Theoret. Chim. Acta (Berl.) 49, 131–142 (1978)

THEORETICA CHIMICA ACTA

On a Formal Treatment of the Rayleigh–Schrödinger Perturbation Theory

Dietrich Haase

Institut für Quantenchemie der Freien Universität Berlin, Holbeinstraße 48, D-1000 Berlin 45

The formal treatment of the Rayleigh–Schrödinger perturbation theory based on a first-order iteration procedure as described in a previous publication is discussed with reference to the properties of the terms of a Taylor series. The formalism is generalized to allow for multiple perturbation.

Key words: Rayleigh–Schrödinger perturbation theory

1. Introduction

For the eigenvalues and eigenvectors of a perturbed Hamiltonian in Rayleigh–Schrödinger (RS) perturbation theory, formal solutions can be derived by different procedures such as the direct solving of the set of perturbation equations [1–3], operator techniques [4–6], or analytical methods [7, 8]. Based on a first-order iteration procedure, a new concise and useful operator treatment of the RS perturbation theory was developed in collaboration with E. Ruch [9]. Although there are similarities in form, this treatment and the Lie-algebraic treatments of the perturbation theory [10, 11] are quite different perturbation schemes. The first-order perturbation iteration method called FOPIM [12, 3] for the derivation of explicit solutions of the RS perturbation equations also differs from our method since in that procedure, in contrast to ours, for each iteration a specific operation is performed.

For non-degenerate states, our formal treatment of the RS perturbation theory is based on the fact that the first-order eigenvector expressed in terms of unperturbed eigenvectors and of the perturbation operator can be used as a "zeroth-order" term for the calculation of an improved "first-order" correction which is equal to the second-order eigenvector. This process can be iterated. Thus, the n'th-order term $f^{(n)}$ of the power series expansion $\sum_n \lambda^n f^{(n)}$ for a RS eigenvalue or eigen-

0040-5744/78/0049/0131/$02.40

vector is expressible in the compact form $f^{(n)}=(1/n!)\not{p}^n f^{(0)}$, where $\not{p}$ is a conveniently defined operator. The procedure permits the determination of the explicit expressions for the energies and eigenvectors in non-degenerate RS perturbation theory in a systematic and easily understandable way.

The method can be used correspondingly to determine, for instance, the RS expressions for projectors onto degenerate subspaces. There exists an analogous treatment of the Dirac perturbation theory for state vectors, for the evolution operator, the statistical operator, or for quantum mechanically conserved quantities [13].

The formalism was proved to be well adapted to certain physical questions in that it provides insights into structural features of perturbation expressions which, in general, might not be provided by the usual formulation such as the connection between the n'th-order and $(n+1)$'th-order terms. This was demonstrated by a simple proof of a linked cluster theorem [9, 14] and the quantum mechanical adiabatic theorem [13]. In these applications, extensive formal calculations could be replaced by a direct conclusion. Another advantage of the iteration method over the usual methods is that a detailed approximation calculus can be systematically carried through. For instance, the quantum mechanical expression for the electric polarizability can be calculated from the expression for the electric dipole moment, the expression for the hyperpolarizability from that for the polarizability etc. [9].

In the present paper, we discuss the iteration method [9] with reference to the properties of the terms of a Taylor series. The restriction on the spectrum of the unperturbed Hamiltonian to have no continuous part as was assumed in the earlier work can be dropped so long as the unperturbed energy eigenvalue considered belongs to the discrete spectrum; furthermore, the perturbed Hamiltonian may be a general analytic function of the perturbation parameter rather than a linear function. The formalism can easily be generalized to allow for multiple perturbation. We shall develop the formal solutions for the energies and corresponding eigenvectors in non-degenerate RS double perturbation theory by introducing two suitable operators corresponding to the two perturbations. In the Appendix, the relations which define the operator $\not{p}$ are summarized as well as some of the properties of this operator.

2. Formal Treatment of RS Single Perturbation Theory

The RS perturbation theory is concerned with the determination of the isolated eigenvalues E_s and eigenvectors $|s\rangle$ of a Hamiltonian $\mathscr{H}=\sum_n \lambda^n \mathscr{H}^{(n)}$ which, for $\lambda\to 0$, tend to known isolated solutions $E_s^{(0)}$ and $|s^0\rangle$, respectively, of the unperturbed Hamiltonian $\mathscr{H}^{(0)}$; the $\mathscr{H}^{(n)}$ are Hermitian operators and λ is a real parameter. It is assumed that the structure of the operators $\mathscr{H}^{(n)}$ permits the energies E_s and the vectors $|s\rangle$ to be expressed also as power series in λ, so that

$$\mathscr{H}=\sum_n \lambda^n \mathscr{H}^{(n)}, \qquad E_s=\sum_n \lambda^n E_s^{(n)}, \qquad |s\rangle=\sum_n \lambda^n |s^n\rangle. \tag{1}$$

The RS procedure, as usually presented, consists in substituting these series into the equation $\mathscr{H}|s\rangle = E_s\,|s\rangle$, in solving recursively the set of perturbation equations obtained, and in expressing the n'th-order solutions $E_s^{(n)}$ and $|s^n\rangle$ in terms of the operators $\mathscr{H}^{(j)}$, $j=0, 1 \ldots n$, and of unperturbed quantities [3].

2.1. First-Order Iteration Procedure

Writing now the power series (1) in the form

$$f(\lambda)=\sum_n \lambda^n f_n(\lambda)\Big|_{\lambda=0}=\sum_n \lambda^n f^{(n)} \qquad \text{with} \qquad f_n(\lambda)=\frac{1}{n!}\frac{\partial^n f(\lambda)}{\partial \lambda^n},\quad f_n(0)=f^{(n)},$$

we arrive at

$$\frac{\partial}{\partial\lambda} f_n(\lambda)\Big|_{\lambda=0}=(n+1)f^{(n+1)}. \tag{2a}$$

This formula forms the basis for our discussion of the RS perturbation theory. According to (2a), the first derivative of a product $\prod_i g_i^{(n_i)}$ of n_i'th-order perturbation terms $g_i^{(n_i)}$ is obtained by substituting for all i the sum $g_i^{(n_i)}+\lambda(n_i+1)g_i^{(n_i+1)}$ for $g_i^{(n_i)}$, collecting the terms of the first order in λ, and dropping λ; for example, we have

$$\begin{aligned} &f^{(n)}=g_1^{(n_1)}g_2^{(n_2)} \quad (n_1+n_2=n)\\ &\Rightarrow (n+1)f^{(n+1)}=(n_1+1)g_1^{(n_1+1)}g_2^{(n_2)}+(n_2+1)g_1^{(n_1)}g_2^{(n_2+1)}. \end{aligned} \tag{2b}$$

Thus, the transformation from $f^{(n)}$ to $f^{(n+1)}$ corresponds to using $f^{(n)}$ as a "zeroth-order" term and to calculating the "first-order" perturbation correction to it. Thereby, the first-order perturbation correction to a sum is the sum of the corresponding corrections of the summands and the first-order correction to a product is the sum of those products, each of which coincides with the given product except for one factor which is replaced by its first-order correction.

Let $f^{(1)}$ be equal to the RS expression for a solution $|s^1\rangle$; the factors $g_i^{(n_i)}$ then are unperturbed vectors $|s^0\rangle$, $\langle s^0|$, the operators $\mathscr{H}^{(1)}$ and $(\mathscr{H}^{(0)}-E_s^{(0)})^{-1}$ etc.; the corrections $g_i^{(n_i+1)}$ are again expressible in such terms and the operator $\mathscr{H}^{(2)}$. Thus, the repeated treatment of $|s^1\rangle$ according to the simple procedure (2b) permits the construction of the explicit expressions for particular n'th-order solutions $|s^n\rangle$. Given an expression for $|s^1\rangle$, the perturbation energies $E_s^{(n)}$ can be computed correspondingly from $E_s^{(0)}=\langle s^0|\mathscr{H}^{(0)}|s^0\rangle$.

The procedure is applicable only if the perturbation corrections to the number 1 and to the unity operator are defined to vanish and if, on the other hand, all zero terms are included which are determined by the perturbation equations and have non-vanishing corrections. Such terms appear in the degenerate case; for example, the difference $E_{s_\rho}^{(0)}-E_{s_\sigma}^{(0)}$ is equal to zero for $E_{s_\rho}^{(0)}=E_{s_\sigma}^{(0)}=E_s^{(0)}$, but $E_{s_\rho}^{(n)}-E_{s_\sigma}^{(n)}$ need not vanish. Therefore, the method provides no advantages for the determination of the eigenvalues and eigenvectors in degenerate RS perturbation theory.

If, however, $E_s^{(0)}$ is a non-degenerate eigenvalue, the first-order expression $|s^1\rangle = \mathcal{Q}_s^{(0)} \mathcal{H}^{(1)} |s^0\rangle$ is determined by just the first-order Schrödinger perturbation equation; no zero terms whose corrections could contribute are provided by the perturbation equations. For constructing the perturbation vectors $|s^n\rangle$ in arbitrarily high order by our procedure, it is thus sufficient to know this explicit expression for $|s^1\rangle$, the operators $\mathcal{H}^{(j)}$, and the first-order correction to the reduced resolvent

$$\mathcal{Q}_s^{(0)} = (1 - |s^0\rangle\langle s^0|) \frac{1}{E_s^{(0)} - \mathcal{H}^{(0)}} (1 - |s^0\rangle\langle s^0|). \tag{3a}$$

$\mathcal{Q}_s^{(1)}$ is evaluated again according to (2b). Using the equation $1 = [A^{-1}]^{(0)} A^{(0)}$ which gives $[A^{-1}]^{(1)} = -[A^{-1}]^{(0)} A^{(1)} [A^{-1}]^{(0)}$, one finds

$$\mathcal{Q}_s^{(1)} = \mathcal{Q}_s^{(0)} (\mathcal{H}^{(1)} - E_s^{(1)}) \mathcal{Q}_s^{(0)} - \mathcal{Q}_s^{(0)} \mathcal{Q}_s^{(0)} \mathcal{H}^{(1)} |s^0\rangle\langle s^0| - |s^0\rangle\langle s^0| \mathcal{H}^{(1)} \mathcal{Q}_s^{(0)} \mathcal{Q}_s^{(0)}. \tag{3b}$$

If the spectrum of $\mathcal{H}^{(0)}$ has no continuous part as was assumed in [9], the operator $\sum_{t(\neq s)} |t^0\rangle\langle t^0|$ could have been used instead of $1 - |s^0\rangle\langle s^0|$.

For example, the explicit expression for the second-order eigenvector $|s^2\rangle$ is determined from

$$|s^1\rangle = \mathcal{Q}_s^{(0)} \mathcal{H}^{(1)} |s^0\rangle$$

as follows

$$|s^2\rangle = \tfrac{1}{2} \{ \mathcal{Q}_s^{(1)} \mathcal{H}^{(1)} |s^0\rangle + 2 \mathcal{Q}_s^{(0)} \mathcal{H}^{(2)} |s^0\rangle + \mathcal{Q}_s^{(0)} \mathcal{H}^{(1)} |s^1\rangle \}.$$

This vector satisfies the condition of normalization of $|s\rangle$ to unity in second order and fulfils the equation $\mathrm{Im}\langle s^0 \mid s^2\rangle = 0$. The example gives rise to consider the normalization and the specific structure of the eigenvector computed by means of our first-order iteration method.

1) The perturbed eigenvector $|s\rangle = \sum_n \lambda^n |s^n\rangle$ is normalized to unity if $|s^0\rangle$ is normalized to unity and if the n'th-order term $|s^n\rangle$, $n \geq 2$, is constructed by means of the repeated transformation of $|s^1\rangle$ according to (2b). This follows from the fact that the normalization condition for the n'th order, differentiated according to Eq. (2b), is transformed into that for the $(n+1)$'th order:

$$\sum_k^{0,n} \langle s^k \mid s^{n-k}\rangle = \delta_{0n} \longrightarrow (n+1) \sum_k^{0,n+1} \langle s^k \mid s^{n+1-k}\rangle = 0.$$

The procedure, therefore, does not permit the use of the intermediate normalization condition sometimes adopted which means that $|s^n\rangle$ for all $n \geq 1$ is taken to be orthogonal to $|s^0\rangle$. In case two unperturbed vectors are orthogonal, the corresponding perturbed vectors are also orthogonal.

2) The expression for the n'th-order vector $|s^n\rangle$, $n \geq 2$, including its phase terms, is determined uniquely by the iteration method starting from $|s^1\rangle = \mathcal{Q}_s^{(0)} \mathcal{H}^{(1)} |s^0\rangle$. Since $|s^1\rangle$ is chosen to have the property

$$\mathrm{Im}\langle s^0 \mid s^1\rangle = 0, \tag{4a}$$

the higher-order terms $\mathrm{Im}\langle s^0|s^n\rangle$ are also fixed by the procedure. We find, from (2a) and (2b),

$$\mathrm{Im}\langle s^0 \mid s^2\rangle = 0, \qquad \mathrm{Im}\langle s^0 \mid s^3\rangle = -\tfrac{1}{3}\mathrm{Im}\langle s^1 \mid s^2\rangle,$$
$$\mathrm{Im}\langle s^0 \mid s^4\rangle = -\tfrac{1}{2}\mathrm{Im}\langle s^1 \mid s^3\rangle, \ldots \tag{4b}$$

The vectors thus obtained have the form

$$\begin{aligned} |s^1\rangle &= |\tilde{s}^1\rangle, \\ |s^2\rangle &= |\tilde{s}^2\rangle, \\ |s^3\rangle &= |\tilde{s}^3\rangle - \tfrac{i}{3}\mathrm{Im}\langle s^1 \mid s^2\rangle |s^0\rangle, \\ |s^4\rangle &= |\tilde{s}^4\rangle - \tfrac{i}{3}\mathrm{Im}\langle s^1 \mid s^2\rangle |\tilde{s}^1\rangle - \tfrac{i}{2}\mathrm{Im}\langle s^1 \mid s^3\rangle |s^0\rangle, \ldots \end{aligned} \tag{4c}$$

The vectors $|\tilde{s}^n\rangle$ are the conventional n'th-order solutions for the RS eigenvector $|\tilde{s}\rangle = \sum_n \lambda^n |\tilde{s}^n\rangle$ satisfying the conditions of normalization to unity for the various orders as well as the equations $\mathrm{Im}\langle s^0 \mid \tilde{s}^n\rangle = 0$. As can be seen from the Eqs. (4c), our procedure, in general, does not take $|\tilde{s}^n\rangle$ into $(n+1)|\tilde{s}^{n+1}\rangle$.

The specific structure of the vectors $|s^n\rangle$ in Eq. (4c) results from the following fact. In contrast to the solution for the energy, the solution for a corresponding eigenvector is not uniquely determined by the perturbation equations since an eigenvector is only defined up to a multiplicative constant. Thus, besides the solutions $|\tilde{s}^n\rangle$, the linear combinations

$$\lambda^n |s^n\rangle = \lambda^n \left(|\tilde{s}^n\rangle + \sum_k^{1,n} a^{(k)} |\tilde{s}^{n-k}\rangle \right), \quad n = 1, 2, \ldots N \tag{5a}$$

with arbitrary coefficients $a^{(k)}$ also satisfy the perturbation equations up to and including the N'th order; we have specified $|s^0\rangle = |\tilde{s}^0\rangle$. If we then require the perturbation solutions $|s\rangle$ and $|\tilde{s}\rangle$ for the eigenvector to be normalized to unity, the terms $a^{(n)}$ turn out to have fixed real parts, but arbitrary imaginary parts expressing the fact that the perturbation equations allow a n'th-order eigenvector to have an imaginary multiple of the unperturbed eigenvector added to it. The multiplicative constant is reduced to a phase factor,

$$1 + \sum_n^{1,\infty} \lambda^n a^{(n)} = \exp(\mathrm{i}\alpha), \qquad \alpha = \sum_n^{1,\infty} \lambda^n \alpha^{(n)} \tag{5b}$$

with real $\alpha^{(n)}$'s; the various normalized solutions differ from one another as do $|\tilde{s}\rangle$ and $|s\rangle = |\tilde{s}\rangle \exp(\mathrm{i}\alpha)$. Let the solutions $|\tilde{s}^n\rangle$ satisfy $\mathrm{Im}\langle s^0 \mid \tilde{s}^n\rangle = 0$ for all n, then, because of Eq. (5a), the imaginary part of $a^{(n)}$ has the form

$$\begin{aligned} \mathrm{Im}\, a^{(1)} &= \mathrm{Im}\langle s^0 \mid s^1\rangle, \\ \mathrm{Im}\, a^{(n)} &= \mathrm{Im}\langle s^0 \mid s^n\rangle - \sum_k^{1,n-1} (\mathrm{Im}\, a^{(k)}) \langle s^0 \mid \tilde{s}^{n-k}\rangle, \quad n \geqslant 2. \end{aligned} \tag{5c}$$

Given the eigenvector $|s\rangle$ through the $(n-1)$'th order as well as the term $\mathrm{Im}\langle s^0 \mid s^n\rangle$, according to Eq. (4a) and (4b), respectively, we can calculate $\mathrm{Im}\, a^{(n)}$ and, from Eq. (5b), the terms $\alpha^{(n)}$ and $\mathrm{Re}\, a^{(n)}$. Thus, Eq. (5a) is reduced to Eq. (4c).

Since the solutions $|s\rangle$ and $|\tilde{s}\rangle$ of (4c) differ only in a phase factor, we have

$$\sum_{k}^{0,n} \langle s^k| \ldots |s^{n-k}\rangle = \sum_{k}^{0,n} \langle \tilde{s}^k| \ldots |\tilde{s}^{n-k}\rangle.$$

The perturbation energies $E_s^{(n)}$, therefore, are computable by the new scheme from $\langle s^0|\mathscr{H}^{(0)}|s^0\rangle$ with the aid of the vectors $|\tilde{s}^k\rangle$, as expected.

Thus, the explicit RS perturbation expressions can be constructed systematically by the first-order iteration procedure, according to (2b). For the derivation of increasingly higher-order expressions, the extent of the corresponding calculations increases rapidly and the method soon becomes unpractical. Its essential advantage, however, is that it reflects the connection of the RS perturbation series terms in n'th order and $(n+1)$'th order. A formalization of the procedure is achieved by the use of a suitably defined operator.

2.2. *The Operator* $\not{p}$

In [9], the operator $\not{p}$ was introduced to formulate the first-order iteration procedure in a simple and compact form. $\not{p}$ is defined to transform $|s^0\rangle$ into $\mathscr{Q}_s^{(0)}\mathscr{H}^{(1)}|s^0\rangle$, $\mathscr{H}^{(n)}$ into $(n+1)\mathscr{H}^{(n+1)}$ and to act on a function of such quantities as a differential operator; its formal properties are summarized in the Appendix. As is shown there, the total Hamiltonian can be put in the form

$$\mathscr{H} = \exp(\lambda \not{p}) \mathscr{H}^{(0)} \exp(-\lambda \not{p}). \tag{6a}$$

In the non-degenerate case, the RS expressions for the n'th-order terms $|s^n\rangle$ and $E_s^{(n)}$ can be constructed by the repeated application of $\not{p}$ to the unperturbed terms

$$|s^n\rangle = \frac{1}{n!} \not{p}^n |s^0\rangle, \qquad E_s^{(n)} = \frac{1}{n!} \not{p}^n E_s^{(0)}; \tag{6b}$$

it follows that

$$|s\rangle = \exp(\lambda \not{p})|s^0\rangle, \qquad E_s = \exp(\lambda \not{p}) E_s^{(0)}. \tag{6c}$$

If $E_s^{(0)}$ is an eigenvalue of $\mathscr{H}^{(0)}$ with finite degeneracy, it was observed [9] that the operator $\not{p}$ can be used to construct the RS projector $\mathscr{P}_s = \sum_n \lambda^n \mathscr{P}_s^{(n)}$ whose zeroth order $\mathscr{P}_s^{(0)}$ is the projector onto the subspace of $\mathscr{H}^{(0)}$ corresponding to $E_s^{(0)}$. We have

$$\mathscr{P}_s^{(n)} = \frac{1}{n!} [\not{p}, \mathscr{P}_s^{(0)}]_n, \qquad \mathscr{P}_s = \exp(\lambda \not{p}) \mathscr{P}_s^{(0)} \exp(-\lambda \not{p}). \tag{6d}$$

The formalism can be extended to similar quantum mechanical expressions which are invariant in each order under rotations of the basis in degenerate subspaces. For example, the perturbation corrections to $\mathrm{tr}\,(\mathscr{P}_s^{(0)}\mathscr{F}^{(0)})$ can be calculated by means of $\not{p}$ if the commutators $[\not{p}, \mathscr{F}^{(0)}]_n$ of $\not{p}$ with the operator $\mathscr{F}^{(0)}$ are known.

3. Formal Treatment of RS Double Perturbation Theory

The formalism presented for the RS single perturbation theory can easily be extended to the case of the double perturbation theory where

$$\mathscr{H}=\mathscr{H}^{(0)}+\sum_{m}^{1,\infty}\lambda_1^m\mathscr{H}_1^{(m)}+\sum_{n}^{1,\infty}\lambda_2^n\mathscr{H}_2^{(n)} \tag{7a}$$

with λ_1, λ_2 being different perturbation parameters and the $\mathscr{H}_i^{(p)}$ being Hermitian operators. We shall discuss the double perturbation corrections to a non-degenerate eigenvalue $E_s^{(00)}$ and the corresponding eigenvector $|s^{00}\rangle$ of $\mathscr{H}^{(0)}$; the perturbed eigenvalue E_s and the perturbed eigenvector $|s\rangle$ are assumed to be expansible in double power series in the parameters λ_1 and λ_2:

$$E_s=\sum_{m,n}\lambda_1^m\lambda_2^n E_s^{(mn)}, \qquad |s\rangle=\sum_{m,n}\lambda_1^m\lambda_2^n|s^{mn}\rangle. \tag{7b}$$

If $E_s^{(00)}$ is degenerate, such an expansion is, in general, not valid. By inserting the series expansions (7a, b) into the Schrödinger equation, one gets a system of perturbation equations which is usually solved as in the case of the single perturbation [3].

We now introduce two left operators $\not{p}_1, \not{p}_2$. In addition to their definitions which are equivalent to those of the operator $\not{p}$ in the case of a single perturbation (see Appendix), we need a definition of their sum and product as well as of the application of $\not{p}_i$ to $\mathscr{H}_j^{(n)}$ with $i\neq j$; obviously, we define

$$\not{p}_i|s^{00}\rangle=\mathscr{Q}_s^{(00)}\mathscr{H}_i^{(1)}|s^{00}\rangle, \qquad \mathscr{Q}_s^{(00)}=\frac{1-|s^{00}\rangle\langle s^{00}|}{E_s^{(00)}-\mathscr{H}^{(0)}};$$

$$[\not{p}_i, \mathscr{H}_j^{(n)}]=0 \quad \text{if} \quad i\neq j; \tag{8}$$

$$(\not{p}_i+\not{p}_j)f=\not{p}_i f+\not{p}_j f, \qquad \not{p}_i\not{p}_j f=\not{p}_i(\not{p}_j f).$$

Thus, the operators $\not{p}_1$ and $\not{p}_2$ can repeatedly be applied to $\mathscr{H}^{(0)}$, $|s^{00}\rangle$ and $E_s^{(00)}$. Using Eq. (6a) and the relation

$$[\not{p}_i, [\not{p}_j, \mathscr{H}^{(0)}]]=0 \quad \text{if} \quad i\neq j,$$

we can put the Hamiltonian in the form

$$\begin{aligned}\mathscr{H}&=\exp(\lambda_1\not{p}_1)\exp(\lambda_2\not{p}_2)\mathscr{H}^{(0)}\exp(-\lambda_2\not{p}_2)\exp(-\lambda_1\not{p}_1)\\&=\exp(\lambda_2\not{p}_2)\exp(\lambda_1\not{p}_1)\mathscr{H}^{(0)}\exp(-\lambda_1\not{p}_1)\exp(-\lambda_2\not{p}_2)\\&=\exp(\lambda_1\not{p}_1+\lambda_2\not{p}_2)\mathscr{H}^{(0)}\exp(-\lambda_1\not{p}_1-\lambda_2\not{p}_2).\end{aligned} \tag{9}$$

To construct formal solutions for $|s^{mn}\rangle$ by the repeated application of the operators $\not{p}_1, \not{p}_2$ to $|s^{00}\rangle$, we choose the first-order solutions as

$$|s^{10}\rangle=\not{p}_1|s^{00}\rangle, \qquad |s^{01}\rangle=\not{p}_2|s^{00}\rangle.$$

We consider the (1, 1)'th-order solutions

$$|s^{11}\rangle_{(12)}=\not{p}_1\not{p}_2|s^{00}\rangle, \qquad |s^{11}\rangle_{(21)}=\not{p}_2\not{p}_1|s^{00}\rangle;$$

they differ in phase terms: using the equations $\mathrm{Im}\langle s^{00}|\not{p}_i s^{00}\rangle = 0$, $i=1,2$, we have for instance,

$$\not{p}_1 \mathrm{Im}\langle s^{00} \mid s^{01}\rangle = \mathrm{Im}\langle s^{10} \mid s^{01}\rangle + \mathrm{Im}\langle s^{00} \mid s^{11}\rangle_{(12)} = 0$$

and get

$$[\not{p}_1, \not{p}_2]|s^{00}\rangle = i\alpha_s^{(11)}|s^{00}\rangle, \qquad \alpha_s^{(11)} = 2\ \mathrm{Im}\langle s^{01} \mid {}^{10}\rangle.$$

The terms $\not{p}_1^m \not{p}_2^n |s^{00}\rangle$, therefore, are not invariant under distinguishable permutations of the $\not{p}_i$'s. The expression for a (m, n)'th-order eigenvector $|s^{mn}\rangle$ is a sum of products of factors each of which, besides of $|s^{00}\rangle$, remains unchanged if $[\not{p}_1, \not{p}_2]$ is applied. Thus, using the equation

$$[\not{p}_1, \not{p}_2] f \circ g = ([\not{p}_1, \not{p}_2] f) \circ g + f \circ [\not{p}_1, \not{p}_2] g,$$

we find

$$[\not{p}_1, \not{p}_2]|s^{mn}\rangle = i\alpha_s^{(11)}|s^{mn}\rangle. \tag{10a}$$

Equations of this type are not to be confused with equations such as $\partial^2|s\rangle/\partial\lambda_1\partial\lambda_2 = \partial^2|s\rangle/\partial\lambda_2\partial\lambda_1$ for $\lambda_1 = \lambda_2 = 0$.

The terms $|s^{mn}\rangle$ satisfy the set of perturbation equations obtained from $\mathscr{H}|s\rangle = E_s|s\rangle$; thus, only those sequences of operators of the type $\not{p}_1^m \not{p}_2^n$ applied to $|s^{00}\rangle$ can generate the vectors $|s^{mn}\rangle$ which, when applied to $\mathscr{H}^{(0)}|s^{00}\rangle = E_s^{(00)}|s^{00}\rangle$, formally yield these perturbation equations. The following sequences of operators fulfil this condition:

$$\frac{1}{m!n!}\not{p}_1^m \not{p}_2^n, \qquad \frac{1}{m!n!}\not{p}_2^n \not{p}_1^m, \qquad \frac{1}{(m+n)!}\sum_p \not{p}_1^m \not{p}_2^n, \quad m, n = 0, 1, 2, \ldots$$

In the first product, all the $\not{p}_1$ operators stand to the left of the $\not{p}_2$ operators; this order is changed in the second product; in the third term, the summation is to be extended over all distinguishable permutations of the $\not{p}_1$ and $\not{p}_2$ operators. In this way, different sequences of vectors $|s^{mn}\rangle$ are defined,

$$\begin{aligned}
|s^{mn}\rangle_{(12)} &= \frac{1}{m!n!}\not{p}_1^m \not{p}_2^n |s^{00}\rangle, \\
|s^{mn}\rangle_{(21)} &= \frac{1}{m!n!}\not{p}_2^n \not{p}_1^m |s^{00}\rangle, \\
|s^{mn}\rangle_{\Sigma} &= \frac{1}{(m+n)!}\sum_p \not{p}_1^m \not{p}_2^n |s^{00}\rangle,
\end{aligned} \tag{10b}$$

leading to perturbed eigenvectors $|s\rangle$ which differ in their phase factors only:

$$\begin{aligned}
|s\rangle_{(12\varepsilon)} &= \exp(\lambda_1 \not{p}_1) \exp(\lambda_2 \not{p}_2)|s^{00}\rangle, \\
|s\rangle_{(21)} &= \exp(\lambda_2 \not{p}_2) \exp(\lambda_1 \not{p}_1)|s^{00}\rangle, \\
|s\rangle_{\Sigma} &= \exp(\lambda_1 \not{p}_1 + \lambda_2 \not{p}_2)|s^{00}\rangle,
\end{aligned} \tag{10c}$$

where

$$|s\rangle_{(12)} = \exp(i\alpha_s^I)|s\rangle_{(21)} = \exp(i\alpha_s^{II})|s\rangle_\Sigma .$$

These vectors are normalized to unity if $|s^{00}\rangle$ is normalized to unity. The order of sequence of the factors in the operator $\exp(\lambda_i \not{P}_i)\exp(\lambda_j \not{P}_j)$, $i \neq j$, indicates that the perturbation $\sum_n \lambda_j^n \mathscr{H}_j^{(n)}$ in all orders is to be considered before $\sum_n \lambda_i^n \mathscr{H}_i^{(n)}$.

As the phases in $\not{P}_1|s^{00}\rangle$ and $\not{P}_2|s^{00}\rangle$ are fixed, the different phase factors $\exp(i\alpha_s^I)$ and $\exp(i\alpha_s^{II})$ are uniquely determined. Using the Campbell–Hausdorff formula

$$\exp(A)\exp(B) = \exp(A+B+\tfrac{1}{2}[A,B]+\tfrac{1}{12}[[A,B],B]-\tfrac{1}{12}[[A,B],A]+\cdots),$$

we have

$$\begin{aligned}|s\rangle_{(12)} &= \exp(\lambda_1\lambda_2[\not{P}_1,\not{P}_2]+\tfrac{1}{2}\lambda_1^2\lambda_2[\not{P}_1,[\not{P}_1,\not{P}_2]]+\cdots)|s\rangle_{(21)}\\ &= \exp(\tfrac{1}{2}\lambda_1\lambda_2[\not{P}_1,\not{P}_2]+\tfrac{1}{3}\lambda_1^2\lambda_2[\not{P}_1,[\not{P}_1,\not{P}_2]]+\cdots)|s\rangle_\Sigma ;\end{aligned}$$

from (10a), the power series for α_s^I and α_s^{II} could explicitly be calculated.

Since the perturbation energies $E_s^{(mn)}$ are uniquely determined by the perturbation equations and do not depend on phases, the term $\not{P}_1^m \not{P}_2^n E_s^{(00)}$ is equal to $\not{P}_2^n \not{P}_1^m E_s^{(00)}$ for each pair of indices (m, n) and, therefore, independent of the order of sequence in which the operators $\not{P}_1, \not{P}_2$ are taken into account; hence

$$[\not{P}_1, \not{P}_2]E_s^{(mn)} = 0. \tag{11a}$$

Thus, the explicit expressions for the (m, n)'th-order energies $E_s^{(mn)}$ can be constructed as follows:

$$E_s^{(mn)} = \frac{1}{m!n!}\mathscr{P}\not{P}_1^m \not{P}_2^n E_s^{(00)} = \begin{cases} \dfrac{1}{m}\not{P}_1 E_s^{(m-1,n)} & \text{for} \quad m \geqslant 1 \\ \dfrac{1}{n}\not{P}_2 E_s^{(m,n-1)} & \text{for} \quad n \geqslant 1, \end{cases} \tag{11b}$$

where $\mathscr{P}\not{P}_1^m \not{P}_2^n$ denotes any permutation of the operators $\not{P}_1, \not{P}_2$. We can write the perturbed eigenvalue E_s in the form

$$\begin{aligned}E_s &= \exp(\lambda_1 \not{P}_1)\exp(\lambda_2 \not{P}_2)E_s^{(00)} = \exp(\lambda_2 \not{P}_2)\exp(\lambda_1 \not{P}_1)E_s^{(00)}\\ &= \exp(\lambda_1 \not{P}_1 + \lambda_2 \not{P}_2)E_s^{(00)}.\end{aligned} \tag{11c}$$

The Eq. (11b) presents a well-known interchange theorem. For $m, n \geqslant 1$, the application of $\not{P}_1$ to $(1/m)E_s^{(m-1,n)}$ can be replaced by the application of $\not{P}_2$ to $(1/n)E_s^{(m,n-1)}$.

Thus, in RS double perturbation theory, we have obtained formal solutions for an energy eigenvector and the corresponding eigenvalue whose zeroth order is non-degenerate. The double perturbation formulas for other quantum mechanical expressions such as projectors can be derived as in analogous cases of the single perturbation theory. The procedure can easily be extended to the case of more than two perturbations.

4. Conclusion

The formal RS perturbation theory for non-degenerate energy states presented for the case of a single perturbation has been generalized to allow for multiple perturbation. By the use of suitable operators which are defined to generate the first-order perturbation corrections with respect to the different perturbations, formal solutions of the RS perturbation equations for the terms of the power series expansions of energy eigenvalues and eigenvectors are obtained in a compact form. The explicit perturbation expressions are constructible systematically by the repeated application of the operators to the corresponding unperturbed terms. Thereby, the uniqueness of the RS energies implies their independence of the order of sequence in which the operators are applied to the zeroth-order eigenvalue; the eigenvector, however, depends on this sequence and the various solutions differ in phase terms.

The perturbation expressions formulated in this way provide insights into some of their interesting features as was illustrated by examples in [9] and [14] which indicate where this treatment of perturbation theory has advantages over the usual formulation. The formalism should be applied to further practical examples such as the calculation of multipole moments of compound systems. It seems to be useful to extend the investigations to diagrammatic methods of perturbation theory.

Acknowledgement. The financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

Appendix

We summarize the definition and some of the properties of the operator $\not{p}$ [9].

Let the Hamiltonian $\mathscr{H}$ be given as a power series $\sum_n \lambda^n \mathscr{H}^{(n)}$ and $E_s^{(0)}$ be an isolated eigenvalue of $\mathscr{H}^{(0)}$. The left operator $\not{p}$ is then defined as follows:

a) $\not{p}|s^0\rangle = \mathscr{Q}_s^{(0)} \mathscr{H}^{(1)} |s^0\rangle, \qquad \not{p}\langle s^0| = \langle s^0| \mathscr{H}^{(1)} \mathscr{Q}_s^{(0)}, \qquad \mathscr{Q}_s^{(0)} = \dfrac{1 - |s^0\rangle\langle s^0|}{E_s^{(0)} - \mathscr{H}^{(0)}};$

b) $(\not{p}\mathscr{H}^{(n)}) = (n+1)\mathscr{H}^{(n+1)};$

c) $\not{p}(f+g) = \not{p}f + \not{p}g;$

d) $\not{p}f \circ g = (\not{p}f) \circ g + f \circ \not{p}g;$

e) $\not{p}^n f = \not{p}^m (\not{p}^{n-m} f), \qquad \not{p}^0 f = f;$

f) $\not{p}^n f = 0, \; n \geqslant 1, \quad$ if f remains uncorrected, e.g. $(\not{p}^n 1) = 0.$

With $f \circ g$ we denote any product of f and g; the notation $(\not{p}f) \ldots$ indicates that $\not{p}$ is applied to f only and not to factors on the right side of the brackets; for the sake of consistency, the number 1 and the unity operator belong to the coefficient domain of $\not{p}$ as well as zero terms which are not determined by the perturbation procedure.

If the application of $\not{p}$ to the operator $\mathscr{A}$ is defined, its application to the inverse operator $\mathscr{A}^{-1}$ is obtained from

$$0=\not{p}\mathscr{A}^{-1}\mathscr{A}=(\not{p}\mathscr{A}^{-1})\mathscr{A}+\mathscr{A}^{-1}\not{p}\mathscr{A}.$$

In this way, the application of $\not{p}$ to $\mathscr{Q}_s^{(0)}$ and, consequently, to the RS expressions for $E_s^{(n)}$ an) $|s^n\rangle$ is defined in arbitrarily high order.

Introducing the sum

$$\mathscr{S}_\lambda^{(n)}=\sum_k^{0,n}\frac{\lambda^k}{k!}\not{p}^k$$

as in the discussion of properties of derivations [15] and, assuming $\not{p}$ formally to be nilpotent to the $(n+1)$'th power, i.e. $\not{p}^{n+1}=0$, we find

$$\mathscr{S}_\lambda^{(n)}f\circ g=(\mathscr{S}_\lambda^{(n)}f)\circ\mathscr{S}_\lambda^{(n)}g \quad \text{if} \quad \not{p}^{n+1}=0$$

and for $n\to\infty$

$$\exp(\lambda\not{p})f\circ g=(\exp(\lambda\not{p})f)\circ\exp(\lambda\not{p})g.$$

Because $\not{p}^n 1=0$, $n\geqslant 1$, we have the result that the operator $\mathscr{S}_\lambda^{(n)}$ transforms unperturbed orthonormal eigenvectors with isolated eigenvalues into vectors which are orthonormal up to and including the n'th order.

The application of $\not{p}$ to an operator $\mathscr{O}$ is described by the commutator $[\not{p},\mathscr{O}]$ as can be seen by the definitions c) and d). The definition e) now reads

$$(\not{p}^n\mathscr{O})=[\not{p},\mathscr{O}]_n \quad \text{with} \quad [\not{p},\mathscr{O}]_0=\mathscr{O},\ [\not{p},\mathscr{O}]_n=[\not{p},[\not{p},\mathscr{O}]_{n-1}]$$

and the definitions b) can be put in the form

$$[\not{p},\mathscr{H}^{(0)}]_n=n!\mathscr{H}^{(n)}.$$

For the Hamiltonian we thus get the representation

$$\mathscr{H}=\sum_n\frac{\lambda^n}{n!}[\not{p},\mathscr{H}^{(0)}]_n=\exp(\lambda\not{p})\mathscr{H}^{(0)}\exp(-\lambda\not{p}).$$

If $E_s^{(0)}$ is a degenerate eigenvalue, the operator $\not{p}$, in contrast to the non-degenerate case, does not take a "correct" zeroth-order eigenvector $|s_\rho^0\rangle$ into the corresponding first-order expression as follows from the definition a). As discussed in [9], the operator $\not{p}$, however, takes the projector $\mathscr{P}_s^{(0)}$ onto the subspace of $\mathscr{H}^{(0)}$ corresponding to $E_s^{(0)}$ into the first-order term $\mathscr{P}_s^{(1)}$ of the perturbation series for $\mathscr{P}_s=\sum_n\lambda^n\mathscr{P}_s^{(n)}$; furthermore, $\not{p}$ transforms $\mathscr{P}_s^{(n)}$ into $(n+1)\mathscr{P}_s^{(n+1)}$ provided that the n'th-order term is written in the explicitly invariant form. This condition takes into account the difference between $\not{p}|s_\rho^0\rangle$ and $|s_\rho^1\rangle$ and means that a term of the type

$$\sum_\rho|s_\rho^0\rangle E_{s_\rho}^{(1)}\langle s_\rho^0|=\sum_\rho|s_\rho^0\rangle\langle s_\rho^0|\mathscr{H}^{(1)}|s_\rho^0\rangle\langle s_\rho^0|$$

has to be replaced by

$$\sum_{\rho,\sigma}|s_\rho^0\rangle\langle s_\rho^0|\mathscr{H}^{(1)}|s_\sigma^0\rangle\langle s_\sigma^0|.$$

As $\langle s_\rho^0 | \mathscr{H}^{(1)} | s_\sigma^0 \rangle = 0$ for $\sigma \neq \rho$, both expressions differ in additive zero terms only, the perturbation corrections of which, in general, do not vanish.

References

1. Kemble, E. C.: The fundamental principles of quantum mechanics, chap. 11. New York: Dover Publ. 1958
2. Wigner, E. P.: Group theory and its application to the quantum mechanics of atomic spectra, chap. 5. New York–London: Academic Press 1959
3. Hirschfelder, J. O., Byers Brown, W., Epstein, S. T.: Advan. Quantum Chem. **1**, 255 (1964)
4. Löwdin, P.-O.: J. Math. Phys. **3**, 969 (1962)
5. Hirschfelder, J. O.: Intern. J. Quantum Chem. **3**, 731 (1969)
6. Silverstone, H. J.: J. Chem. Phys. **54**, 2325 (1971)
7. Messiah, A.: Quantum mechanics, Vol. II, chap. 16. Amsterdam: North-Holland 1965
8. Silverstone, H. J., Holloway, T. T.: J. Chem. Phys. **52**, 1472 (1970)
9. Haase, D., Ruch, E.: Theoret. Chim. Acta (Berl.) **29**, 235 (1973)
10. Primas, H.: Rev. Mod. Phys. **35**, 710 (1963)
11. Murray, F. J.: J. Math. Phys. **3**, 451 (1962)
12. Hirschfelder, J. O.: J. Chem. Phys. **39**, 2099 (1963)
13. Haase, D.: Fortschr. Phys. **24**, 37 (1976)
14. Haase, D., Ruch, E.: Theoret. Chim. Acta (Berl.) **29**, 189 (1973)
15. Jacobson, N.: Lie algebras, pp. 7–9. New York: Interscience Publ. 1966

Received September 23, 1977/March 6, 1978