)

as is implied in (2.95), the set of BW perturbation equations

\[
\begin{align*}
H_0 \phi_i^{(1)} + V \phi_i = & \quad \epsilon_i^{(1)} \phi_i + E_i \phi_i^{(1)} \\
H_0 \phi_i^{(2)} + V \phi_i^{(1)} = & \quad \epsilon_i^{(2)} \phi_i + E_i \phi_i^{(2)} \\
& \quad \cdots \\
H_0 \phi_i^{(r)} + V \phi_i^{(r-1)} = & \quad \epsilon_i^{(r)} \phi_i + E_i \phi_i^{(r)}
\end{align*}
\]

(2.97)
contains the unknown $E_i(\lambda)$ and hence the expansions in (2.96) are not true power-series ones; we should better write (2.96) as

$$
E_i(\lambda) = \epsilon_i + \lambda \epsilon_i^{(1)} + \sum_{k=2}^{\infty} \lambda^k \epsilon_i^{(k)}(E_i) \\
\psi_i(\lambda) = \phi_i + \sum_{k=1}^{\infty} \lambda^k \phi_i^{(k)}(E_i)
$$

(2.98)

showing explicitly the functional dependences of correction-coefficients.

b) It is also difficult to appreciate the meaning of terms like $\{\epsilon_i^{(m)}(m \gg 2)\}$ or $\{\phi_i^{(m)}(m \gg 1)\}$ unless $E_i(\lambda)$ is exactly known beforehand (in which case probably there is no point of talking about PT!). From a practical point of view, indeed, a value ascribed to some BW correction term, say $\epsilon_i^{(m)}(m \gg 2)$ for the energy, must refer to the level of approximation involved in $E_i(\lambda)$ contained in it.

Of course, there are other ways of deriving BWPT$^{78,79}$, but these two above-mentioned features are very characteristic of this formalism and owe their origin, in our approach, from the fact that we have deliberately suppressed the $\lambda$-dependence of $E_i(\lambda)$ occurring at the second part of the RHS of (2.95). In fact, this very strategy distinguishes BWPT from the more familiar RS one. However, in spite of these two tickling features, not usually emphasised (but see ref.80), and a number of other practical difficulties$^{44}$, the BW formalism has drawn some attention$^{77,81}$.

In particular, since the BW energy series

$$
E_i(\lambda) = \epsilon_i + \lambda \epsilon_i^{(1)} + \lambda^2 \sum_j \frac{V_{ij} V_{ik}}{E_i - \epsilon_j} + \lambda^3 \sum_{j,k} \frac{V_{ij} V_{ik} V_{ik}}{(E_i - \epsilon_j)(E_i - \epsilon_k)} + \cdots
$$
with \( V_{ij} = \langle \phi_i | V | \phi_j \rangle \), on truncation at some \( m \)th order, along with an expansion of \( B_i \) appearing in the denominators of the correction terms to the same order, gives us the RS series correctly to \( m \)th order plus a definite remainder containing higher-order terms, such a series in argued to have better convergence properties \(^{21,79}\) for one hopes to have extracted some higher-order RS effects into lower orders of a BW scheme through the remainder term (but see ref.\(^{85}\)). Clearly, such a conclusion is based more on intuition than on a rigorous theoretical feeling as one cannot be sure \textit{a priori} whether the very remainder contains important higher-order RS effects which are not cancelled \(^{51}\). A second advantage of the BW theory is its straightforward applicability to cases associated with degeneracy. Finally, it seems also interesting to apply the BW theory to the ground state \((i=0)\) at least, where it is known that the theory offers upper bounds \(^{51,77,82}\) to \( E_o(\lambda) \) if the \( m \)th partial sum

\[
E_o(m) = E_0 + \lambda \epsilon_0^{(1)} + \sum_{k=2}^{m} \lambda^k \epsilon_0^{(k)} (E_o) \tag{2.99}
\]

is computed in a self-consistent fashion to some odd order \((m = 2n + 1; n = 0, 1, \ldots)\). In fact, such a self-consistency scheme really amounts to writing (2.99) in a modified form:

\[
E_o(m) = E_0 + \lambda \epsilon_0^{(1)} + \sum_{k=2}^{m} \lambda^k \epsilon_0^{(k)} (E_o(m)). \tag{2.100}
\]

However, we like to emphasize at this point that this self-consistency scheme (to be called hereafter the scheme \((i)\)) is not the only possible procedure that one \textit{has} to adopt in order to perform a BW calculation.
Essentially, there may be four possible ways of calculating the BW partial energy-sum, given by (2.99), depending on the value of $E_o$ to be chosen in order to calculate $\{\epsilon_o^{(k)}\}$:

a) to adopt the scheme (i), i.e., to use an equation of the form (2.100),

b) to consider an equation of the form (2.99) where one uses the exact $E_o$, to be known beforehand, to compute the correction terms like $\epsilon_o^{(k)}$; adoption of such a strategy is only of pedagogical interest but may be useful in studying the true convergence behaviour of the BW energy-series and this will be referred to as scheme (ii),

c) to follow the Morse-Feshbach (MF) approach, the scheme (iii), where the $m$th partial energy-sum is calculated using the equation

$$
E_o(m) = \epsilon_o + \lambda \epsilon_o^{(1)} + \sum_{k=2}^{m} \lambda^k \epsilon_o^{(k)}(E_o(m-k))
$$

(2.101)

in order to keep track of the ordering parameter $\lambda$,

d) to substitute some approximate value of energy, say $E_o^a$, for $E_o$, for the calculation of $\epsilon_o^{(k)}$ to arrive at the equation (scheme (iv))

$$
E_o(m) = \epsilon_o + \lambda \epsilon_o^{(1)} + \sum_{k=2}^{m} \lambda^k \epsilon_o^{(k)}(E_o^a)
$$

(2.102)

We shall denote the energy-values obtained through these different schemes respectively by $E_o^{BW}(m)$, $E_o(m)$, $E_o^{MF}(m)$ and $E_o^a(m)$. It is clear that either of these schemes may be followed for the calculations of excited states as well, i.e., for any $E_i(m)$, but here we shall be concerned mainly with the ground state.
The upper-bound property

The proof that the odd-order sums of the BW energy-series provide upper bounds to the exact ground-state energy proceeds as follows\(^{51,52}\). We choose the trial function

\[
\Psi_o = \phi_o + \sum_{k=1}^{n} \lambda^k \phi_o^{(k)} = \Psi_o(n)
\]

and construct the expectation-value equation to obtain

\[
\mathcal{E}_o = \frac{\langle \Psi_o | H | \Psi_o \rangle}{\langle \Psi_o | \Psi_o \rangle} = \frac{\mathcal{E}_o(2n+1) + \mathcal{E}_o(S_n-1)}{S_n} \geq \mathcal{E}_o,
\]

where \(S_n = \langle \Psi_o | \Psi_o \rangle\) and the last inequality follows from the Rayleigh-Ritz principle. Since usually one follows the scheme (i) to obtain \(\mathcal{E}_o(2n+1)\), which is actually referring to \(\mathcal{E}_o^{BW}(2n+1)\), it is then justifiable on this very ground to consider the \(\mathcal{E}_o\) appearing in \(\{\phi_o^{(k)}\}\) and \(\{\epsilon_o^{(k)}\}\), and hence also the one at the second part of the numerator in (2.104), as referring to \(\mathcal{E}_o^{BW}(2n+1)\) so that the inequality (2.104) now reads

\[
\mathcal{E}_o^{BW} = \mathcal{E}_o^{BW}(2n+1) \geq \mathcal{E}_o,
\]

which is to be proved.

At this stage, we note that (a) if one sticks to the scheme (ii), it immediately follows from (2.104) that then the relation

\[
\mathcal{E}_o \neq \mathcal{E}_o(2n+1) \geq \mathcal{E}_o
\]

is obeyed which implies that the upper-bound nature of the odd-order sums of the BW energy-series will hold also when one adopts the scheme (ii) and
(b) if one substitutes for the $E_o$ in $\{\phi_o^{(k)}\}$ and $\{\epsilon_o^{(k)}\}$ some $E_o^{a}$, following the attitude (iv), the inequality (2.104) will then read

$$\bar{E}_o^{a} = E_o^{a}(2n+1) + E_o^{a}(S_n^n - 1) \geq E_o.$$  \hspace{1cm} (2.107)

Similarly, when the trial function is chosen, instead of (2.103), in the form

$$\bar{\psi}_o (\mu) = \phi_o + \sum_{k=1}^{E} \lambda_j^k \phi_o^{(k)}$$  \hspace{1cm} (2.108)

and is thus rendered more flexible to improve the upper bound, one obtains

$$\bar{E}_o (\mu) = \frac{\langle \bar{\psi}_o (\mu) | H | \bar{\psi}_o (\mu) \rangle}{\langle \bar{\psi}_o (\mu) | \bar{\psi}_o (\mu) \rangle} = \frac{E_o (\mu) + \eta_0 (S_j^n - 1)}{S_j^n} \geq E_o,$$

$$\ldots$$  \hspace{1cm} (2.109)

with $S_j^n = \langle \bar{\psi}_o (\mu) | \bar{\psi}_o (\mu) \rangle$ and

$$E_o (\mu) = \epsilon_o + \lambda \epsilon_o^{(1)} + 2 \sum_{i=1}^{n} \lambda_i^{i+1} \epsilon_o^{(i+1)}$$

$$+ \sum_{i,j=1}^{n} \lambda_i^{i+j} \epsilon_o^{(i+j+1) - \epsilon_o^{(i+j)}}.$$  \hspace{1cm} (2.110)

Here again $E_o (\mu)$ turns out to be an upper bound to $E_o$ if we agree to identify the $E_o$ appearing in $\{\epsilon_o^{(k)}\}$ and $\{\phi_o^{(k)}\}$ as $E_o (\mu)$, following thus again the scheme (i), for we then obtain from (2.109)

$$\bar{E}_o^{BW} (\mu) = E_o^{BW} (\mu) \geq E_o.$$  \hspace{1cm} (2.111)

As above, if we instead follow the scheme (iv), the inequality (2.109) takes the form
From (2.112), one quickly finds that if the expressions concerned are constructed using the exact $E_o$, in keeping with the notion of scheme (ii), the relation becomes

$$\overline{E}_o(\mu) \neq E_o(\mu) \geq E_o,$$  \hspace{1cm} (2.113)

as may be checked by putting $E_o$ in place of $E^a_o$ in (2.112). The inequality (2.113) is of a form similar to (2.106) and again shows that the upper-bound property of $E_o(\mu)$ defined by (2.110) will also hold for any set of $\{\mu_i\}$ if one follows the strategy (ii). Further exploitations of these inequalities would concern us a bit later.

**The MF procedure**

Although in connection with the BW theory the MF approach is usually quoted, we remarked before that this procedure is in some sense different and corresponds to scheme (iii), a rather special one. Thus the $n$th partial sum of the wavefunction series for some $i$th state here reads

$$\psi_i(n) = \phi_i + \lambda \phi_i^{(1)}(E_i(n-1)) + \lambda^2 \phi_i^{(2)}(E_i(n-2)) + \ldots + \lambda^n \phi_i^{(n)}(\epsilon_i),$$ \hspace{1cm} (2.114)

and hence, if we consider also the MF energy-series of the form (2.101) for some $i$th state, the set of perturbation equations takes the form

$$H_0 \phi_i^{(1)} + V \phi_i = \epsilon_i^{(1)} \phi_i + E_i(m-1) \phi_i^{(1)}$$
$$H_0 \phi_i^{(2)} + V \phi_i^{(1)}(E_i(m-2)) = \epsilon_i^{(2)} \phi_i + E_i(m-2) \phi_i^{(2)}$$
$$\ldots \ldots \ldots \ldots$$
$$H_0 \phi_i^{(m)} + V \phi_i^{(m-1)}(\epsilon_i) = \epsilon_i^{(m)} \phi_i + \epsilon_i \phi_i^{(m)}.$$ \hspace{1cm} (2.115)
when we wish to calculate the energy-series to \( m \)th order. But, if we follow the conventional scheme, these equations become

\[
\begin{align*}
H_0 \phi_i^{(1)} + V \phi_i^{(1)} &= \epsilon_i^{(1)} \phi_i^{(1)} + E_i^{(m)} \phi_i^{(1)} \\
H_0 \phi_i^{(2)} + V \phi_i^{(1)} &= \epsilon_i^{(2)} \phi_i^{(1)} + E_i^{(m)} \phi_i^{(2)} \\
& \vdots \\
H_0 \phi_i^{(m)} + V \phi_i^{(m-1)} &= \epsilon_i^{(m)} \phi_i^{(1)} + E_i^{(m)} \phi_i^{(m)}; \\
\end{align*}
\]

the corresponding perturbation equations for the schemes (ii) and (iv) follow respectively by putting \( E_i^a \) and \( E_i^a \) in place of \( E_i^{(m)} \) appearing in (2.116).

It is now clear, looking at the set of perturbation equations, that the **MF procedure is very different from the other three schemes**, for in this scheme one uses different approximate values of \( E_i \) occuring at different orders of the perturbation equations as also different expressions for \( \phi_i^{(k)} \) appearing at the \( k \)th and \( (k+1) \)th order equations. Thus, for example, in the MF procedure one can not construct energy-correction terms to \( (2n+1) \)th order from a knowledge of the wavefunction-series to \( n \)th order, with a hermitian \( V \). This particular advantage continues to exist in the BW framework only when one uses the same value of \( E_i \) throughout, as in (2.116), whatever may be the level of approximation involved. Clearly, then, from a practical point of view, the MF scheme (iii) is much less advantageous compared to either of the other three. Also, as the proof of the upper-boundedness goes, it is apparent that this property will not hold in the MF formulation. Finally, from (2.101), it immediately follows that to second order, the MF procedure differs in no way from the standard RS formulation.
The above difficulties, however, may be bypassed if one uses, instead of (2.101), the equation

$$B_i(m) = \varepsilon_i + \lambda \varepsilon_i^{(1)} + \sum_{k=2}^{m} \lambda^k \varepsilon_i^{(k)}(B_i(m-1)), \quad (2.117)$$

as was advocated later in ref. 79 to be more convenient for practical purposes (note: it may be termed as the convenient MF (CMF) approach). We shall denote such an expression for energy by $B_i^{\text{CMF}}(m)$. Since now one uses the same approximation $B_i^{\text{CMF}}(m-1)$ to $B_i$ in all the perturbation equations to $m$th order, it is easy to see that the above-mentioned advantage is regained and, going to second order only, one can also distinguish it from the RS theory. However, the upper-bound property, if any, is yet to be established. Noting that this CMF approach is essentially equivalent to scheme (iv), with $B_i^{\text{CMF}}(m-1)$ replacing $B_i^a$, it becomes clear from (2.107) that the bound property does not exist in this case, in general. These subtle aspects of the MF procedure have not been discussed previously and that is why misunderstandings with the MF approach cropped up. Thus, in ref. 21 it was remarked that the upper-bound property of the odd-order sums of the BW series for $B_i$ escaped notice of Morse and Feshbach. Furthermore, a disagreement concerning the upper-boundedness for the Mathieu problem (third-order) led the author of ref. 21 to suspect the correctness of the numerical data presented in ref. 79. Actually, Table 9.1 in ref. 79 was constructed on the basis of (2.117) and not following scheme (i) or (ii), so that the upper-bound property does not definitely follow. Unfortunately, however, the table concerned is slightly in error. We, hence, present the corrected data in our Table 10 in the notation of ref. 79 to avoid any further confusion on this point. We emphasize that although in this special case the third-order sums lie above
Table 10. The corrected Morse-Feshbach Table.

<table>
<thead>
<tr>
<th>s</th>
<th>( b_0^{(0)} )</th>
<th>( b_0^{(1)} )</th>
<th>( b_0^{(2)} )</th>
<th>( b_0^{(3)} )</th>
<th>( b_0^{(4)} )</th>
<th>Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0</td>
<td>0.10000</td>
<td>0.09872</td>
<td>0.09875</td>
<td></td>
<td>0.09875</td>
</tr>
<tr>
<td>1.0</td>
<td>0</td>
<td>0.50000</td>
<td>0.46429</td>
<td>0.46965</td>
<td></td>
<td>0.46896</td>
</tr>
<tr>
<td>2.0</td>
<td>0</td>
<td>1.00000</td>
<td>0.83333</td>
<td>0.89197</td>
<td>0.87338</td>
<td>0.87823</td>
</tr>
<tr>
<td>4.0</td>
<td>0</td>
<td>2.00000</td>
<td>1.00000</td>
<td>1.77778</td>
<td>1.15252</td>
<td>1.54486</td>
</tr>
</tbody>
</table>

* See Table 9.1 of ref. 79.
the exact ground-state energy, this does not hold universally and
a closer scrutiny in this respect will be found in what follows.

**Bracketing, ARSPT and the associated inequalities**

It follows from (2.107) that

\[ E_{a}(2n+1) > \left( E_{o} - E_{a} \right) S_{n}^{a} + E_{o} . \]  

(2.118)

So, if we choose \( E_{o} \leq E_{a} \), subject to the requirement that \( S_{n}^{a} \geq 1 \), we quickly find that the following inequality

\[ E_{a}(2n+1) > E_{o} \]  

(2.119)

is satisfied, i.e., \( E_{o}(2n+1) \) turns out to be again an upper bound to \( E_{o} \). The case \( n=1 \) is of special importance in this context, for then the condition \( S_{n}^{a} \geq 1 \) is always satisfied. Thus, we may remark that especially the third-order HW energy will always lie above the exact ground-state energy if \( E_{o} \) appearing in the \( |\epsilon_{o}^{(k)}| \) is replaced by any approximate \( E_{a} \), provided that \( E_{a} \leq E_{o} \). This condition is sufficient, but not necessary, as is evident from (2.118). One may, hence, obtain a good upper bound to \( E_{o} \) if some lower bound to the same is available, thus bracketing the lowest eigenvalue. It may be mentioned that this bracketing scheme is more economic than the one of Lowdin\(^{74}\) which requires knowledge of the inverse of the matrix \( QHQ \) where \( Q \) is the projector onto the space orthogonal to \( \phi_{o} \). Also, one may be provided with a still tighter upper bound by exploiting the increased flexibility of the trial function like (2.108). Thus, from (2.112) we find

\[ E_{o}^{a}(\mu) > E_{o} \]  

(2.120)

if we choose \( E_{o}^{a} \leq E_{o} \) and also that \( S_{\mu}^{a} \geq 1 \), which is obviously
again satisfied for \( n = 1 \). Variation of the parameters \( \mu \) in 
\( E^a_0(\mu) \) would understandably then lead to a much tighter upper bound. Focussing our attention on the Mathieu problem and the Table 10, it may now seem convincing to one why the NF approach (actually CMP) provides upper bounds at the third order, i.e., why \( b_0^{(3)} \geq b_0 \) (note that \( b_0^{(2)} < b_0 \)). It may be checked that the same lower bound \( b_0^{(2)} \), if employed in (2.120), with \( \mu_2 = \mu_3 = \ldots = 0 \), and the resulting expression for \( E_0^a(\mu) \) is optimised, one obtains as an upper bound to \( E_0 \), noting (2.120), the estimate 
\( E_0^a(\mu^{opt.}) = 1.60000 \), a remarkably tighter bound compared to \( b_0^{(3)} \), for \( s = 4.0 \).

Thus far, we were concerned with the bracketing of \( E_0 \) and for such a purpose we required knowledge of some lower bound to \( E_0 \). However, more often, one does not have such an information at hand, primarily because lower bounds are, in general, difficult to obtain theoretically. So, now we shall discuss a problem of wider interest. This problem is concerned with the extent of accuracy of some \( E_0^a \), obtained either theoretically or experimentally. For this purpose, we shall exploit the inequalities obtained to compute bounds to the error \( \Delta_a \) involved in \( E_0^a \), where 
\( \Delta_a = (E_0 - E_0^a) \).

From (2.118), it follows that 
\[ \Delta_a \leq \frac{E_0^a(2n+1) - E_0^a}{S_n^a} \]  
(2.121)

Although this bound is one-sided, a 'good' bound may well serve our purpose and it may be achieved by varying \( n \). Moreover, from (2.112) we obtain a more flexible, and hence better, form of inequality that reads
\[ \Delta_a \leq \frac{E_a^a(\mu) - E_o^a}{s_a^a} \]  \hspace{1cm} (2.122)

where both \( n \) and the set \( \{ \mu_i \} \) can be varied. It seems of value to investigate how far these bounds are useful for practical purposes. For such a calculational demonstration, we choose a simple test problem given by the Hamiltonian

\[ H(\lambda) = (-\nabla^2 + x^2) + \lambda x^2 \]  \hspace{1cm} (2.123)

with the bracketted part denoting the unperturbed problem. The ground-state RS energy-series for this problem corresponds to the function \((1+\lambda)^{1/2}\) which diverges if \( \lambda > 1 \). In such situations, it is the usual practice to sum the \( E_o(\lambda) \) series (RS) up to the numerically smallest term (the by now familiar asymptotic sum) neglecting the remainder whose partial sums behave obstinately. Thus, at \( \lambda = 2 \), one obtains the value 2.0000 as the asymptotic estimate. Considering this value as \( E_o^a \), we find from (2.122) with \( n = 1 \), by varying \( \mu_1 \), a good estimate of the remainder \( \Delta_a \) which amounts to \( \Delta_a \leq -0.2403 \) at \( \mu_1 = 0.4 \). This value of \( \Delta_a \) serves two purposes; first, it convincingly demonstrates that although the partial sums of a divergent series behave uncontrollably, the total sum may well be finite; second, from the relation

\[ \Delta_a = E_o - E_o^a \]

one finds \( E_o \leq 1.7597 \) and hence the next calculation of the same sort may be performed using \( E_o^a = 1.7597 \). Still better results would be obtained if one varies \( n \) in (2.108) while constructing the bounds from (2.122).

It may be mentioned that although \( E_o^a \) may be any good estimate of \( E_o \) obtained by any scheme whatsoever, we have considered it as a RS perturbative estimate (truncated) corresponding to some divergent series representation for a special purpose, to
investigate how far an asymptotic sum is useful. This may at once by seen by computing $A_n$. Keeping it in mind that currently the use of several summation techniques (Ch. 3) have been shown to be quite promising in handling divergent series, one may also employ the inequalities we obtained to estimate the relative efficiency of various such schemes. We may add that crucial in this issue of estimating a finite total sum is the nature of the perturbed Hamiltonian and not that how dangerously the series representation corresponding to $B_o$ diverges. Thus, for the $H(\lambda)$ in (2.123), the RS energy-series

$$B_o(\lambda) = 1 + \lambda/2 - \lambda^2/8 + \lambda^3/16 - \ldots \quad (2.124)$$

diverges right from beginning at $\lambda = 5$, but we find, from (2.122), $A_n < -0.8421$ at $\lambda_1 = 0.1$ for $B_o^{a} = 3.5$, the first-order sum (upper bound). This implies $B_o < 2.6579$. But the applicability of (2.104) or (2.109), and hence of inequalities, requires that $H(\lambda)$ must be bounded from below. Thus, one may conclude that the series (2.124) represents some finite value, and this value, one might hope, may be somehow approximated; this is important and desirable too.
REFERENCES

Perturbation Theory for Linear Operators, 1966 (Springer: Berlin); N. Reed and B. Simon, Methods of Modern Mathematical
Physics, IV: Analysis of Operators, 1978 (Academic: New York);
etc. and refs. therein.

2. Kato\(^1\), G. Speisman, Phys. Rev. 107, 1180, 1957; W.A. Bingel,
Int. J. Quantum Chem. 2, 101, 1968; and R. Ahlrichs, ibid,
Int. J. Quantum Chem. Whitton, 6, 1173, 1972; etc.

3. Kato\(^1\)'s book, pp.410-413; for later numerical and analytical
work on Z\(^-1\)PT, see e.g. A. Forman and G.G. Hall, J. Mol. Spectr.
B. Brandas and O. Goscinski, Int. J. Quantum Chem. 4, 571,
1970; ibid, 68, 59, 1972; R. Ahlrichs, Phys. Rev. A5, 605,
1972 (with correction of an error in the work of Forman-Hall);
etc.


Rev. A1, 552, 1970; see also other refs. in Silverman\(^6\).

F.H. Stillinger and D.K. Stillinger, Phys. Rev. A10, 1109,
1974; F.H. Stillinger and T.A. Weber, ibid, A10, 1122, 1974;
D.R. Herrick and F.H. Stillinger, J. Chem. Phys. 62, 4360,
1975; J.N. Silverman and J.C. van Leuven, ibid, 67, 2955, 1977;

7. Silverman.


9. B. Simon, Ann. Phys. (NY) 58, 76, 1970; Int. J. Quantum Chem. 21, 3, 1982; Reed-Simon; J. Killingbeck, Rept. Progr. Phys. 40, 963, 1977 (and earlier refs. therein); A.V. Turbiner, Preprint INT-F-117, 1979; etc. In simpler terms, the SCA implies that the perturbing potential should not be more singular than the unperturbed potential part so that the physics with $+\lambda V$ and $-\lambda V$ would not drastically differ.


16. R. Ahlrichs and F. Claverie, Int. J. Quantum Chem. 6, 1001, 1972 (for the delta-function model of $H_2^+$, they found that over a region of internuclear separation $R < R_C$).

17. Ref.13; Killingbeck; Silverman-van Leuven.

19. Reed-Simon, p.60; Killingbeck; etc.


21. See, e.g., Killingbeck.

22. It might be argued that one could have also chosen the 'intermediate normalization', \( \langle \psi^* | \psi \rangle = 1 \). The point is, one can equivalently choose either \( \langle \psi^* | \psi \rangle = 1 \) or \( \langle \psi | \psi \rangle = 1 \); here this equivalence is lost.

23. K. Bhattacharya, Chem. Phys. Letts. 80, 257, 1981 (the present discussion follows closely this work).


29. In this respect, the 'no-wavefunction' (HVT) approach to RSPT is of course a disadvantage (see Sec.1.2.1(c)).

30. For \( E_o^4(\lambda) \), the RSPT answer is wrong and why it is so we have discussed. A recent observation (R.P. Saxena and V.S. Verma, J. Phys. A : 15, L149, 1982) is also of interest in this context; it has been demonstrated here that for negative \( \lambda \) if one constructs a series in \( (-\lambda)^{-\frac{1}{2}} \), encouraging results can be obtained. The moral is, \( \lambda \) cannot be a proper parameter then; for, if it were so, we should have exact results for
either sign of \(\lambda\). Not even \((-\lambda)^{-\frac{1}{2}}\) is the proper parameter for the corresponding series also fails to work for small \(\lambda\); see also ref. 31.

37. S. Wilson, Preprint, 1979, Daresbury Laboratory, WL/SCI/P 198T.
38. K. Bhattacharyya, Int. J. Quantum Chem. 20, 1273, 1981; here we proceed slightly differently and correct some errors which we have detected later.
48. See e.g. ref. 43; S. Wilson and M.F. Guest, J. Phys. B: 14, 1709, 1981, and refs. in these articles.
51. (Various such approximants can also be found in) A. Dalgarno, in Quantum Theory, Vol.1, 1961 (ed. D.R. Bates; Academic: New York).


63. V. Rojansky, Introductory Quantum Mechanics, 1959 (Prentice-Hall; Englewood Cliffs, NJ); ref.46.


77. Leinaas-Kuo.

78. J.E. Lennard-Jones, Proc. R. Soc. 120, 598, 1930.

79. P.M. Morse and H. Feshbach, Methods of Theoretical Physics, Part 2, 1953 (McGraw-Hill: Kogakusha).


