

ABOUT APPROXIMATION OF CONVERGENCE SUPEROPERATORS IN QUANTUM PERTURBATION THEORY

D. BIELIŃSKA-WĄŻ^a, I. PAIDAROVÁ^b AND PH. DURAND^c

^aInstytut Fizyki, Uniwersytet Mikołaja Kopernika
Grudziądzka 5, 87-100 Toruń, Poland

^bJ. Heyrovský Institute, Academy of Sciences of the Czech Republic
Dolejškova 3, 182 23 Prague 8, Czech Republic

^cLaboratoire de Physique Quantique, Unité Associée au CNRS no. 505
Université Paul Sabatier, 31062 Toulouse Cedex 4, France

(Received August 2, 1996; revised version February 6, 1997)

Perturbation methods are generally used for solving wave operator equations associated with the determination of effective Hamiltonians. In many cases the standard Rayleigh–Schrödinger and Brillouin–Wigner series either converge slowly or diverge. Therefore it is necessary to modify or to renormalize the standard wave equations. For that purpose derivative and convergence superoperators within the Rayleigh–Schrödinger and Brillouin–Wigner formalisms were introduced. A new efficient approximation for convergence superoperators is investigated in this paper. Its application to a model system of N non-interacting molecules shows that this approximation can overcome convergence difficulties.

PACS numbers: 31.15.Md

1. Introduction

The aim of this paper is to contribute to convergence studies in quantum perturbation theory [1, 2]. The convergence properties of the standard Rayleigh–Schrödinger and Brillouin–Wigner perturbation expansions are slow and for many systems the series diverge. In order to overcome the convergence difficulties many methods have been developed, for example, infinite summations, Padé approximations, diagrammatic expansions and infinite partial summations to all orders that, however, cannot be considered as the final development of the theory.

A powerful approach is to use derivative and convergence superoperators introduced by Durand et al. [3]. These superoperators enable us to investigate and extend the convergence properties of the standard Rayleigh–Schrödinger and

Brillouin–Wigner schemes. However, the direct use of these superoperators in solving wave operator equation, associated with the determination of an effective Hamiltonian, implies difficult operator inversions. In order to avoid them it is necessary to use an approximation. In this article we propose a new approximation of the convergence superoperator. Its efficiency is tested on a numerical example and a model system of N non-interacting molecules.

2. The method

Let us assume that the Hamiltonian H is partitioned into the unperturbed Hamiltonian H_0 and a perturbation V

$$H = H_0 + V. \quad (1)$$

The Hilbert space is split into a finite n -dimensional model space, which is spanned by a set of eigenvectors of H_0 , and the outer space. Two projection operators P_0 and Q_0 , with the property $P_0 + Q_0 = 1$, are associated with these two spaces. In this paper we have in view to solve the Schrödinger equation for molecules by the configuration interaction (CI) method. In this method a finite number of atomic orbitals generates a finite number of electronic configurations, therefore both the outer space and the entire Hilbert space have finite dimensions and all operators and superoperators are represented by finite matrices. In the following, operators and superoperators will be represented by roman and calligraphic letters, respectively. Let us consider the Schrödinger equation

$$H|\Psi_i\rangle = E_i|\Psi_i\rangle, \quad i = 1, \dots, n. \quad (2)$$

The exact solutions $|\Psi_i\rangle$ can be written as

$$|\Psi_i\rangle = \Omega|i\rangle, \quad i = 1, 2, \dots, n, \quad (3)$$

where Ω is the wave operator [2]. The result of the action of Ω on an unperturbed state $|i\rangle$ is $|\Psi_i\rangle$, and Ω obeys the intermediate normalization

$$P_0\Omega = P_0. \quad (4)$$

Using Eqs. (3) and (4) the Schrödinger equation (2) can be transformed into the basic wave operator equation [4, 5]

$$H\Omega = \Omega H\Omega. \quad (5)$$

The importance of this equation, which depends quadratically on Ω , has been recently stressed by Löwdin [6], who advocates new investigations of this non linear equation. This paper provides new tools for solving Eq. (5). Equation (5) can also be written in the form

$$H\Omega = \Omega H^{\text{eff}}. \quad (6)$$

Equation (6) generalizes the Schrödinger equation for n eigenvalues. The effective Hamiltonian in Eq. (6) is written

$$H^{\text{eff}} = P_0 H \Omega, \quad (7)$$

H^{eff} is defined within the model space and provides n exact eigenvalues. It is useful to define the reduced wave operator X by writing

$$\Omega = (P_0 + Q_0)\Omega = P_0 + X. \quad (8)$$

Multiplying Eq. (5) from the left by Q_0 and from the right by P_0 and using expression (8), Eq. (5) can be written in the following form:

$$Q_0(1 - X)H(1 + X)P_0 = 0. \quad (9)$$

Equation (9) can be transformed into perturbative wave operator expressions which can always be written in the form

$$X = f(X), \quad (10)$$

where the operator $f(X)$ is a function of X [5]. The two most basic expressions of $f(X)$ are associated with the Rayleigh–Schrödinger (RS) and Brillouin–Wigner (BW) schemes. In the case of a one-dimensional space, $f(X)$ is given by

$$f(X) = \frac{Q_0}{E_0 - H_0}(1 - X)V(1 + X)P_0, \quad (11)$$

for the RS scheme and by

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

for the BW scheme. $\frac{Q_0}{E_0 - H_0}$ in (11) is a short hand notation for the inversion of the projected operator $Q_0(E_0 - H_0)Q_0$ in the outer space. E_0 in (11) is the zero order energy, i.e., the energy of the unperturbed state, and E in (12) is the exact energy. For greater values of n , the expressions of $f(X)$ that generalize (11) and (12) can be found in Ref. [7]. The standard RS and BW perturbation expansions arise from the iterative solutions of (10). Unfortunately, the series very often converge slowly or diverge. In order to avoid these difficulties Eq. (10) has to be modified [3]. The general idea is to transform the wave operator equation (10) with the aim to reduce as much as possible the dependence of $f(X)$ on X . This can be done formally by introducing a superoperator \mathcal{A} acting in the vectorial space of all operators coupling the model space and the outer space. Subtracting $\mathcal{A}X$ from both sides of Eq. (10) leads to the modified wave operator equation

$$X = \frac{1}{1 - \mathcal{A}}(f(X) - \mathcal{A}X). \quad (13)$$

When $\mathcal{A} = 0$ this equation reduces to the standard perturbation expression (10). The best choice of \mathcal{A} is associated with the Newton–Raphson solution of Eq. (10). For that purpose Eq. (10) is linearized in the neighborhood of X

$$X + \Delta X = f(X + \Delta X) = f(X) + \mathcal{A}\Delta X + \dots, \quad (14)$$

where ΔX is a small variation of X and \mathcal{A} is the derivative superoperator acting in the vectorial space of all operators which couple the model space and the outer space. \mathcal{A} generalizes the concept of derivative $f'(x)$ associated with the Newton–Raphson solution of the algebraic equation $x = f(x)$. Expressions of \mathcal{A} within the RS and BW schemes can be found in Ref. [7]. The iterative solution of Eq. (13) defines a quadratically convergent Newton–Raphson procedure. Equation (13) can be also written in the form

$$X = X - \mathcal{C}[X - f(X)], \quad (15)$$

where \mathcal{C} is the convergence superoperator

$$\mathcal{C} = \frac{1}{1 - \mathcal{A}}. \quad (16)$$

The iterative solution of Eq. (15) is not practicable because the inversion of $1 - \mathcal{A}$ would be as difficult as solving the Schrödinger equation. For that reason it is necessary to use for \mathcal{C} an approximate expression. For instance, \mathcal{C} can be approximated by a polynomial expansion [7]

$$\mathcal{C} = \frac{1}{1 - \mathcal{A}} = \sum_{k=0}^m c_k \mathcal{A}^k. \quad (17)$$

In this paper we propose another approximation of \mathcal{C} . Let us write

$$\Delta \mathcal{A} = \mathcal{A} - \mathcal{A}_0, \quad (18)$$

where \mathcal{A} is the exact derivative superoperator and \mathcal{A}_0 is a reasonable approximation of \mathcal{A} . The convergence superoperator can be expanded in terms of $\Delta \mathcal{A}$

$$\mathcal{C} = \frac{1}{1 - \mathcal{A}_0} + \frac{1}{1 - \mathcal{A}_0} \Delta \mathcal{A} \frac{1}{1 - \mathcal{A}_0} + \dots, \quad (19)$$

or more generally as

$$\mathcal{C} = \frac{1}{1 - \mathcal{A}} = \sum_{k=0}^m \left[\frac{1}{1 - \mathcal{A}_0} \Delta \mathcal{A} \right]^k \frac{c_k}{1 - \mathcal{A}_0}. \quad (20)$$

If we take $\mathcal{A}_0 = 0$, expression (20) reduces to the polynomial form (17). In standard perturbation theory the convergence pattern associated with the iterative solution of Eq. (10) depends on the eigenvalues of \mathcal{A} . If their moduli are small with respect to 1 then the RS and BW perturbation series converge, but if they are close to 1, convergence difficulties appear. In our approach the convergence properties of expansion (19) have to be discussed in terms of the eigenvalues of the superoperator $\frac{1}{1 - \mathcal{A}_0} \Delta \mathcal{A}$, the moduli of which must remain as small as possible with respect to 1. The best choice is to choose \mathcal{A}_0 as close as possible to \mathcal{A} under the constraint that one must be able to invert $1 - \mathcal{A}_0$. There are many efficient ways to do that. The Brillouin–Wigner approach is especially attractive since the inversion of superoperators reduce to the inversion of operators. This scheme will be used in the next section in the two numerical applications. For large CI calculations the almost band structure of the H matrix makes useful the introduction of an intermediate space which greatly improves the convergence properties [8]. These developments are far beyond the scope of this paper.

3. Numerical illustration and discussion

The convergence properties discussed above are presented on a purely numerical example and for a model system of N non-interacting molecules which is frequently used as a benchmark for investigating the convergence properties of CI methods [9].

3.1. Numerical example

The components of the Hamiltonian matrix are given by $H_{ij} = i$, if $i = j$ and $H_{ij} = \lambda$, if $i \neq j$, for $i, j = 1, 2, \dots, N$. The calculation were done with $\lambda = 0.6$ and $N = 100$. Table I presents the results obtained for the lowest eigenvalue E_1 . The two lowest eigenvalues E_1, E_2 , are investigated simultaneously in Table II. The comparison with the Newton–Raphson procedure shows that the convergence properties are significantly improved by adding the first order term in expression (19).

TABLE I

Energies E_1 obtained by the iterative solution of Eq. (15) ($N = 100$) using three different C operators.

iteration	E^a	E^b	E^c
1	-0.86386	-0.86386	-0.86386
2	0.71640	0.71640	0.71640
3	0.49299	0.56023	0.54433
4	0.55287	0.54215	0.54081
5	0.53775	0.54090	
6	0.54160	0.54082	
7	0.54061	0.54081	
8	0.54086		
9	0.54080		
10	0.54082		
11	0.54081		

$^aC = \frac{1}{1-\mathcal{A}_0}$,

$^bC = \frac{1}{1-\mathcal{A}_0} + \frac{1}{1-\mathcal{A}_0} \Delta \mathcal{A} \frac{1}{1-\mathcal{A}_0}$, \mathcal{A}_0 is the matrix \mathcal{A} from the first iteration,

c Newton-Raphson procedure.

TABLE II

Energies E_1 and E_2 obtained by the iterative solution of Eq. (15).

iteration	E_1^a	E_2^a	E_1^b	E_2^b	E_1^c	E_2^c
1	-1.29677	0.93269	-1.29677	0.93269	-1.29677	0.93269
2	0.61493	1.88371	0.61493	1.88371	0.61493	1.88371
3	0.51422	1.44389	0.55398	1.64183	0.54568	1.58090
4	0.54504	1.59277	0.54159	1.56858	0.54082	1.55955
5	0.53992	1.54928	0.54085	1.56039	0.54081	1.55954
6	0.54099	1.56263	0.54081	1.55962		
7	0.54078	1.55860	0.54081	1.55954		
8	0.54082	1.55982				
9	0.54081	1.55945				
10	0.54081	1.55956				
11	0.54081	1.55953				
12	0.54081	1.55954				

$^a, ^b, ^c$ see footnotes to Table I.

3.2. N non-interacting molecules ($N = 8$)

The description of this model system can be found in Ref. [7]. We assume that the first $N-1$ molecules are identical and that their matrix representation in the basis of the Hartree-Fock ground state and of the first excited state is given by A , the matrix representation for the N -th molecule is given by B (arbitrary units)

$$A = \begin{bmatrix} 0 & \lambda \\ \lambda & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & \lambda \\ \lambda & \mu \end{bmatrix}.$$

The size of the full CI matrix is 2^N . It is obtained by the tensorial product of $N-1$ times the matrix A by the matrix B . For this model system the total energies are known. For the ground state and for the first excited state they are given by

$$E_1 = (N-1) \frac{1}{2} \left(1 - \sqrt{1 + 4\lambda^2} \right) + \frac{1}{2} \left(\mu - \sqrt{\mu^2 + 4\lambda^2} \right), \quad (21)$$

$$E_2 = (N-1) \frac{1}{2} \left(1 - \sqrt{1 + 4\lambda^2} \right) + \frac{1}{2} \left(\mu + \sqrt{\mu^2 + 4\lambda^2} \right). \quad (22)$$

When all molecules are identical ($\mu = 1$), the energy of the ground state is far from the energy of the first excited state. Results for the ground state with $\lambda = 0.3$ and $\mu = 1$ are given in Table III. If $\mu \simeq 0$ the two lowest states are almost degenerate. The results corresponding to this case ($\lambda = 0.3$, $\mu = 0$) which is relevant to a two-dimensional effective Hamiltonian are reported in Table IV. From the comparison of Tables III and IV with Tables I and II it is seen that in both examples (one and two eigenvalues) the convergence properties are quite similar. An extended numerical study by varying the parameters λ and μ has confirmed the

TABLE III

Energy of the ground state E_1 for a model system of eight identical non-interacting molecules obtained by the iterative solution of Eq. (15) using three different C operators, ($\lambda = 0.3$, $\mu = 1$).

iteration	E^a	E^b	E^c
1	-0.72000	-0.72000	-0.72000
2	-0.65519	-0.65519	-0.65519
3	-0.66629	-0.66454	-0.66478
4	-0.66451	-0.66476	-0.66476
5	-0.66480		
6	-0.66475		
7	-0.66476		

^{a,b,c}see footnotes to Table I.

TABLE IV

Energy of the ground state E_1 and of the first excited state E_2 for a model system of eight molecules obtained by iterative solution of Eq. (15) using three different C operators, ($\lambda = 0.3$, and $\mu = 0$).

iteration	E_1^a	E_2^a	E_1^b	E_2^b	E_1^c	E_2^c
1	-0.93000	-0.33000	-0.93000	-0.33000	-0.93000	-0.33000
2	-0.87586	-0.27364	-0.87586	-0.27364	-0.87586	-0.27364
3	-0.88572	-0.28303	-0.87832	-0.28143	-0.88168	-0.28166
4	-0.87825	-0.28143	-0.87943	-0.28166	-0.88167	-0.28167
5	-0.88447	-0.28171	-0.88016	-0.28167		
6	-0.87935	-0.28166	-0.88065	-0.28167		
7	-0.88358	-0.28167	-0.88098	-0.28167		
8	-0.88009	-0.28167	-0.88120	-0.28167		
9	-0.88297	-0.28167	-0.88135	-0.28167		
10	-0.88059	-0.28167	-0.88145	-0.28167		
11	-0.88255	-0.28167	-0.88152	-0.28167		
12	-0.88094	-0.28167	-0.88157	-0.28167		
13	-0.88227	-0.28167	-0.88160	-0.28167		
14	-0.88117	-0.28167	-0.88162	-0.28167		
15	-0.88208	-0.28167	-0.88163	-0.28167		
16	-0.88133	-0.28167	-0.88164	-0.28167		
17	-0.88195	-0.28167	-0.88165	-0.28167		
18	-0.88144	-0.28167	-0.88166	-0.28167		
19	-0.88186	-0.28167				
20	-0.88151	-0.28167				
21	-0.88180	-0.28167				

^{a,b,c}see footnotes to Table I.

validity of the convergence scheme presented in this paper. In all cases the approximation $C = \frac{1}{1-\mathcal{A}_0} + \frac{1}{1-\mathcal{A}_0} \Delta \mathcal{A} \frac{1}{1-\mathcal{A}_0}$ is efficient. It provides results which converge almost as well as the Newton-Raphson procedure using the exact expression of the convergence operator.

The approximation proposed in this paper seems to be more efficient than the polynomial approximation used in Ref. [3]. This work should be considered as a pilot study checking its efficiency on simple model systems. For actual systems with large CI matrix representations the direct inversion $1-\mathcal{A}_0$ is an impossible task. As mentioned above, the solution could replace the superoperator \mathcal{A}_0 by its projection in a reduced vectorial space of a lower dimension. Work in this direction is in progress.

References

- [1] J. Killingbeck, *Rep. Prog. Phys.* **40**, 963 (1977).
- [2] I. Lindgren, J. Morrison, *Atomic Many-Body Theory*, Springer, Berlin 1977.
- [3] Ph. Durand, J. Šavrdá, I. Paidarová, *Theor. Chim. Acta* **88**, 243 (1994).
- [4] Ph. Durand, *Phys. Rev. A* **28**, 3184 (1983).
- [5] Ph. Durand, J.-P. Malrieu, in: *Ab initio Methods in Quantum Chemistry*, Ed. K.P. Lawley, Vol. 1, Wiley, New York 1987.
- [6] P. Löwdin, *Int. J. Quantum Chem.* **55**, 77 (1995).
- [7] Ph. Durand, I. Paidarová, *Int. J. Quantum Chem.* **58**, 341 (1996).
- [8] J.P. Malrieu, Ph. Durand, J.P. Daudey, *J.Phys. A, Math. Gen.* **18**, 809 (1985).
- [9] J.P. Malrieu, *J. Chem. Phys.* **70**, 4405 (1979).