Extension of the perturbational-variational Rayleigh-Ritz formalism to large order

1983 J. Phys. A: Math. Gen. 16 3471

(http://iopscience.iop.org/0305-4470/16/15/013)

J. Phys. A: Math. Gen. **16** (1983) 3471–3483. Printed in Great Britain

Extension of the perturbational-variational Rayleigh–Ritz formalism to large order

Jeremiah N Silverman

Max-Planck-Institut für Strahlenchemie, D-4330 Mülheim (Ruhr), Federal Republic of Germany

Received 7 December 1982, in final form 14 June 1983

Abstract. The previously derived perturbational-variational Rayleigh–Ritz (PVRR) matrix formalism for calculating the Rayleigh–Schrödinger (RS) perturbation series of the normal modes of astrophysical and other classical systems is greatly strengthened so as to make it a powerful tool for applying large-order perturbation theory to the study of *both* classical and quantum mechanical eigenvalue problems; the underlying generalised Sturm–Liouville type perturbed eigenvalue equation is encountered in many branches of physics. The original formalism is first modified by introducing a number of auxiliary matrices which render it amenable to the recursive calculation of the RS series of eigenvalues, eigenvectors, and related expectation values. It is then extended to large perturbational order by deriving and incorporating the generalised PVRR remainder and Hellmann–Feynman theorems; these PVRR matrix theorems include the corresponding RS quantum mechanical theorems as special cases. The resultant modified and extended formalism is applicable to an arbitrary Rayleigh–Ritz (linear variational) *Ansatz* and enables, in a single computer run, the accurate recursive calculation to extremely high perturbational order of the PVRR series of all eigenstates.

1. Introduction

In the context of a generalised Sturm–Liouville type (see e.g. Courant and Hilbert 1953) perturbed eigenvalue equation of classical origin, a perturbational-variational Rayleigh–Ritz (PVRR) matrix formalism has previously been developed (Silverman and Sobouti 1978) and applied (Sobouti and Silverman 1978, Sobouti 1978) to the calculation of the Rayleigh–Schrödinger (RS) perturbation series of normal modes (eigenvalues and eigenvectors) in certain oscillating astrophysical systems; excellent results were obtained in this manner. The PVRR approach, however, is not limited to astrophysical problems since the Sturm–Liouville eigenvalue equation is encountered in many branches of physics. In this paper, we modify and extend the PVRR formalism so as to make it a powerful tool for applying large-order perturbation theory (LOPT) to the study of *both* classical and quantum mechanical systems.

LOPT has been a subject of increasing interest since the fundamental work of Bender and Wu (1969) and of Simon (1970) on the anharmonic oscillator in connection with quantum field theory. The principal goals of LOPT are to generate very high-order (typically, 100th) RS perturbation series with great accuracy, to study their asymptotic properties with the view in mind of determining the functional behaviour of the solutions, and to sum the series exactly. These developments have recently been thoroughly reviewed by Zinn-Justin (1981), Simon (1982), and Čížek and Vrscay

(1982). Until now, LOPT has not been applied to classical eigenvalue problems. Indeed, its application has been restricted thus far to a few one-particle quantum mechanical systems, such as the anharmonic oscillator and the Stark and Zeeman effects for the hydrogen atom, for which the exact RS series can be determined analytically to arbitrarily high order. Although these series are strongly divergent, they can be summed by a number of powerful LOPT methods, e.g. the recently introduced summability procedure of Silverman (1983a) and references cited therein.

Evidently, it would be of considerable interest to broaden the scope of LOPT to both classical and larger quantal systems. For such systems, however, it is not yet possible to obtain the exact RS series analytically to high order and one must employ numerical approximation methods which combine the variational principle with perturbation theory. The extended PVRR procedure presented here represents such a numerical method, which generalises and strengthens the earlier Dalgarno–Drake (DD) and other closely related quantum mechanical matrix formulations of RS perturbation theory (Dalgarno and Drake 1969, Imamura 1968, Brändas and Goscinski 1970, Carbó 1970, 1972, Carbó and Gallifa 1972). Further, it has been shown (Brändas and Goscinski 1970, Sanders 1972) that the Hylleraas–Scherr–Knight (HSK) variational-perturbational procedure (Hylleraas 1930, Scherr and Knight 1963) and the DD matrix formulations are, in fact, entirely equivalent if a fixed linear basis is used throughout in the former. Both the HSK and DD methods have been widely applied to compute low-to-moderate-order RS $1/Z$ series (where Z is the nuclear charge) of the non-relativistic electronic states of M-electron atomic isoelectronic sequences for $M \leq 4$. Recently, it has been theoretically demonstrated (Silverman 1981) that the radii of convergence of the RS $1/Z$ series are quite small. Since this and other available evidence suggest that RS series with poor convergence properties are the rule rather than the exception, LOPT studies are of great importance.

Conventional RS perturbation theory as applied to the time-independent Schrödinger equation (see, e.g., the review of Hirschfelder *et al* 1964) is inefficient with respect to LOPT because evaluation of the RS series of the exact eigenfunction through nth order, $n = 0, 1, \ldots$, only suffices to compute the RS series of the exact eigenvalue through $(n+1)$th order. Similarly, combination (Carr 1957) of the Hellmann–Feynman theorem (Hellmann 1937, Feynman 1939) with RS perturbation theory, although a valuable supplementary procedure, also only yields the eigenvalue to one order higher than the eigenfunction.

Crucial to the implementation of LOPT is the well known RS quantum mechanical theorem (Hylleraas 1930, Dalgarno and Stewart 1956, Dupont-Bourdelet *et al* 1960, Scherr and Knight 1963, Hirschfelder *et al* 1964, Löwdin 1965, Hirschfelder 1969) which states that the exact eigenfunction series through nth order yields the exact eigenvalue series through $(2n+1)$th order. The theorem also applies to the approximate quantities furnished by the HSK and DD approaches (Brändas and Goscinski 1970, Watson and O'Neil 1975), to the perturbation series obtained within Hartree–Fock (Langhoff *et al* 1966) and extended Hartree–Fock (Coulson and Hibbert 1967, Hibbert 1967) theory, and to the perturbational-variational series derived from an arbitrary variational *Ansatz* containing any number and type of adjustable variational parameters (Silverman and van Leuven 1967, 1968). In accordance with our previous usage (Silverman and van Leuven 1967, Epstein 1968), we refer to these $(2n+1)$ theorems collectively as the *remainder theorem* although some authors designate them as Wigner theorems. The general content of the remainder theorem easily follows

(Sinanoğlu 1961, Epstein 1968) from the stationary nature of both exact and optimal variational solutions to the Schrödinger equation; it is more troublesome, however, to obtain the explicit form of the theorem for arbitrary orders, and the previous derivations of the quantum mechanical versions involve a rather complex and lengthy sequence of algebraic manipulations.

Neither the remainder nor the Hellmann–Feynman theorem has previously been presented within the framework of PVRR. In the present work, both theorems are derived in generalised matrix form and incorporated into the PVRR procedure, thus extending the formalism; in particular, the generalised PVRR remainder theorem is derived in a new and elegant manner. These generalised PVRR theorems include the corresponding RS quantum mechanical theorems as special cases.

This article is organised as follows. The generalised perturbed eigenvalue problem is introduced in § 2; the previous basic PVRR formalism is modified in § 3; the PVRR matrix theorems are derived in § 4; and the extended PVRR formalism is described in § 5. Finally, in § 6 we briefly discuss our results.

2. The generalised perturbed eigenvalue problem

Consider the generalised Sturm–Liouville type perturbed eigenvalue equation

$$(\mathscr{H}-\varepsilon^s\sigma)|\psi^s\rangle=0, \qquad s=1,2,\ldots, \tag{1}$$

where $\mathscr{H}$ is a linear second-order Hermitian operator, σ a positive definite weighting operator dependent upon the configuration coordinates of the system, and $|\psi^s\rangle$ and ε^s are, respectively, the eigenfunction and eigenvalue of the sth discrete eigenstate. In classical mechanics, (1) occurs frequently in vibration problems where it is introduced by the theory of small oscillations; for example, in the previously mentioned astrophysical problems, $\mathscr{H}$ is a second-order integrodifferential operator, σ is the matter density, and the normal modes of the system are described by the eigendisplacement vectors $|\psi^s\rangle\exp(\mathrm{i}\omega^s t)$ where the angular frequencies ω^s are given by $\omega^s=(\varepsilon^s)^{1/2}$. In quantum mechanics, (1) simplifies since σ is invariably unity; undoubtedly, the most important example is the time-independent Schrödinger equation where $\mathscr{H}$ is the Hamiltonian operator and the ε^s are the energy levels. Throughout, we primarily focus our attention on the more general classical eigenvalue equation; our results, however, can be readily specialised to the quantum mechanical case.

Formally, the perturbation enters (1) through a real coupling parameter λ, independent of the configuration coordinates of the system, which is embedded in $\mathscr{H}$ and σ. Thus, we can write

$$\mathscr{H}=\mathscr{H}(\lambda), \qquad \sigma=\sigma(\lambda), \tag{2a, b}$$

whence

$$|\psi^s\rangle=|\psi^s(\lambda)\rangle, \qquad \varepsilon^s=\varepsilon^s(\lambda). \tag{2c, d}$$

The nature of λ depends, of course, upon the specific problem at hand but, in general, two quite different categories occur. In the first category, λ is a natural parameter which can vary continuously or discretely in a neighbourhood of $\lambda=0$, thus parametrising a family of solutions; in the second category, λ is introduced to serve as a dummy ordering parameter in a perturbational treatment so that it can only assume the physically significant value of unity. For generality, we treat the more difficult case

of natural λ; subsequently, in § 5, we show how the resultant PVRR formalism can be simplified for dummy λ.

Except for a few simple systems such as those previously mentioned, (1) can neither be solved exactly in closed form nor via RS perturbation theory in series form. As is well known, in the overwhelming majority of cases, one must resort to expanding the $|\psi^s\rangle$ in terms of a complete basis set. Let the $|\varphi^t\rangle$, $t = 1, 2, \ldots$, be an arbitrary complete *discrete* set of linearly independent basis functions which span the Hilbert space of the $|\psi^s\rangle$ and satisfy the same boundary conditions as the latter. In general, the $|\varphi^t\rangle$ are selected to be independent of λ and, hence, cannot be orthonormal† with respect to the weighting function $\sigma(\lambda)$. Note that unlike conventional RS perturbation theory, the basis set are not taken to be the eigenfunctions of $\mathscr{H}_0 = \mathscr{H}(0)$ as the complete spectrum of these may not be known and/or may include continuum solutions; the advantages of using such an arbitrary discrete basis for atomic and molecular calculations have previously been noted by Shull and Löwdin (1959), Blinder (1960) and more recently by Ladik and Čížek (1980). Now, expand the $|\psi^s\rangle$ as

$$|\psi^s\rangle = \sum_t |\varphi^t\rangle C^{ts}, \tag{3}$$

where the C^{ts} are linear expansion coefficients which form a column vector C^s for a given s. Introduce the square matrices H, S, C, and E, where

$$H = [H^{ts}], \qquad H^{ts} = \langle\varphi^t|\mathscr{H}|\varphi^s\rangle, \tag{4a, b}$$

$$S = [S^{ts}], \qquad S^{ts} = \langle\varphi^t|\sigma|\varphi^s\rangle, \tag{5a, b}$$

C is the matrix formed by collecting the column vectors C^s,

$$C = [C^s] = [C^{ts}], \tag{6a}$$

and E is the diagonal matrix of the eigenvalues ε^s,

$$E = [\varepsilon^s \delta_{ts}]. \tag{6b}$$

Then, in the usual manner, one obtains the familiar matrix eigenvalue equation

$$HC = SCE, \tag{7a}$$

which collects equations (1) in matrix representation for all s; the classical (7*a*) is more general, however, than the formally equivalent quantum mechanical matrix equation constructed with a non-orthonormal basis as here S is not merely the overlap matrix of the basis functions. Since H and S are Hermitian and, further, S is positive definite, it follows from matrix theory that a nonsingular solution C of (7*a*) exists which simultaneously diagonalises S to the unit matrix I and H to the eigenvalue matrix E:

$$C^\dagger S C = I, \qquad C^\dagger H C = E. \tag{7b), (8}$$

In passing, we note that the basis set $|\varphi^t\rangle$ can always be initially orthonormalised with respect to σ to form a new set of λ-dependent basis functions. Although this would eliminate the S matrix from (7), it would do so at the cost of introducing a more complicated λ-dependency in the correspondingly transformed H matrix; in the case

† The Sturm–Liouville generalised orthonormality condition for the basis set is $\langle\varphi^t|\sigma|\varphi^u\rangle = \delta_{tu}$ where δ_{tu} is the Kronecker delta.

of natural λ, this would largely divest the PVRR formalism of its computational simplicity. We return to this point in § 5.

Equations (7) and (8) are exact since they are of infinite order involving expansions in terms of a complete discrete set. In the classical Rayleigh–Ritz (RR) procedure (see, e.g., Gould 1957), the expansion (3) is approximated by truncation to a finite number of terms, say N, yielding an RR *Ansatz*; in quantum mechanics, of course, this procedure is commonly termed the method of configuration interaction. All matrices correspondingly truncate to Nth order and may be denoted by $\hat{H}$, $\hat{S}$, $\hat{C}$, and $\hat{E}$. Due to the variational nature of the problem, (7) and (8) remain unchanged in form although now applying to the RR-approximated finite matrices. For a *fixed* value of λ, (7*a*) can be solved for $\hat{E}$ and $\hat{C}$ via the standard RR procedure imposing normalisation of the $\hat{C}^s$ with respect to $\hat{S}$ through (7*b*). In this paper, we deal with the practical problem of strengthening the PVRR formalism for application to an arbitrary RR *Ansatz*. Nevertheless, all subsequent results hold equally for both exact and RR-approximated solutions for all states; therefore, we need not distinguish between them and, in what follows, we suppress the carets for brevity.

Finally, in quantum mechanical applications, one has frequent occasion to compute the expectation value for the sth state, $\langle w\rangle^s$, of an arbitrary operator w independent of λ. Let $\langle w\rangle$ and $[X]^{\text{diag}}$ denote diagonal matrices which are constructed, respectively, from the $\langle w\rangle^s$ for all states and from the diagonal elements of an arbitrary square matrix X. Then, in compact RR notation, we have for a given value of λ,

$$\langle w\rangle = [C^{\dagger} W C]^{\text{diag}}, \tag{9}$$

where

$$W = [W^{ts}], \qquad W^{ts} = \langle \varphi^t | w | \varphi^s \rangle. \tag{10a, b}$$

3. Modification of the basic PVRR formalism

Unlike the above RR approach, PVRR focuses attention on the λ-dependency of all matrices in (7)–(9); this λ-dependency is an immediate consequence of (2)–(6). Let $\mathcal{H}$ and σ be analytic functions of natural λ in a neighbourhood of $\lambda = 0$. Then $\mathcal{H}$ and σ admit convergent power-series expansions in λ,

$$\mathcal{A}(\lambda) = \sum_{j=0}^{\infty} \mathcal{A}_j \lambda^j, \qquad \mathcal{A} = \mathcal{H}, \sigma; \tag{11a}$$

although in the majority of quantum mechanical applications, λ only occurs linearly in the Hamiltonian $\mathcal{H}$, the general power series is frequently encountered in classical problems. Assume that similar RS convergent λ-expansions exist for $|\psi^s\rangle$ and ε^s. Thus, we take

$$\mathcal{B}(\lambda) = \sum_{j=0}^{\infty} \mathcal{B}_j \lambda^j, \qquad \mathcal{B} = |\psi^s\rangle, \varepsilon^s, \qquad s = 1, 2, \ldots. \tag{11b}$$

We have assumed that the power series of (11) are convergent, but this is not essential to the PVRR formalism since, as previously mentioned, powerful LOPT methods exist for summing slowly convergent or divergent RS series. Note that the power series in (11*a*) are regarded as known while those of (11*b*) are to be determined via PVRR. To this end, substitute (11) appropriately into (3)–(6) which yields the PVRR λ-expansions

for the various matrices,

$$A(\lambda) = \sum_{j=0}^{\infty} A_j \lambda^j, \qquad A = H, S, \tag{12a}$$

$$A_j = [A_j^{ts}], \qquad A_j^{ts} = \langle \varphi^t | \mathscr{A}_j | \varphi^s \rangle, \tag{12b, c}$$

and

$$B(\lambda) = \sum_{j=0}^{\infty} B_j \lambda^j, \qquad B = C, E, \langle w \rangle, \tag{12d}$$

where the computationally simple structure of the matrices H_j and S_j is a direct consequence of the fact that the $|\varphi^t\rangle$ are independent of λ; further, it follows from (12) that the H_j and S_j are individually Hermitian and that the E_j are diagonal with elements ε_j^s. On substituting the λ-expansions of H, S, C, and E into (7) and equating coefficients of like powers of λ, one obtains the hierarchy of coupled PVRR equations,

$$\sum_{k=0}^{j} H_k C_{j-k} = \sum_{k=0}^{j} \sum_{i=0}^{j-k} S_k C_i E_{j-k-i}, \qquad j = 0, 1, \ldots, \tag{13a}$$

$$\sum_{k=0}^{j} \sum_{i=0}^{j-k} C_k^{\dagger} S_i C_{j-k-i} = I \delta_{0j}, \qquad j = 0, 1, \ldots, \tag{13b}$$

upon which much of the subsequent PVRR formalism is based. The zeroth-order equations,

$$H_0 C_0 = S_0 C_0 E_0, \qquad C_0^{\dagger} S_0 C_0 = I, \tag{14a, b}$$

are solved for E_0 and C_0 by standard diagonalisation where the elements ε_0^s of E_0 are the roots of the secular determinant

$$|H_0 - \varepsilon_0 S_0| = 0. \tag{14c}$$

In the present treatment, it is required that the ε_0^s form a nondegenerate spectrum. Since, in general, the basis set $|\varphi^t\rangle$ are not the eigenfunctions of $\mathscr{H}_0$, one is at liberty to define λ and shift its origin as one pleases. Thus, in the PVRR approach, zero-order exact or near degeneracy can always be simply removed by Taylor-expanding $\mathscr{H}$ and σ, (11*a*), about some other suitable choice of origin, $\lambda_0 \neq 0$, to generate a new H_0 and S_0 with a nondegenerate set of ε_0^s; it is assumed throughout that this has been done if necessary. As is well known, removal of zero-order degeneracy in conventional RS perturbation theory is far more troublesome.

In the original PVRR procedure of Silverman and Sobouti (1978), the E_j and C_j, $j = 1, 2, \ldots$, are determined successively, order-by-order, by premultiplying (13*a*) with $C_0^{\dagger}$ and considering, respectively, the diagonal and off-diagonal elements of the resultant equations; their formulation, however, is not well suited for computer programming, particularly for large-order PVRR series. We now modify this basic formalism so as to render it amenable for recursive calculations to arbitrarily high order; we also obtain the PVRR series of the quantum mechanical expectation values of (9) as these have not been previously considered. Our results can be expressed succinctly in a form suitable for the *concurrent* determination of the PVRR series for

all N states by introducing the auxiliary matrices P_j and Q_j defined as

$$P_j \equiv C_0^{\dagger}\Big(\sum_{k=1}^{j} H_k C_{j-k} - \sum_{k=0}^{j}\sum_{i=0}^{j-k}(1-\delta_{k0}\delta_{i0})(1-\delta_{k0}\delta_{ij})S_k C_i E_{j-k-i}\Big), \qquad j=1,2,\ldots, \tag{15}$$

and

$$Q_j^{tu} \equiv (\varepsilon_0^u - \varepsilon_0^t)^{-1} P_j^{tu}, \qquad t \neq u, \qquad j=1,2,\ldots,. \tag{16a}$$

$$Q_j^{tt} \equiv -\frac{1}{2}\Big(\sum_{k=0}^{j}\sum_{i=0}^{j-k}(1-\delta_{k0}\delta_{i0})(1-\delta_{kj}\delta_{i0})C_k^{\dagger}S_i C_{j-k-i}\Big)^{tt}, \qquad j=1,2,\ldots. \tag{16b}$$

Then, it is easily shown that the modified basic formalism yields

$$E_j = [P_j]^{\mathrm{diag}}, \qquad j=1,2,\ldots, \tag{17a}$$

$$C_j = C_0 Q_j, \qquad j=1,2,\ldots, \tag{17b}$$

$$\langle w\rangle_j = \Big[\sum_{k=0}^{j} C_k^{\dagger} W C_{j-k}\Big]^{\mathrm{diag}}, \qquad j=0,1,\ldots. \tag{17c}$$

Note that both the off-diagonal and diagonal elements of the P_j and Q_j are required where, in (17*b*), the diagonal elements of Q_j impose the orthonormality conditions of (13*b*) on the C_j. In analogy with conventional RS theory, it follows from (15) and (17*a*) that calculation of the E_j with this procedure requires a knowledge of the lower-order C_k and E_k through $k=j-1$, while from (17*c*), calculation of the $\langle w\rangle_j$ for an arbitrary expectation value requires all lower order C_k through $k=j$.

In the following section, we derive two PVRR matrix theorems, valid for all N levels of an arbitrary RR *Ansatz*, which enable the basic PVRR formalism to be extended to the efficient calculation of very high order E_j.

4. Derivation of PVRR matrix theorems

4.1. Generalised PVRR remainder theorem

Consider the case of the even orders, E_{2n}, $n=1,2,\ldots$. Premultiply the first n equations of (13*a*), $j=0,1,\ldots,n-1$ by $-C_{2n-j}^{\dagger}$, *omit* the $(n+1)$th equation, $j=n$, premultiply the next n equations, $j=n+1, n+2,\ldots,2n$ by $C_{2n-j}^{\dagger}$, and write down the resulting $2n$ product-equations in a triangular array, retaining only the diagonal elements. This yields

$$\begin{aligned}
&-C_{2n}^{s\dagger}G_0C_0^s &&= 0\\
&-C_{2n-1}^{s\dagger}G_0C_1^s - C_{2n-1}^{s\dagger}G_1C_0^s &&= 0\\
&\quad\vdots\\
&-C_{n+1}^{s\dagger}G_0C_{n-1}^s - C_{n+1}^{s\dagger}G_1C_{n-2}^s - \ldots - C_{n+1}^{s\dagger}G_{n-1}C_0^s &&= 0\\
&C_{n-1}^{s\dagger}G_0C_{n+1}^s + C_{n-1}^{s\dagger}G_1C_n^s + \ldots + C_{n-1}^{s\dagger}G_{n-1}C_2^s + C_{n-1}^{s\dagger}G_nC_1^s + \ldots &&= 0\\
&\quad\vdots\\
&C_1^{s\dagger}G_0C_{2n-1}^s + C_1^{s\dagger}G_1C_{2n-2}^s + \ldots + C_1^{s\dagger}G_{n-1}C_n^s + C_1^{s\dagger}G_nC_{n-1}^s + \ldots &&= 0\\
&C_0^{s\dagger}G_0C_{2n}^s + C_0^{s\dagger}G_1C_{2n-1}^s + \ldots + C_0^{s\dagger}G_{n-1}C_{n+1}^s\\
&\qquad + C_0^{s\dagger}G_nC_n^s + \ldots + C_0^{s\dagger}G_{2n}C_0^s &&= 0
\end{aligned} \tag{18a}$$

where for brevity of presentation we have introduced the Hermitian matrix G defined as

$$G \equiv H - \varepsilon^s S, \tag{19a}$$

with the PVRR λ-expansion,

$$G(\lambda) = \sum_{j=0}^{\infty} G_j \lambda^j, \qquad G_j = H_j - \sum_{i=0}^{j} \varepsilon_i^s S_{j-i}. \tag{19b,c}$$

Now sum the array (18a) in a column-by-column manner, using the Hermitian character of all triple-product terms. Cancellation in pairs of all terms in C_{n+1}^s, $C_{n+2}^s, \ldots, C_{2n}^s$ evidently occurs and one easily obtains

$$\sum_{j=1}^{n} \sum_{k=n-j}^{n-1} C_k^{s\dagger} G_j C_{2n-j-k}^s + \sum_{j=n+1}^{2n} \sum_{k=0}^{2n-j} C_k^{s\dagger} G_j C_{2n-j-k}^s = 0, \qquad n = 1, 2, \ldots. \tag{20a}$$

Equation (20a) is one formulation of the desired remainder theorem for the even orders of the eigenvalue series since the highest-order eigenvector appearing is C_n^s and the highest order eigenvalue is ε_{2n}^s (in G_{2n}); we shall shortly bring this expression into more convenient form for recursive calculations for all N states simultaneously.

In the case of the odd orders, E_{2n+1}, $n = 1, 2, \ldots$, we proceed in a slightly modified manner. Premultiply the first $(n+1)$ equations of (13a), $j = 0, 1, \ldots, n$ by $-C_{2n+1-j}^{\dagger}$ and the second $(n+1)$ equations, $j = n+1, n+2, \ldots, 2n+1$ by $C_{2n+1-j}^{\dagger}$, producing this time $(2n+2)$ product-equations; as before, we write these down in a triangular array, retaining only the diagonal terms. Thus,

$$\begin{aligned}
&-C_{2n+1}^{s\dagger} G_0 C_0^s && = 0\\
&-C_{2n}^{s\dagger} G_0 C_1^s - C_{2n}^{s\dagger} G_1 C_0^s && = 0\\
&\quad\vdots\\
&-C_{n+1}^{s\dagger} G_0 C_n^s - C_{n+1}^{s\dagger} G_1 C_{n-1}^s - \ldots - C_{n+1}^{s\dagger} G_n C_0^s && = 0\\
&C_n^{s\dagger} G_0 C_{n+1}^s + C_n^{s\dagger} G_1 C_n^s + \ldots + C_n^{s\dagger} G_n C_1^s + C_n^{s\dagger} G_{n+1} C_0^s && = 0\\
&\quad\vdots\\
&C_1^{s\dagger} G_0 C_{2n}^s + C_1^{s\dagger} G_1 C_{2n-1}^s + \ldots + C_1^{s\dagger} G_n C_n^s + C_1^{s\dagger} G_{n+1} C_{n-1}^s + \ldots && = 0\\
&C_0^{s\dagger} G_0 C_{2n+1}^s + C_0^{s\dagger} G_1 C_{2n}^s + \ldots + C_0^{s\dagger} G_n C_{n+1}^s\\
&\qquad + C_0^{s\dagger} G_{n+1} C_n^s + \ldots + C_0^{s\dagger} G_{2n+1} C_0^s && = 0.
\end{aligned} \tag{18b}$$

Summation of (18b) in the same manner as (18a) yields

$$\sum_{j=1}^{n} \sum_{k=n+1-j}^{n} C_k^{s\dagger} G_j C_{2n+1-j-k}^s + \sum_{j=n+1}^{2n+1} \sum_{k=0}^{2n+1-j} C_k^{s\dagger} G_j C_{2n+1-j-k}^s = 0, \qquad n = 1, 2, \ldots, \tag{20b}$$

where all terms in C_{n+1}^s, $C_{n+2}^s, \ldots, C_{2n+1}^s$ have cancelled in pairs. In (20b), which expresses the remainder theorem for the odd orders of the eigenvalue series, the highest-order eigenvector appearing is C_n^s and the highest-order eigenvalue is ε_{2n+1}^s.

Equations (20a) and (20b) are now solved respectively for ε_{2n}^s and ε_{2n+1}^s. The results can be expressed compactly in a recursive all-states form by introducing the

auxiliary matrices R_{2n} and R_{2n+1} defined as

$$R_{2n} \equiv \sum_{j=1}^{n} \sum_{k=n-j}^{n-1} C_k^{\dagger}\left(H_j C_{2n-j-k} - \sum_{i=0}^{j} S_{j-i} C_{2n-j-k} E_i\right) + \sum_{j=n+1}^{2n} \sum_{k=0}^{2n-j} C_k^{\dagger}\left(H_j C_{2n-j-k} - \sum_{i=0}^{n} S_{j-i} C_{2n-j-k} E_i\right), \qquad n = 1, 2, \ldots,$$

$$R_{2n+1} \equiv \sum_{j=1}^{n} \sum_{k=n+1-j}^{n} C_k^{\dagger}\left(H_j C_{2n+1-j-k} - \sum_{i=0}^{j} S_{j-i} C_{2n+1-j-k} E_i\right) + \sum_{j=n+1}^{2n+1} \sum_{k=0}^{2n+1-j} C_k^{\dagger}\left(H_j C_{2n+1-j-k} - \sum_{i=0}^{n} S_{j-i} C_{2n+1-j-k} E_i\right), \qquad n = 1, 2, \ldots, \tag{21}$$

where only the diagonal elements are required. Then,

$$\begin{aligned} E_{2n} &= [R_{2n}]^{\text{diag}}, \\ E_{2n+1} &= [R_{2n+1}]^{\text{diag}}, \end{aligned} \qquad n = 1, 2, \ldots. \tag{22}$$

Equations (21) and (22), which are the final form of the generalised PVRR remainder theorem, represent a new result. It is important to note that not only the C_k but also the E_k, $k = n+1, n+2, \ldots, 2n-1$, have been eliminated from the R_{2n} and R_{2n+1} matrices; the removal of these E_k results from the automatic orthonormalisation via (16*b*) and (17*b*) of the eigenvector matrix C through nth order with respect to S when each successive order C_n is computed. Thus, calculation of the E_{2n} and E_{2n+1} from (22) requires only a knowledge of the lower-order C_k and E_k through $k = n$; this elimination of terms leads to a substantial reduction in computational labour, particularly for large n.

The essential simplicity of our derivation of the completely generalised classical remainder theorem should be contrasted with the complexity of the previous derivations of the various specialised quantum mechanical remainder theorems (e.g., Dupont-Bourdelet *et al* 1960). By setting the $S_j = I\delta_{0j}$ in (21), one recovers the most general PVRR quantum mechanical remainder theorem for an orthonormal basis set independent of λ; in particular, for the common special case of linearly perturbed Hamiltonians, $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1\lambda$, $\mathcal{H}_j = 0$, $j = 2, 3, \ldots$, these expressions further reduce to

$$\begin{aligned} E_{2n} &= \left[C_n^{\dagger} H_1 C_{n-1} - \sum_{j=1}^{n} \sum_{k=n-j}^{n-1} C_k^{\dagger} C_{2n-j-k} E_j\right]^{\text{diag}}, \\ E_{2n+1} &= \left[C_n^{\dagger} H_1 C_n - \sum_{j=1}^{n} \sum_{k=n+1-j}^{n} C_k^{\dagger} C_{2n+1-j-k} E_j\right]^{\text{diag}}, \end{aligned} \qquad n = 1, 2, \ldots. \tag{23}$$

Equation (23) generalises the previous result of Hirschfelder *et al* (1964) for the RS series of exact solutions to the PVRR series of the N levels of an arbitrary RR *Ansatz*.

4.2. Generalised PVRR Hellmann–Feynman theorem

Our starting point is the generalised Hellmann–Feynman theorem (Silverman 1983b) which can be written in the all-states matrix formulation,

$$\frac{\partial E}{\partial \lambda} = \left[C^{\dagger}\left(\frac{\partial H}{\partial \lambda} C - \frac{\partial S}{\partial \lambda} CE\right)\right]^{\text{diag}}. \tag{24}$$

The formal equivalent of (24) has previously been derived in the context of the Schrödinger equation by Löwdin (1959); in quantum mechanics, however, as previously mentioned, S is merely the overlap matrix of the basis functions, and the possible λ-dependency of S can only enter through the basis set.

Substitute the λ-expansions of all quantities into (24) and equate coefficients of like powers of λ. By introducing the auxiliary matrix F_j defined as

$$F_j \equiv j^{-1} \sum_{k=0}^{j-1} \sum_{i=0}^{j-k-1} (i+1) C_k^{\dagger} \Big(H_{i+1} C_{j-k-i-1} - S_{i+1} \sum_{m=0}^{j-k-i-1} C_m E_{j-k-i-m-1} \Big), \qquad j = 1, 2, \ldots, \tag{25a}$$

one can express the E_j compactly as

$$E_j = [F_j]^{\text{diag}}, \qquad j = 1, 2, \ldots. \tag{25b}$$

Equations (25) are the generalised PVRR Hellmann–Feynman theorem, a new result. In analogy with the R_{2n} and R_{2n+1}, (21), only the diagonal elements of the F_j need be computed. As in the case of (17a), the calculation of the E_j via (25) requires a knowledge of all C_k and E_k through $k = j-1$. Note that equations (25) reduce to the most general quantum mechanical formulation of the PVRR Hellmann–Feynman theorem for an orthonormal basis independent of λ on setting $S_j = I\delta_{0j}$. In the special case of linearly perturbed Hamiltonians, these expressions further reduce to

$$E_j = \Big[j^{-1} \sum_{k=0}^{j-1} C_k^{\dagger} H_1 C_{j-k-1} \Big]^{\text{diag}}, \qquad j = 1, 2, \ldots, \tag{26}$$

which generalises the result of Carr (1957) for the RS series of exact solutions to the PVRR series of all levels of an arbitrary RR *Ansatz*.

5. Extended PVRR formalism

We now extend the modified PVRR formalism of § 3 to large order by means of the generalised Hellmann–Feynman and remainder theorems; the most efficient way of accomplishing this is exhibited schematically in table 1. The extended PVRR calculations assume a natural cyclic structure where each cycle consists of three parts, the first pertaining to the modified basic formalism, the second to the Hellmann–Feynman theorem, and the third to the remainder theorem. After the nth cycle is completed, the C_j, E_j, and $\langle w \rangle_j$ will have been computed for $j = 0, 1, \ldots, n$; in particular, E will have been determined through nth order by three independent methods, thus affording a valuable check on the internal consistency and accuracy of the PVRR calculations. Further, the E_j, $j = n+1, n+2, \ldots, 2n+1$, will have been generated via the remainder theorem. The three parallel calculations of the E_j, $j = 1, 2, \ldots, n$, involve a relatively modest additional computational effort since in each order, only the diagonal elements of the F_j, R_{2j}, and R_{2j+1} matrices are required.

The present formalism entails the concurrent calculation of the PVRR series of the N lowest states of a given symmetry where N is the degree of truncation of the basis set. Although the PVRR series are perturbationally exact within the framework of a given finite RR *Ansatz*, they are, of course, variational approximations to the corresponding exact RS series. For systems where one requires a large N to obtain accurate variational convergence, the concurrent calculation for all N states becomes impracticable due to computer storage limitations and the PVRR series of the high-lying states

Table 1. Cyclic flow of extended PVRR calculations incorporating the Hellmann–Feynman and remainder theorems.

j	Equations	Highest-order input	Auxiliary matrices	Output
0	(14), (17c)			$E_0, C_0, \langle w\rangle_0$
1	(17)	E_0, C_0	P_1, Q_1	$E_1, C_1, \langle w\rangle_1$
	(25)	E_0, C_0	F_1	E_1
	(22)	E_1, C_1	R_2, R_3	E_2, E_3
2	(17)	E_1, C_1	P_2, Q_2	$E_2, C_2, \langle w\rangle_2$
	(25)	E_1, C_1	F_2	E_2
	(22)	E_2, C_2	R_4, R_5	E_4, E_5
	⋮	⋮	⋮	⋮
n	(17)	E_{n-1}, C_{n-1}	P_n, Q_n	$E_n, C_n, \langle w\rangle_n$
	(25)	E_{n-1}, C_{n-1}	F_n	E_n
	(22)	E_n, C_n	R_{2n}, R_{2n+1}	E_{2n}, E_{2n+1}

are then of limited interest due to their generally poorer variational convergence. In such cases, one can readily reduce the N-states formalism from first order onwards to the calculation of the PVRR series of one state at a time for any desired values of $s \leqslant N$; the full-scale solution of the zeroth-order equations (14) for all N states must, of course, still be obtained.

Finally, we deal with the problem, previously touched upon in § 2, of eliminating S from (7a) and, hence, the S_j from the PVRR formalism. Consider the case of dummy λ. Here, the series expansions of $\mathscr{H}$ and σ, (11a), are purely formal and result from some arbitrary partitioning scheme. We are at liberty, then, to take σ as independent of λ. Alternatively, for natural λ, σ may also be independent of λ in certain systems (e.g., all quantum mechanical systems). In these special but widely encountered cases, one can advantageously carry out the desired simplification of the formalism via the Löwdin (1950) technique of symmetric orthonormalisation; this procedure utilises a transformation matrix $S^{-1/2}$ with the property $S^{-1/2}SS^{-1/2} = I$. Equations (7) are reduced in this manner to the canonical form,

$$\tilde{H}\tilde{C} = \tilde{C}E, \qquad \tilde{C}^{\dagger}\tilde{C} = I, \tag{27a, b}$$

where the transformed $\tilde{H}$ and $\tilde{C}$ are related to H and C by

$$S^{-1/2}HS^{-1/2} = \tilde{H}, \qquad C = S^{-1/2}\tilde{C}, \tag{28a, b}$$

and the λ-expansions of E and $\tilde{C}$ are to be obtained from (27) with the PVRR procedure. In the special cases mentioned above, $S^{-1/2}$ is independent of λ so there is no difficulty† in forming the hierarchy of PVRR equations from (27a) corresponding to (13a). The remainder of the formalism then goes through unchanged and all subsequent PVRR equations are simply obtained by respectively replacing the H_j, S_j, and C_j with $\tilde{H}_j$, $\delta_{0j}I$, and $\tilde{C}_j$ in our previous general results.

† In the general case of natural λ, it follows from (2b) that $S^{-1/2}$ is dependent upon λ. Thus, to form the λ-expansion of $\tilde{H}$, (28a), required for input, one would have to apply the PVRR formalism initially to generate the λ-expansion of $S^{-1/2}$. The computational labour required for this step alone is of the same order of magnitude as that required for the complete PVRR solution of the original non-orthonormal problem.

6. Discussion

In this paper, we have greatly strengthened the original PVRR formalism of Silverman and Sobouti (1978). To accomplish this, we have first modified the basic formalism so as to render it highly adaptable to computer programming; we have then extended its scope to LOPT studies of both classical and quantum mechanical eigenvalue problems by deriving and incorporating the generalised PVRR remainder and Hellmann–Feynman theorems. The modified and extended formalism is extremely compact because the explicit formulations of conventional RS perturbation theory, which become increasingly cumbersome in higher order, have been completely avoided through the introduction of several auxiliary matrices. In a single computer run, the new procedure permits the accurate recursive calculation to very high perturbational order of the PVRR series of all N levels of an arbitrary RR *Ansatz*.

We have thoroughly tested and verified all phases of the extended PVRR formalism by applying it to the calculation of $1/Z$ expansions for atomic isoelectronic sequences and to the perturbational study of certain molecular problems; these results, which illustrate the power of the new procedure, will be presented elsewhere.

Acknowledgments

The author wishes to thank Professor O E Polansky for his interest in this research and for his support.

References

Bender C M and Wu T T 1969 *Phys. Rev.* **184** 1231–60
Blinder S M 1960 *J. Chem. Phys.* **32** 111–5
Brändas E and Goscinski O 1970 *Phys. Rev.* A **1** 552–60
Carbó R 1970 *Theor. Chim. Acta* **17** 74–80
—— 1972 *Int. J. Quantum Chem.* **6** 609–11
Carbó R and Gallifa R 1972 *Nuovo Cimento* **10B** 576–82
Carr W J, Jr 1957 *Phys. Rev.* **106** 414–7
Čížek J and Vrscay E R 1982 *Int. J. Quantum Chem.* **21** 27–68
Coulson C A and Hibbert A 1967 *Proc. Phys. Soc.* **91** 33–43
Courant R and Hilbert D 1953 *Methods of Mathematical Physics* vol I (New York: Interscience) pp 275–396
Dalgarno A and Drake G W F 1969 *Chem. Phys. Lett.* **3** 349–50
Dalgarno A and Stewart A L 1956 *Proc. R. Soc.* A **238** 269–75
Dupont-Bourdelet F, Tillieu J and Guy J 1960 *J. Phys. Radium* **21** 776–8
Epstein S T 1968 *J. Chem. Phys.* **48** 4725–6
Feynman R P 1939 *Phys. Rev.* **56** 340–3
Gould S H 1957 *Variational Methods for Eigenvalue Problems* (Toronto: University of Toronto Press)
Hellmann H 1937 *Einführung in die Quantenchemie* (Leipzig: Deuticke) pp 285–6
Hibbert A 1967 *Proc. Phys. Soc.* **91** 819–24
Hirschfelder J O 1969 *Int. J. Quantum Chem.* **3** 731–48
Hirschfelder J O, Brown W B and Epstein S T 1964 *Adv. Quantum Chem.* **1** 255–374
Hylleraas E A 1930 *Z. Phys.* **65** 209–25
Imamura A 1968 *Molec. Phys.* **15** 225–38
Ladik J and Čížek J 1980 *J. Chem. Phys.* **73** 2357–9
Langhoff P W, Lyons J D and Hurst R P 1966 *Phys. Rev.* **148** 18–25

Löwdin P-O 1950 *J. Chem. Phys.* **18** 365–75
—— 1959 *J. Molec. Spectrosc.* **3** 46–66
—— 1965 *J. Math. Phys.* **6** 1341–53
Sanders F C 1972 *Chem. Phys. Lett.* **17** 291–3
Scherr C W and Knight R E 1963 *Rev. Mod. Phys.* **35** 436–42
Shull H and Löwdin P-O 1959 *J. Chem. Phys.* **30** 617–26
Silverman J N 1981 *Phys. Rev.* A **23** 441–54
—— 1983a *Phys. Rev.* A **28** 498–501
—— 1983b *Int. J. Quantum Chem.* submitted for publication
Silverman J N and van Leuven J C 1967 *Phys. Rev.* **162** 1175–94
—— 1968 *Phys. Rev.* **169** 1398
Silverman J N and Sobouti Y 1978 *Astron. Astrophys.* **62** 355–63
Simon B 1970 *Ann. Phys., NY* **58** 76–136
—— 1982 *Int. J. Quantum Chem.* **21** 3–25
Sinanoğlu O 1961 *J. Chem. Phys.* **34** 1237–40
Sobouti Y 1978 *Astron. Astrophys.* **70** 665–75
Sobouti Y and Silverman J N 1978 *Astron. Astrophys.* **62** 365–74
Watson D K and O'Neil S V 1975 *Phys. Rev.* A **12** 729–35
Zinn-Justin J 1981 *Phys. Rep.* **70** 109–67