Theor Chim Acta (1994) 88: 243–256

Theoretica Chimica Acta

Convergence studies in quantum perturbation theory

Philippe Durand[1], **Jiří Savrda**[1,*], **Ivana Paidarová**[2]

[1] Laboratoire de Physique Quantique, Unité de recherche Associée au CNRS n°505 IRSAMC, Université Paul Sabatier, 118 route de Narbonne F-31062 Toulouse Cedex, France

[2] J. Heyrovský Institute, Academy of Sciences of the Czech Republic, Dolejškova 3, 182 23 Prague 8, Czech Republic

Received 4 February 1993/Accepted 28 September 1993

Summary. Wave operator equations associated with the determination of energies are generally solved by perturbation methods [1]. However, it is well known that for most actual systems the standard Rayleigh–Schrödinger and Brillouin–Wigner series have slow convergence properties. We suggest that one way how to deal with this problem is to modify or to renormalize the standard wave equations. For that purpose we introduce new *derivative* and *convergence* operators associated with the basic Rayleigh–Schrödinger and Brillouin–Wigner formalisms. Since the direct use of these operators would imply difficult operator inversions, we investigate the efficiency of various approximations of the derivative operator. It is shown that these approximations can overcome convergence difficulties and also open the way to systematic derivations of infinite partial summation schemes. Our approach is also discussed with respect to the standard diagonalization procedure of Davidson. Three simple model systems are investigated numerically.

Key words: Perturbation theory – Convergence – Davidson method.

There is a large literature concerning perturbation theory in Quantum Mechanics [1]. The standard Rayleigh–Schrödinger and Brillouin–Wigner perturbation expansions are presented in most elementary text books on Quantum Mechanics. Starting from a zero-order unperturbed description, wavefunctions and energies are expanded by perturbation. It is well known that the convergence properties of these series are rather slow. For many actual applications the series diverge. Although powerful methods have been developed for improving the convergence properties, such as infinite summations and Padé approximants, a general algebraic scheme for systematically improving the perturbation expansions is still lacking. The situation is worse in many-body perturbation theory where diagrammatic expansions and infinite partial summations to all orders cannot be considered as the final development of the theory.

The aim of this paper is to bring sound foundations and new results on perturbation theory. This will be done by introducing new concepts of *derivative*

---

[*] On leave from J. Heyrovský Institute, Prague

and *convergence* operators which are able to predict and extend the convergence properties of the standard Rayleigh–Schrödinger and Brillouin–Wigner schemes. These operators will allow to generalize efficiently the standard perturbation schemes in a systematic way. New perturbation-iteration schemes will be presented and it will be shown that the exact inversion of these operators leads to quadratically converging methods. Obviously, this cannot be done for actual systems since the exact inversion of these operators would be as difficult as solving the Schrödinger equation. Nevertheless, the methods based on approximate derivative and convergence operators are useful and powerful. Some new wave operator equations and methods for their solution will be presented.

The paper is organized as follows. Notation and basic wave operator equations are introduced in section 1. The wave operator formalism will be used throughout this paper since it can be generalized easily to the theory of effective and intermediate Hamiltonians that will be investigated later. Wave operator equations modified by means of exact or approximate derivative operators will be presented in section 2. Finally, a numerical illustration will be given in section 3 for three simple model systems.

1 Wave operators equations (Rayleigh–Schrödinger and Brillouin–Wigner)

Let us consider the Schrödinger equation

$$H|\psi\rangle = E|\psi\rangle. \tag{1}$$

It is convenient to introduce a wave operator Ω which allows to pass from a zero-order unperturbed normalized state $|0\rangle$to the exact solution [2]

$$|\psi\rangle = \Omega|0\rangle, \quad \Omega \equiv \Omega P_0. \tag{2}$$

$P_0 = |0\rangle\langle 0|$ is the projector associated with the unperturbed state $|0\rangle$ and

$$P_0\Omega = P_0. \tag{3}$$

Hereafter, we will keep the intermediate normalization. For an extended study of stationary perturbation theory outside the intermediate normalization see reference [3]. The wave operator can be split into P_0 and the reduced wave operator X

$$\Omega = (P_0 + Q_0)\Omega = P_0 + X, \tag{4}$$

where $Q_0 = 1 - P_0$ is the projector associated with the orthogonal complement. Using (2) and (3) the Schrödinger equation (1) can be transformed into the two basic wave operator equations [4, 5]

$$H\Omega = \Omega H\Omega, \tag{5}$$

$$H\Omega = E\Omega, \tag{6}$$

where the exact energy is given by

$$E = \langle 0|H\Omega|0\rangle = \langle 0|H(1 + X)|0\rangle. \tag{7}$$

Expressions (5) and (6) are the starting points for obtaining the standard Rayleigh–Schrödinger (RS) and Brillouin–Wigner (BW) expansions. Hereafter it will be assumed that the Hamiltonian is partitioned into an unperturbed Hamiltonian H_0 and a perturbation V:

$$H = H_0 + V. \tag{8}$$

Multiplying both sides of (5) by Q_0 and P_0 and using (4) leads to

$$Q_0(1-X)H(1+X)P_0 = 0. \tag{9}$$

Rayleigh–Schrödinger expansion

Expression (9) can be transformed into the perturbation expression [4]

$$X = \frac{Q_0}{E_0 - H_0}(1-X)V(1+X)P_0, \tag{10}$$

where $E_0 = \langle 0|H_0|0\rangle$ is the zero-order unperturbed energy. The wave operator X is obtained in the RS approach by iterative solution of Eq. (10). Starting from $X = 0$ one easily obtains

$$X = \frac{Q_0}{E_0 - H_0} VP_0 + \frac{Q_0}{E_0 - H_0} V \frac{Q_0}{E_0 - H_0} VP_0 - \frac{Q_0}{(E_0 - H_0)^2} VP_0VP_0 + \cdots \tag{11}$$

Brillouin–Wigner expansion

Multiplying both sides of equation (6) by Q_0 and using (4) and (8) we get the operator equation

$$X = \frac{Q_0}{E - H_0} V(1+X)P_0. \tag{12}$$

As above, the perturbation expression for X is easily obtained

$$X = \frac{Q_0}{E - H_0} VP_0 + \frac{Q_0}{E - H_0} V \frac{Q_0}{E - H_0} VP_0 + \cdots \tag{13}$$

In the next section, the operator equations (10) and (12) will be modified to obtain better convergence properties by introducing *derivative* and *convergence* operators.

2 Modified wave operator equations

It is convenient to write the two basic Bloch-like equations (10) and (12) in the unique form

$$X = f(X), \tag{14}$$

where

$$f(X) = \frac{Q_0}{E_0 - H_0}(1-X)V(1+X)P_0 \tag{15}$$

in the RS approach and

$$f(X) = \frac{Q_0}{E - H_0} V(1+X)P_0 \tag{16}$$

for the BW scheme for which E is given by (7).

The iterative solution of equation (14) with $f(X)$ given by (15) and (16) provides the standard RS and BW perturbation expansions recalled in Sect. 1. It is well known that these methods often converge very slowly, even not linearly. Although expression (14) is rather formal, it has the advantage that, beyond the usual perturbation expansions reviewed in section 1, it suggests more powerful methods for determining X. In order to reduce the dependency of $f(X)$ on X the equation (14) will be modified. This can be done formally by introducing an operator A acting in the vectorial space of all operators coupling the subspace spanned by the unperturbed space $|0\rangle$ and its orthogonal complement. It is assumed that all the vectorial spaces are of finite dimension. Substracting the identity $AX = AX$ from both sides of (14) gives

$$(1 - A)X = f - AX \tag{17}$$

and

$$X = \frac{1}{1 - A}(f - AX). \tag{18}$$

Hereafter f or $f(X)$ we will use without distinction. The modified (renormalized) wave equation (18) is the fundamental equation of this paper. Up to now the operator A has not been defined. For reasons that will be given below we call it the *derivative* operator. It was stated only that it must reduce as much as possible the dependency on X of the right-hand side of Eq. (18). Many useful expressions of A are possible related to the Newton–Raphson scheme.

Newton–Raphson Derivative Operator

Since the iterative method of resolution of equation (14) which provides the RS and BW-like expansions converge slowly, at best linearly, the next step is to look for methods which converge quadratically or at least quasi-quadratically. For that purpose, Eq. (14) is linearized in the neighborhood of X

$$\begin{aligned} X + \Delta X &= f(X + \Delta X), \\ &= f(X) + A\Delta X + \cdots \end{aligned} \tag{19}$$

ΔX denotes a small variation of X, and A is a linear operator (superoperator) acting in the vectorial space of all operators coupling the subspace spanned by the unperturbed space $|0\rangle$ and its orthogonal complement. A generalizes the concept of derivative of an ordinary algebraic function $f(x)$. This operator so called derivative operator issuing from $f(X)$ will play a central role in investigation of the convergence properties of perturbation series and will provide more efficient schemes for solving Eq. (14). The two operators associated with the RS and BW formalisms can easily be derived from expressions (15) and (16). Some elementary algebraic calculations lead to

$$A = \frac{Q_0}{E_0 - H_0}[(1 - X)V - \langle 0|V(1 + X)|0\rangle]Q_0 \tag{20}$$

and

$$A = \frac{Q_0}{E - H_0}(1 - f)VQ_0 \tag{21}$$

for the RS and BW schemes, respectively. To our knowledge, the two above expressions do not seem to have been previously reported. Neglecting the higher order terms in (19) one gets

$$\Delta X = -\frac{1}{1-A}(X-f). \tag{22}$$

With (22) the equation (14) can be transformed

$$X = X + \Delta X = X - \frac{1}{1-A}(X-f) \tag{23}$$

and can also be put down in the form

$$X = \frac{1}{1-A}(f - AX)\,. \tag{24}$$

This expression (24) of X is of the form anticipated by (18), the derivative operator A being now defined by (20) and (21) for the RS and BW formalisms, respectively. Since in the solution of the wave operator equations, the derivative operator A plays the same role as the derivative $f'(x)$ associated with the Newton–Raphson solution of the ordinary algebraic equation $x = f(x)$, the convergence properties associated with the iterative solution of Eq. (24) can be discussed in terms of the eigenvalues of A. If their moduli $|\lambda|$ are small with respect to one, the RS and BW perturbation series will converge very fast. However, convergence difficulties can be expected for $|\lambda| \cong 1$. Note that if we stay within a perturbative approach, the exact derivative operators (20) and (21) are identical up to the first order

$$A = \frac{Q_0}{E_0 - H_0} V Q_0. \tag{25}$$

Eigenvalues and eigenvectors of A can be expressed by the equation

$$\frac{Q_0}{E_0 - H_0} V Q_0 |\phi\rangle = \lambda |\phi\rangle, \tag{26}$$

which can be transformed into

$$Q_0\left(H_0 + \frac{V}{\lambda}\right)Q_0 |\phi\rangle = E_0 |\phi\rangle. \tag{27}$$

It results from the above equations that considerable convergence difficulties will appear when $\lambda \cong 1$. In this case, Eq. (27) shows that eigenstates $|\phi\rangle$ have energies close to the energy of the exact solution $|\psi\rangle$. Thus the eigenvectors belonging to the outer space could strongly mix up with the unperturbed state $|0\rangle$. If E_0 is the lowest unperturbed eigenvalue of H_0, A can be transformed into a hermitian operator by a similiarity transformation

$$\frac{Q_0}{\sqrt{E_0 - H_0}} V \frac{Q_0}{\sqrt{E_0 - H_0}}. \tag{28}$$

The above operator has the same eigenvalues as A given by (25).

In the following, it is convenient to define a *convergence* operator by

$$C = \frac{1}{1 - A}. \tag{29}$$

As A, the convergence operator C is an exact operator (superoperator) acting in the vectorial space of all operators coupling the unperturbed state $|0\rangle$ to the outer (complementary) subspace.

Expressions (22) and (23) can be written as

$$\Delta X = -C(X - f), \tag{30}$$

$$X = X - C(X - f). \tag{31}$$

The iterative solution of Eqs. (30) and (31) defines a quadratically convergent Newton–Raphson procedure. If $X^{(n)}, f^{(n)}, A^{(n)}, C^{(n)}$ are the values of X, f, A, C obtained at the nth iteration, the new value of X at iteration $n + 1$ is given by

$$X^{(n+1)} = X^{(n)} - C^{(n)}(X^{(n)} - f^{(n)}) \tag{32}$$

and

$$X^{(n+1)} = \frac{1}{1 - A^{(n)}}(f^{(n)} - A^{(n)}X^{(n)}). \tag{33}$$

The iterative solution of (33) is not generally very useful since the inversion of $1 - A$ would be as difficult as solving the Schrödinger equation. Note that quadratic convergence can only be obtained if operators A are the exact derivative operators given by expressions (20) and (21) for the RS and BW formalisms, respectively. However, the equations (32) and (33) have a larger validity. They are exact equations in which A and C can be a priori any linear operator. For that reason, they will be good starting expressions for derivating approximate Newton–Raphson methods. There are many ways to define approximate Newton–Raphson methods. One may, for example, use simplified expressions of A that can easily be inverted. One can also use a polynomial approximation of the convergence operator while keeping exact expressions of A. This last approach will now be presented since it provides a hierarchy of procedures that generalize the standard diagonalization of Davidson.

Polynomial approximation of 1/1 − A

The convergence operator defined by (29) is approximated by a polynomial $C(A)$

$$\frac{1}{1 - A} \cong C(A) = \sum_{k=0}^{n} c_k A^k. \tag{34}$$

Hereafter we will use systematically the notation $C(A)$ for a polynomial approximation of the convergence operator C. Expressions (30) and (31) can now be rewritten as

$$\Delta X = -C(A)(X - f), \tag{35}$$

$$X = X - C(A)(X - f). \tag{36}$$

It is useful to define a *convergence* polynomial by

$$P(A) = 1 - (1 - A)C(A) \tag{37}$$

and to express $C(A)$ as a function of $P(A)$

$$C(A) = \frac{1 - P(A)}{1 - A}. \tag{38}$$

Replacing in (36) $C(A)$ by expression (38) leads to the new modified wave operator equation

$$X = \frac{1}{1 - A}(f - AX) + \frac{P(A)}{1 - A}(X - f). \tag{39}$$

The comparison of equations (24) and (39) shows that the right-hand side of (39) contains an additional term proportional to $P(A)$. Since A is assumed to be an exact Newton–Raphson derivative operator, $P(A)$ must have eigenvalues as close as possible to zero. The higher the degree of the polynomial $P(A)$, the smaller will be the X-dependency of the right-hand side of (39) and the better will be the convergence of the series arising from the iterative solution of this equation. Obviously the crudest approximation is to take for $C(A)$ a zero-degree polynomial, i.e. to approximate the convergence operator by a constant. We will see in the next paragraph that this approximation, within the BW formalism, is identical to the diagonalization method of Davidson.

Comparison with the Davidson method

If the polynomial $C(A)$ is reduced to a constant, equation (30) indicates that the vectorial subspace, in which H is to be diagonalized, is augmented at each step of the iterative process by the vector

$$|\Delta\psi\rangle = (X - f)|0\rangle. \tag{40}$$

If we work within the Brillouin–Wigner approach, f is given by (16), and (40) can be written as

$$\begin{aligned} |\Delta\psi\rangle &= [X - \frac{Q_0}{E - H_0} V(1 + X)]|0\rangle \\ &= \frac{Q_0}{E - H_0}[(E - H_0)X - V(1 + X)|0\rangle \\ &= \frac{Q_0}{E - H_0}(E - H)(X + P_0)|0\rangle \\ &= \frac{Q_0}{E - H_0}(E - H)|\psi\rangle. \end{aligned} \tag{41}$$

The expression (41) of $|\Delta\psi\rangle$ is identical to the vector $\boldsymbol{q}$ in Ref. [6, p. 108–109]. Thus the method presented in this paper is a generalization of the Davidson method which corresponds to the zero-order approximation of C, i.e. to polynomial $C(A)$ of the degree $n = 0$. The Davidson method is especially efficient when all the eigenvalues of A are typically within the range $[-0.8, 0.8]$, but convergence polynomial of higher order are required when the range of the eigenvalues of A is more extended. This will be illustrated by the last numerical application in Sect. 3. An advantage of our approach is that the whole methodology is not associated with symmetric or hermitian operators. In the Davidson scheme, the residual vector which improves the wavefunction at each iteration is an approximate Newton–Raphson correction derived from the minimization of the Rayleigh quotient. In our wave operator approach we emphasize the role of the wavefunction instead of the energy. The Newton–Raphson scheme arises naturally from the initial wave operator equation. We have considered the two most basic RS and BW schemes but other modified or partially renormalized equations could be useful as it was done in determination of relativistic regular two-component Hamiltonians [7, 8]. An advantage of the derivation of exact or approximate Newton–Raphson schemes from wave operators is that the whole formalism applies independently of the nature of the matrix to be diagonalized that can be real non-symmetrical or even non-hermitian as in Floquet representations using optical potentials for the solution of the time-dependent Schrödinger equation [9, 10].

To demonstrate how our algebraic approach works three simple model examples are shown in the last section.

3 Numerical illustration

Three simple models are presented to illustrate the general scheme presented above. We will show on the simplest two-level system how the passage from the RS to the BW scheme can be interpreted in terms of renormalization. The second example of N non-interacting hydrogen molecules will give an idea of the efficiency of the polynomial approximation presented in Sect. 2. Finally, in the last example a comparison with the Davidson method will be done.

1 Two-level system

Let us consider the two-level Hamiltonian

$$H = |1\rangle\langle 1| + \lambda[|0\rangle\langle 1| + |1\rangle\langle 0|], \tag{42}$$

where λ is a perturbation parameter. The matrix representation of H in the basis of the two unperturbed states is given in Fig. 2a. The advantage of this simple model is that the complementary space spanned by $|1\rangle$ is one-dimensional. It allows a simple algebraic discussion since the wave operator equations are reduced to ordinary algebraic equations. The RS and BW functions are given in Appendix A. The functions f_{RS} and f_{BW} given by A2 and A4 appear as a particular case of the general operator expressions (15) and (16). Then the derivatives f'_{RS} and f'_{BW} given by (A6) and (A7) are particular cases of the derivative operators (20) and (21). The

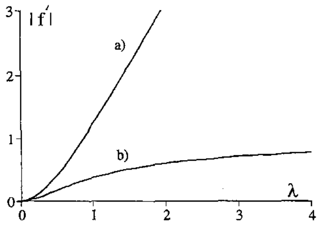

Fig. 1. Graphical representation of the Rayleigh–Schrödinger (RS) and Brillouin–Wigner (BW) derivatives f'_{RS} and f'_{BW} given by (A9) and (A10) of Appendix A. The RS iterative scheme of case a) converges up to $\lambda \approx 0.9$ whereas the BW scheme of case b) is always convergent since $|f'_{BW}|<1$.

graphical representation of f'_{RS} and f'_{BW} given in Fig. 1 displays the convergence properties of the RS and BW methods.

Since the iterative solution of an ordinary algebraic equation $x = f(x)$ converges when $|f'(x)| < 1$, Fig. 1 shows that the RS scheme converges for λ up to $\lambda \approx 0.9$ whereas the convergence of the BW scheme is assured for any value of λ. The two-level system is obviously far from actual applications. We have presented the above elementary model for pedagogical reasons. It illustrates clearly how the convergence of perturbation series can be improved by a purely algebraic approach of infinite summations derived from the knowledge of exact A derivative operators such as those given by (15) and (16).

2 *N non-interacting H_2 molecules*

The model is briefly presented in appendix B and more details can be found in Ref. [11]. The energy of the correlated ground state has been calculated from the Hartree–Fock (HF) wavefunction taken as the zero-order reference state. The matrix representation of the Hamiltonian in the basis of the HF, di-, and quadri-excited states is given in Fig. 2.

$$\underset{\textbf{a}}{\begin{bmatrix} 0 & \lambda \\ \lambda & 1 \end{bmatrix}} \qquad \underset{\textbf{b}}{\begin{bmatrix} 0 & \lambda & \lambda & 0 \\ \lambda & 1 & 0 & \lambda \\ \lambda & 0 & 1 & \lambda \\ 0 & \lambda & \lambda & 2 \end{bmatrix}} \qquad \underset{\textbf{c}}{\begin{bmatrix} 0 & \lambda & \lambda & \lambda & 0 & 0 & 0 & 0 \\ \lambda & 1 & 0 & 0 & \lambda & 0 & \lambda & 0 \\ \lambda & 0 & 1 & 0 & \lambda & \lambda & 0 & 0 \\ \lambda & 0 & 0 & 1 & 0 & \lambda & \lambda & 0 \\ 0 & \lambda & \lambda & 0 & 2 & 0 & 0 & \lambda \\ 0 & 0 & \lambda & \lambda & 0 & 2 & 0 & \lambda \\ 0 & \lambda & 0 & \lambda & 0 & 0 & 2 & \lambda \\ 0 & 0 & 0 & 0 & \lambda & \lambda & \lambda & 3 \end{bmatrix}}$$

Fig. 2. Matrix representation of the model Hamiltonian presented in Appendix B describing N non-interacting H_2 molecules. λ is a perturbation parameter: (a) $N = 1$; (b) $N = 2$; (c) $N = 3$.

For the sake of simplicity, the zero-order HF energy has been taken equal to zero and we have also used arbitrary units. To consider a large perturbation, we have chosen rather arbitrarily $\lambda = 0.6$ which is a value much greater than for actual hydrogen molecules. In all cases the calculations were done by using the perturbation-iteration scheme given by (33). We have calculated a system made up of eight molecules. The calculated energies for the successive iterations are reported in Table 1.

The Brillouin–Wigner series corresponding to the choice $C^{(n)} = 1$ is extremely slowly converging. The Newton–Raphson method with the choice $C^{(n)} = 1/(1 - A^{(n)})$ converges with six exact figures in four iterations. The convergence is clearly quadratic at the end of the iterative process. Between these two extreme situations, the choice of a polynomial convergence operator $C(A)$ leads to intermediate convergences. For a polynomial of the fourth order the results are almost as good as those given in the last column by the Newton–Raphson procedure. At each step of the calculations the best c_k coefficients were optimized by minimizing the euclidian norm $\| X - f \|$.

Table 1. Energies of eight H_2 molecules. The exact figures are underlined (arbitrary units)

Iteration	Brillouin–Wigner	Degree of the polynomial $P(A)$					Newton–Raphson
		0	1	2	3	4	
0	$-\underline{2}.88000$	$-\underline{2}.88000$	$-\underline{2}.88000$	$-\underline{2}.88000$	$-\underline{2}.88000$	$-\underline{2}.88000$	$-\underline{2}.88000$
1	-0.742268	-1.47821	-1.51714	-1.80686	-1.85364	-1.85512	-1.85466
2	-3.36023	-1.72252	$-\underline{2}.18236$	$-\underline{2}.33380$	$-\underline{2}.31051$	$-\underline{2}.30064$	$-\underline{2}.30035$
3	-0.973391	-1.82557	$-\underline{2}.15280$	-2.25172	$-\underline{2.2}5033$	$-\underline{2.24}968$	$-\underline{2.24}967$
4	-3.54737	$-\underline{2}.03419$	$-\underline{2.2}5096$	$-\underline{2.24}932$	$-\underline{2.24}824$	$-\underline{2.24820}$	$-\underline{2.24820}$
5	-1.09896	$-\underline{2}.02169$	$-\underline{2.248}02$	$-\underline{2.24828}$	$-\underline{2.24820}$		
6	-3.59360	$-\underline{2}.28002$	$-\underline{2.24}900$	$-\underline{2.2482}2$			
7	-1.18484	$-\underline{2.2}5646$	$-\underline{2.24820}$	$-\underline{2.24820}$			
8	-3.57152	$-\underline{2.2}5710$	$-\underline{2.248}26$				
9	-1.25193	$-\underline{2.248}34$	$-2.\underline{2482}2$				
10	-3.51806	-2.24946					
11	-1.30866	$-\underline{2.24}735$					
12	-3.45201	$-\underline{2.24}784$					
13	-1.35896	$-\underline{2.24}757$					
14	-3.38273	$-2.\underline{24}787$					
15	-1.40482	$-\underline{2.24}778$					
16	-3.31467	$-\underline{2.248}12$					
17	-1.44732	$-\underline{2.248}18$					
18	-3.24982	$-\underline{2.24820}$					
19	-1.48710						
20	-3.18891						

3 Impurity model

This model illustrates the efficiency of the derivative operator for obtaining solutions which are far from the domain of validity of standard perturbation theory. The matrix representation of the impurity Hamiltonian is given in Fig. 3.

$$\begin{bmatrix} 0 & \lambda & \lambda & \lambda & \lambda & \cdot \\ \lambda & 1 & \mu & 0 & 0 & . \\ \lambda & \mu & 1 & \mu & 0 & . \\ \lambda & 0 & \mu & 1 & \mu & . \\ \lambda & 0 & 0 & \mu & 1 & . \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Fig. 3. Matrix representation of an impurity coupled to the quasi-continuum arising from the Hückel description of a linear polyene. λ characterizes the coupling with the quasi-continuum. 4μ is the width of the Hückel band.

The parameter λ measures the strength of the coupling between the impurity and the outer space and the parameter μ measures the interaction between two neighbouring states in the outer space. It is to be noted that, for this system, the matrix representation of the zero-order derivative operator A, given by expression (25), is identical to the Hückel matrix for a linear polyene [12]. Therefore the eigenvalues of A are known to be within the range $[1 - 2\mu, 1 + 2\mu]$. Thus, it is to be expected that the standard perturbation series and the method of Davidson will be efficient for $\mu < 0.5$. This has been checked numerically. Beyond $\mu = 0.5$, the functional space generated at each iteration by the Davidson method is useless. For the value $\mu = 0.6$, the iterative process does not converge towards the exact value $E = -0.00161$ corresponding to the normalized impurity state $|\psi\rangle = 0.88\,|0\rangle + \cdots$ At the 50th iteration, the Davidson procedure gives the value -0.1905 which is not related to the impurity level and is close to the lowest energy solution $E = -0.1976$. The usefulness of the derivative operator A is illustrated in Table 2.

Table 2. Impurity energy obtained by the iterative solution of Eq. (33). A(exact) and A(zero-order) are given by (21) and (25), respectively. The matrix representation of the impurity model is given in Fig. 3. Calculations were done for a matrix of order 50 with $\lambda = 0.005$ and $\mu = 0.6$. The exact figures are underlined (arbitrary units)

Iteration	A(exact)	A(zero-order)
0	-0.0025	-0.0025
1	$-0.00\underline{1}348$	$-0.00\underline{1}347$
2	$-0.00\underline{16}09$	$-0.00\underline{16}09$
3	$-0.00\underline{1612}$	$-0.00\underline{161}3$
4		$-0.00\underline{1612}$

The use of the perturbation-iteration scheme generated by (33) with, successively, exact and approximate operators A, leads also to the exact impurity level in a few iterations. This solution can also be obtained within the polynomial approximation described in Sect. 2 and numerically illustrated in Table 2 for the model system of N non-interacting H_2 molecules.

Conclusion

This paper has been devoted to the presentation of a new general scheme systematically improving the convergence properties of perturbation series. This has been done by transforming the standard Rayleigh and Brillouin–Wigner wave equations by means of derivative and convergence operators which were given explicitly. The concept of derivative operator generalizes the concept of derivative of an ordinary algebraic function. The advantage of our approach is that it allows to discuss the convergence problems in terms of well defined Newton–Raphson exact operators that can be approximated in many ways providing, for example, efficient infinite summation schemes. From a practical point of view, we have established the relation between our approach and the diagonalization procedure of Davidson. Our method reduces to the Davidson method when the derivative operator is taken equal to zero. Another advantage of our approach is that it is not derived from a variational approach based on the Rayleigh quotient, but from the Newton–Raphson method applied to wave operator equations. This means that our formalism has the same efficiency for both the hermitian and nonhermitian Hamiltonians. Finally, the method can easily be extended to the simultaneous determination of few eigenstates which is relevant to the theory of effective Hamiltonians. Recent progress have already been made in this domain in the determination of regular two-component relativistic Hamiltonians [8]. Other investigations are in progress concerning the convergence properties of effective and intermediate Hamiltonians that will be published in forthcoming papers.

Appendix A

In the basis of the two unperturbed states $|0\rangle$ and $|1\rangle$ the exact Hamiltonian of a two-state system is given by

$$H = |1\rangle\langle 1| + \lambda(|0\rangle\langle 1| + |1\rangle\langle 0|). \tag{A1}$$

λ is a perturbation parameter. For this simple model Hamiltonian the complementary space spanned by $|1\rangle$ is one-dimensional and $f(X)$ given by (15) is reduced then to an ordinary algebraic expression $f(x)$ depending on a scalar variable x

$$f_{\mathrm{RS}} = \lambda(x^2 - 1) \tag{A2}$$

and the Bloch-like equation (14) can be written as

$$x = \lambda(x^2 - 1). \tag{A3}$$

In the Brillouin–Wigner (BW) case Eq. (16) leads to

$$f_{\mathrm{BW}} = -\frac{\lambda}{1 - \lambda x} \tag{A4}$$

and the corresponding Bloch-like expression is

$$x = -\frac{\lambda}{1 - \lambda x}. \tag{A5}$$

The convergence of the iterative solution of (A3) is governed by the value of the derivative of $f(x)$

$$f'_{\mathrm{RS}} = 2\lambda x, \tag{A6}$$

$$f'_{\mathrm{BW}} = -\frac{\lambda^2}{(1-\lambda x)^2}, \tag{A7}$$

in the RS and BW cases, respectively. Since the relevant exact solution of (A3) or (A5) is

$$x = \frac{1-\sqrt{1+4\lambda^2}}{2\lambda} \tag{A8}$$

at convergence the derivatives (A6) and (A7) have the values

$$f'_{\mathrm{RS}} = 1-\sqrt{1+4\lambda^2} \tag{A9}$$

and

$$f'_{\mathrm{BW}} = -\frac{4\lambda^2}{(1+\sqrt{1+4\lambda^2})^2}. \tag{A10}$$

Appendix B

The model of N non-interacting molecules is presented and investigated in reference [11]. Each molecule has only one occupied self consistent field (SCF) molecular orbital (MO) σ and, for simplicity, only one non-occupied virtual MO σ^* is considered per subsystem. The Hartree–Fock ground state can be written as

$$|0\rangle = \prod_{p=1}^{N} a_p^+ a_{\bar{p}}^+ \,|vacuum\rangle. \tag{B1}$$

a_p^+ and $a_{\bar{p}}^+$ are the creation operators of an electron in a σ orbital on site p with a spin α or β. The full configuration interaction (CI) implies di-, quadri-, ... $2N$-excited determinants. For N hydrogen molecules the size of the full CI is 2^N. Figure 2 shows the matrix representation for $N = 1, 2, 3$. λ is a perturbation parameter proportional to the exchange integral associated with the σ and σ^* orbitals.

The correlation energy per molecule, in arbitrary units, is given by

$$E_{\mathrm{corr}} = \tfrac{1}{2}(1-\sqrt{1+4\lambda^2}). \tag{B2}$$

For the N non-interacting molecules the total correlation energy is

$$E_{\mathrm{corr}} = N\tfrac{1}{2}(1-\sqrt{1+4\lambda^2}).$$

The above energy is identical to the total energy since the Hartree–Fock energy is taken to be zero. For more details and a study of the size dependence of delocalized treatments of the correlation problem see Ref. [11].

References

1. Killingbeck J (1977) Rep Prog Phys 40:963
2. Lindgren I, Morrison J (1977) Atomic many-body theory. Springer, Berlin.
3. Kutzelnigg W (1992) Theoret Chim Acta 83:263
4. Durand Ph (1983) Phys Rev A28:3184
5. Durand Ph, Malrieu J-P (1987) in: Lawley KP (ed) Ab initio methods in quantum chemistry I. Wiley, New York

6. Davidson ER, Matrix eigenvector methods in: Diercksen GH and Wilson S (eds), Methods in computational molecular physics. Reidel, Dordrecht, 1983, p 95.
7. Chang C, Pelissier M, Durand Ph (1985) Phys Scr 34:394
8. van Lenthe E, Baerends EJ, Snijders JG (1993) J Chem Phys 99:4597
9. Shirley JH (1965) Phys Rev B138:979
10. Jolicard G (1985) Chem Phys Lett 121:106
11. Malrieu J-P (1979) J Chem Phys 70:4405
12. Heilbronner E, Bock H (1976) The HMO model and its applications. Wiley, New York (1976)