

Convergence property of multireference many-body perturbation theory analyzed by the use of a norm of effective Hamiltonian

Haruyuki Nakano and Shigeru Obara

Department of Chemistry, Faculty of Science, Kyoto University, Kyoto 606, Japan

Received December 22, 1992/Accepted February 13, 1993

Summary. It is proposed to use a norm of a n th order effective Hamiltonian, for analyzing the convergence property of the multireference many-body perturbation theory (MR-MBPT). The utilization of the norm allows us to employ only (1) a *single* number for all the states that we are interested in, and (2) values which decreases from the *positive* side to zero as the order n of the perturbation increases. This characteristic features are in contrast to those in the usually used scheme where *several* numbers, namely, the eigenvalues of the target states, should be used and they may *oscillate* around exact eigenvalues. The present method has been applied to MR-MBPT calculations of the $(\text{H}_2)_2$, CH_2 , and LiH molecules based on the multireference versions of Rayleigh–Schrödinger PT, Kirtman–Certain–Hirschfelder PT, and the canonical Van Vleck PT; and following features are found: (1) the above three versions of the perturbation theories have essentially the same convergence property judged from the lowering of the norm; (2) the lower order truncation of the perturbation series gives reasonable solutions; (3) the norm decreases irrespective of the perturbation expansion being convergent or divergent for the first several orders (up to about the sixth order).

Key words: Convergence property – Perturbation – Multireference – Quasidegenerate – Effective Hamiltonian

1 Introduction

Many-body perturbation theory (MPBT) [1–12] has been utilized as a convenient method of taking into account of the electron correlation beyond the Hartree–Fock approximation. The single reference version (SR-MBPT) [1–3], which is for the system expressed by a single main configuration, is fully established and equipped as a standard tool for program systems such as GAUSSIAN series [13]. Multireference version (MR-MBPT) [4–12] for the system with (quasi-)degeneracy is also getting popular. At actual calculations the perturbation expansion is presumed to be convergent; however, it is not always the case when the potential energy surfaces are evaluated by the method. It is desirable to have a simple and convenient scheme to check whether the

expansion is convergent or not. The purpose of the present work is to propose this kind of a scheme and to exemplify the scheme by investigating the convergence properties of some molecules.

It is too difficult to evaluate the higher order terms by the diagrammatic approach beyond the third and fourth orders for MR-MBPT and SR-MBPT, respectively, and hence few studies have been done by it. On the contrary, the method based on the Hamiltonian matrix is rather straightforward, so many works have been done for both the single reference version [14, 15] and the multireference version [16–18].

In these studies of the multireference version, the convergence properties have been discussed based on the n th order *energies* obtained by a diagonalization of the effective Hamiltonian. Although the energy is one of the most important quantities in practical studies, it is the effective Hamiltonian that is immediately obtainable from the order-by-order calculation of the perturbation expansion. Therefore, in order to investigate the convergence properties, the examination of the effective Hamiltonian itself is more direct than that of the energies. In this article, we investigate the convergence property of the effective Hamiltonian by introducing a norm of the Hamiltonian. We have focused on clarifying the following aspects: (1) how can the convergence property be investigated by the use of the norm of the effective Hamiltonian, (2) how is the convergence property for some typical molecules, and (3) how different is the convergence property for the several versions of the multireference many-body perturbation theories, such as, the Rayleigh–Schrödinger PT(RS-PT) [4–8], the Kirtman–Certain–Hirschfelder PT(KCH-PT) [19–21], and the canonical Van Vleck PT(CVV-PT) [22].

In the subsequent section the norm of the effective Hamiltonian is defined, in the third section computational details are given, in the fourth section the calculated results are discussed, and concluding remarks are drawn in the last section.

2 Norm of the effective Hamiltonian

In RS-PT, KCH-PT, and CVV-PT the total Hamiltonian $H(=H_0 + V)$ is transformed into:

$$K = H_0 + \sum_{n=1}^{\infty} K^{(n)} \quad (1)$$

with $K^{(n)}$ taking a form of 2×2 block diagonal:

$$K^{(n)} = \begin{pmatrix} K_{\text{eff}}^{(n)} & 0 \\ 0 & K_{QO}^{(n)} \end{pmatrix}. \quad (2)$$

Here $K_{\text{eff}}^{(n)}$ is for the n th order term of the states which we are to solve. We define a norm as a square root of the sum of squared absolute values of all elements in the smaller part of the n th order effective Hamiltonian K_{eff} :

$$\|K_{\text{eff}}^{(n)}\| = \left[\sum_{i,j} |(K_{\text{eff}}^{(n)})_{ij}|^2 \right]^{1/2}. \quad (3)$$

The advantages of employing the norm over the usual method, where the n th order eigenvalues are employed after summing up the effective Hamiltonian up to n th order and diagonalizing it, are that (1) the effective Hamiltonian is an

immediately obtained quantity from the order-by-order calculation of the perturbation expansion, so that the norm is expected to reflect more directly the convergence property than the eigenvalues, (2) the convergence property of all the states interested in are treated simultaneously in contrast to the eigenvalue method, where the analysis should be carried out for each state separately, (3) the positiveness of the norm results in a simple behavior of the norm as the order n increases, namely, the norm decreases from a positive value to zero, while in the eigenvalue method the energy may oscillate around the exact value.

There are other definitions of a norm of a matrix: (1) the maximum absolute value of elements, (2) a sum of the absolute values of all the elements, and (3) the maximum absolute value of the eigenvalues of the matrix. They are, however, found not to lead to much different results in the present analysis of the convergence property.

The n th order effective Hamiltonian is, roughly speaking, proportional to V^n so that the logarithm of the n th order norm is expected to decrease proportional to the order of the perturbation. On account of this, the inverse sign of the gradient of $\log \|K_{\text{eff}}^{(n)}\|$ with respect to the perturbation order n , can be considered as a *rate* of convergence.

One might consider that it is better to use the error, namely the difference of the sum up to the n th order contribution from the exact effective Hamiltonian, for the analysis of the convergence property. This notion seems to be plausible, but is not practical, since (1) the calculation of the exact effective Hamiltonian is not a simple task, and (2) the finiteness of the numerical accuracy in a computer hampers an accurate evaluation of the error in particular for higher orders.

3 Calculation

Based on the method given in the previous section, we carried out the perturbation calculations for some molecules, and investigated the convergence properties by the use of the norm of the n th order contribution to the effective Hamiltonian. The scheme of the partitioning of a total Hamiltonian is that by Epstein and Nesbet [25], that is, only the diagonal elements of the total Hamiltonian matrix are used as the unperturbed Hamiltonian matrix H_0 , and all the off-diagonal elements are the perturbation V . The orders of the perturbation expansion calculated are up to 100 for RS-PT and KCH-PT, and 20 for CVV-PT. The lowness of the order in CVV-PT is caused by a very much consumption of a computer time, which originates from a non-existence of a formula irrelevant in the order n of the perturbation. However, the convergence property of the CVV-PT investigated by up to 20th order would be sufficient for the analysis of the convergence property.

As N -electron functions, we used the configuration state functions (CSF's), which satisfy the space- and spin-symmetry, since they result in a smaller size of matrices in the calculation than by the Slater determinants, which are usually used in full CI method. From a practical point of view, the smallness of the size is essential at the calculation of contributions from such higher-order terms as 100.

The present method is applied to lower electronic states of three molecules, namely, 1) dimeric hydrogen molecules ($(\text{H}_2)_2$), 2) methylene (CH_2), and 3) lithium hydride (LiH). Details of calculations, such as, geometries, the basis sets,

the N -electron state functions (CSF's), and the model spaces, are collected in the following:

(1) $(\text{H}_2)_2$ (1A_g). The "P4-model" [23] is used as the conformations of four H atoms. There, four hydrogen atoms conform (i) a square with the distance of 2.0 a.u. and (ii) a rectangle with the sides of 2.0 and 2.6 a.u. Three primitive s -type Gaussian functions are placed on each atom as the basis set. Their orbital exponents are 4.501800, 0.681444, and 0.151398 [24]. All single and double excitations from all the possible 1A_g state functions generated by the use of the lowest eight RHF orbitals, $1a_g$, $2a_g$, $1b_{2u}$, $1b_{3u}$, $2b_{3u}$, $1b_{1g}$, and $2b_{1g}$ are included as the CSF's. The CSF's with configurations $(a_g)^2(b_{2u})^2$ and $(a_g)^2(b_{3u})^2$ are chosen as model-space functions, since they are degenerate in the square conformation.

(2) CH_2 (1A_1). The basis set is MIDI-4 [24]. Two geometries are used. One is the optimized geometry by the RHF calculation, where the CH bond distance and the HCH angle are found to be 2.0963 a.u. and 105.3 degrees, respectively. The other is a geometry having 1.2 times elongated CH bonds. The CSF space consists of all single and double excitations from all possible 1A_1 state functions generated by using the lowest three of a_1 orbitals, two of b_1 orbitals, and one of b_2 orbitals, with $1a_1$ frozen as a core. Two CSF's $(1a_1)^2(2a_2)^2(1b_1)^2(3a_1)^2$ and $(1a_1)^2(2a_1)^2(1b_1)^2(4a_1)^2$ are the functions in the model space.

(3) LiH (${}^1\Sigma^+$). The MIDI-1 [24] augmented with a single p -type Gaussian function ($\zeta = 0.76$) on the lithium atom is used as the basis set. Two geometries are used, where the LiH bond length are optimized by the RHF calculation, 3.1824 a.u., and is elongated to 4.4554 a.u. The CSF's used are full, namely, all the N -electron functions having ${}^1\Sigma^+$ symmetry. The model states are $(1\sigma_1)^2(2\sigma_1)(3\sigma_1)$, $(1\sigma_1)^2(2\sigma_1)(4\sigma_1)$, and $(1\sigma_1)^2(2\sigma_1)(5\sigma_1)$.

Although the CSF spaces of the calculations of $(\text{H}_2)_2$ and CH_2 are not full, it is sufficiently large for the investigation of the convergence properties. Indeed, the convergence property in the calculation of LiH , with a reduced CSF space from the full, is found to be essentially the same.

4 Results and discussion

First of all, to grasp the behavior of the norm in increasing the order n of the perturbation expansion, we give an example: the total Hamiltonian $\mathbf{H} (= \mathbf{H}_0 + \mathbf{V})$ is the following 2×2 matrix:

$$\begin{aligned} \mathbf{H} &= \begin{pmatrix} -a & b \\ b & a \end{pmatrix} \\ &= \begin{pmatrix} -a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix} \quad (a, b > 0), \end{aligned} \quad (4)$$

whose eigenvalues E_{\pm} are:

$$E_{\pm} = \pm(a^2 + b^2)^{1/2}. \quad (5)$$

The n th order effective Hamiltonian $\mathbf{K}_{\text{eff}}^{(n)}$ is now a (1×1) matrix, so that its element is the n th order energy itself and the norm is its absolute value. There is no ambiguity which measure of a norm given previously are used. The perturbation expansion of the energy is just the Taylor expansion of $(1 + b^2/a^2)^{1/2}$ in E_{\pm} :

$$\begin{aligned}
 E_{\pm} &= \pm a \left(1 + \frac{b^2}{a^2} \right)^{1/2} \\
 &= \pm a \left[1 + \frac{b^2}{2a^2} - \frac{b^4}{8a^4} + \cdots + (-1)^{n-1} \frac{(2n-3)!!}{(2n)!!} \left(\frac{b}{a} \right)^{2n} + \cdots \right], \quad (6)
 \end{aligned}$$

which is convergent if b^2/a^2 is less than or equal to one, and is divergent otherwise although $(1 + b^2/a^2)^{1/2}$ itself takes a finite value. As can be seen from the expansion, only the even order of b contributes to the energy. The results are shown in Fig. 1 for the three sets of a and b : (a) $a = 1.0$, $b = 0.5$, (b) $a = 0.5$, $b = 0.5$, and (c) $a = 0.45$, $b = 0.5$. The sets (a), (b), and (c) are the examples of rapid convergence, slow convergence, and slow divergence, respectively. Features found are that (1) in the case of (a), the norm of n th order contribution to the effective Hamiltonian monotonously decrease, (2) the change of the norm seems to have two stages, one is up to first several orders and the other is the subsequent orders, which is more apparent in the cases of (b) and (c), and (3) even in slow divergent case (c) the norm decreases at first several orders.

Now we give the result for the molecules described in the previous section. The norm $\|K_{\text{eff}}^{(n)}\|$ of the n th order contribution $K_{\text{eff}}^{(n)}$ to the effective Hamiltonian K_{eff} has been calculated (Table 1) and depicted in Fig. 2 for up to the 35th order. The tendency of the change of the norm at higher orders than the 35th is essentially the same. All the norms in the figure decrease as n increases, which indicates the perturbation expansion is convergent. Features found from the figure are as follows:

(1) All the norms at the orders higher than the sixth are smaller than those at the second order to the sixth order. This means that the lower order contributions include important interaction in the total Hamiltonian, and therefore, the lower order truncation of the perturbation series gives reasonable solutions of both the ground and the excited states. This result agrees with that of Knowles et al. concluded for the ground state by the single reference perturbation theory [14].

(2) The rate of the lowering of the norm changes at about sixth order as elucidated by solid lines drawn in the figure; initially the lowering is rapid then it becomes slower. This type of a decrease of the rapidity of convergence has been found in the single reference calculations [14]. We note that the behavior of this two-stage lowering is found also in the above example of the 2×2 Hamiltonian matrix.

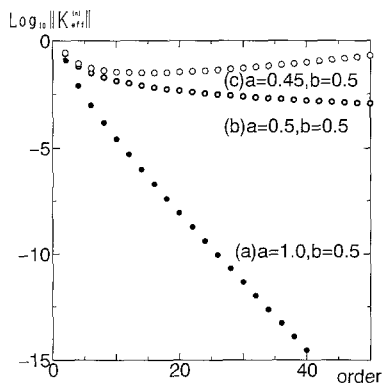
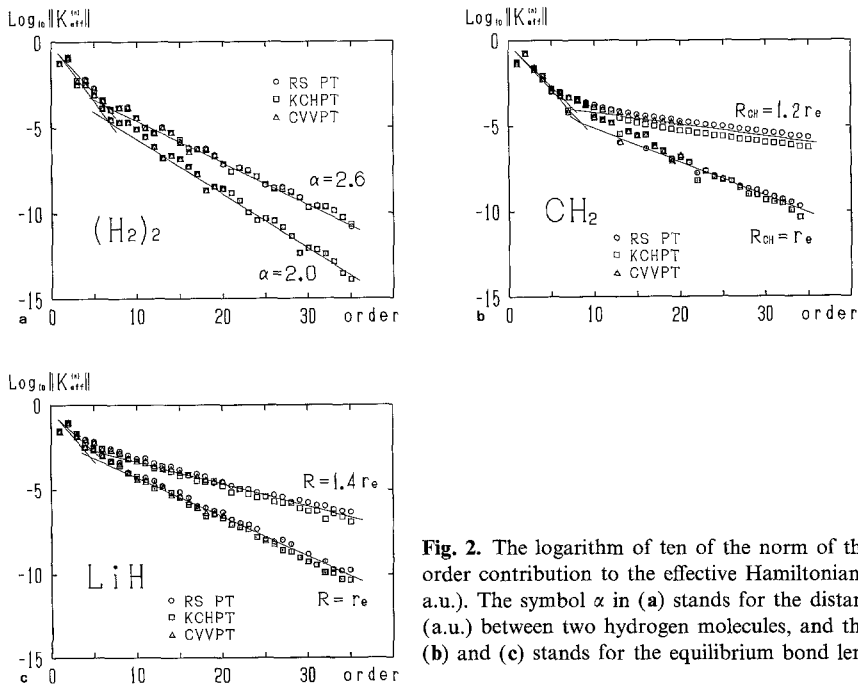


Fig. 1. The logarithm of ten of the norm of n th order contribution to the effective Hamiltonian: (a) rapid convergent case, (b) slowly convergent case, and (c) slowly divergent case

Table 1. The norm $\|K_{\text{eff}}^{(n)}\|$ of the n th order contribution to the effective Hamiltonian (a.u.).

Order	1	2	3	4	5	6	7
			(H ₂) ₂	$\alpha = 2.0$			
RS-PT	0.0604227	0.1152107	0.0035343	0.0037054	0.0009328	0.0001522	0.0000312
KCH-PT	0.0604227	0.1152107	0.0034961	0.0037409	0.0009253	0.0001479	0.0000301
CVV-PT	0.0604227	0.1152107	0.0034879	0.0036976	0.0009255	0.0001525	0.0000326
				$\alpha = 2.6$			
RS-PT	0.0607751	0.1374467	0.0061244	0.0068273	0.0022970	0.0004134	0.0001091
KCH-PT	0.0607751	0.1374465	0.0061049	0.0061695	0.0021047	0.0004392	0.0001319
CVV-PT	0.0607751	0.1374465	0.0056883	0.0064487	0.0022203	0.0003790	0.0001253
			CH ₂	$r = r_e$			
RS-PT	0.0527405	0.1729374	0.0276498	0.0055501	0.0011871	0.0009240	0.0000637
KCH-PT	0.0527405	0.1729324	0.0276563	0.0053336	0.0012430	0.0009987	0.0000879
CVV-PT	0.0527405	0.1729324	0.0275896	0.0054144	0.0011395	0.0009352	0.0000702
				$r = 1.2r_e$			
RS-PT	0.0464480	0.1790188	0.0208590	0.0087538	0.0015451	0.0005650	0.0005028
KCH-PT	0.0464480	0.1790110	0.0208104	0.0086175	0.0015465	0.0006374	0.0004830
CVV-PT	0.0464480	0.1790110	0.0208073	0.0086748	0.0015285	0.0005639	0.0005096
			LiH	$r = r_e$			
RS-PT	0.0303517	0.0766322	0.0162235	0.0036892	0.0028471	0.0010887	0.0005886
KCH-PT	0.0303517	0.0762923	0.0160619	0.0031771	0.0021247	0.0009814	0.0005162
CVV-PT	0.0303517	0.0762923	0.0161345	0.0034645	0.0025535	0.0010383	0.0005182
				$r = 1.4r_e$			
RS-PT	0.0286622	0.0959195	0.0233945	0.0094840	0.0071137	0.0028763	0.0027452
KCH-PT	0.0286622	0.0953640	0.0224555	0.0075893	0.0057613	0.0027333	0.0022064
CVV-PT	0.0286622	0.0953640	0.0229016	0.0085707	0.0065104	0.0026317	0.0022386

**Fig. 2.** The logarithm of ten of the norm of the n th order contribution to the effective Hamiltonian (in a.u.). The symbol α in (a) stands for the distance (a.u.) between two hydrogen molecules, and the r_e in (b) and (c) stands for the equilibrium bond length

(3) There is no apparent difference in the convergence property among the three perturbation methods, namely, RS-PT, KCH-PT, and CVV-PT. Although the norms of the KCH-PT are smaller than those of the RS-PT at the tenth and later orders in CH_2 with the elongated CH bonds their behavior is “parallel”, namely, the differences of the norms do not increase. If they were increased, the KCH-PT can be said to be superior to the others, but they don't. One might consider that there is a geometry where the increase of the differences *does* take place; however, it is not found from the calculations by modifying $R_{\text{H}_2-\text{H}_2}$, $R_{\text{C}-\text{H}}$, and $R_{\text{Li}-\text{H}}$ with an increment of 0.02 a.u.

Zarrabian et al. has pointed out [17] for the $(\text{H}_2)_2$ calculation at the square form that the accuracy of the energy began to decrease from the 16th order, but, in our calculation no such behavior is found and the energy converged smoothly. This difference would come from the difference of the zeroth order N -electron functions.

We have also investigated the cases where the perturbation expansion is slowly divergent by further increasing the bond distances $R_{\text{H}_2-\text{H}_2}$, $R_{\text{C}-\text{H}}$, $R_{\text{Li}-\text{H}}$. The divergence occurs caused mainly by the inadequacy of the model space to the excited states. The change of the norms are shown in Fig. 3, where the results only by the RS-PT method are shown, since other methods give similar results. The norms decrease at first several orders, but, they do not decrease further at later orders, which is in contrast to the convergent cases shown in Fig. 2. The results by Zarrabian et al. noted above corresponds to this case.

It is interesting that even in the case of divergence the norm decreases at lower orders. This appears in the case of CH_2 most clearly. The norm decreases monotonically up to 11th order, and there the energies become very close to the exact values; errors are 1.2×10^{-6} a.u. and -4.4×10^{-5} a.u. for the ground and the excited states, respectively. The CSF functions in the model spaces are not sufficient, which is the cause of the divergence; they, however, remain ones of the important functions in these elongated geometries. Then the norm decreases at the lower orders. Of course, at higher orders non-existence of other important

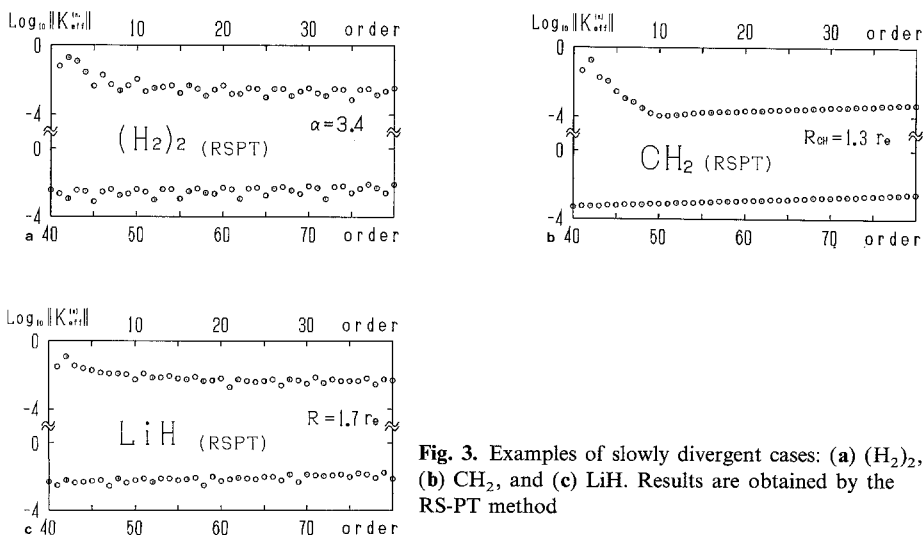


Fig. 3. Examples of slowly divergent cases: (a) $(\text{H}_2)_2$, (b) CH_2 , and (c) LiH . Results are obtained by the RS-PT method

functions in the model space becomes critical, and hence the divergence takes place.

Finally, we note that the above characteristic feature of the behavior of the norm is unaffected by the truncation of the CSF space from the full one, such as in the calculations of $(\text{H}_2)_2$ and CH_2 . We truncated the full CSF space in the LiH calculation in a similar manner to the calculations of $(\text{H}_2)_2$ and CH_2 , and found essentially the same behavior of the norms. The numerical stability of the present results are confirmed by artificial decreases of numerical accuracy at all arithmetic operations in a computer by forcing the last several bits to zero.

5 Concluding remarks

We have proposed to use the norm of the n th order contribution to the effective Hamiltonian for the analysis of the convergence property of the multireference many-body perturbation theory (MR-MBPT), and applied to the Rayleigh–Schrödinger (RS), Kirtman–Certain–Hirschfelder (KCH), and the canonical Van Vleck (CVV) perturbation theories for the calculations of $(\text{H}_2)_2$, CH_2 , and LiH molecules with the Epstein–Nesbet partitioning of the total Hamiltonian. We find that

(1) The lower order truncation of perturbation series gives reasonable solutions as far as the norm of the higher order contribution decrease continuously, namely in the convergent case, as the order n increases, or does not if the norm is almost constant or increases, namely in the divergent case,

(2) The change of the norm seems to have two stages, one is up to about the sixth order and the decrease of the norm is more rapid than in the other stage. In this former stage the norm decreases *even* in the divergent case, and the feature of the series being divergent becomes apparent at the later stage, which also takes place even with the (2×2) Hamiltonian in equation (4).

(3) The above three perturbation theories, namely, Rayleigh–Schrödinger PT, Kirtman–Certain–Hirschfelder PT, and the canonical Van Vleck PT, have essentially the same convergence property both at the convergent and the divergent cases.

It is interesting to apply the present method to the perturbation calculations with the Møller–Plesset partitioning [26] of the total Hamiltonian, and/or non-Hartree–Fock orbitals, such as the orbitals of V^{N-1} [27], V^{N-2} [28] potential, or MCSCF orbitals. We expect the present method, which utilizes the norm of the n th order contribution to the effective Hamiltonian, allows us to analyze the convergence property clearer than by the usual method using the eigenvalues of the effective Hamiltonian.

Acknowledgements. Part of the present calculations is carried out by the use of COLUMBUS [29]. All the calculations were performed on a FACOM VP-400 and a VP-2600 at the Data Processing Center of Kyoto University.

References

1. Brueckner KA (1955) Phys Rev 97:1353; (1955) Phys Rev 100:36
2. Goldstone J (1957) Proc R Soc (London) Ser A 239:267
3. Kelly HP (1969) Adv Chem Phys 14:129
4. Brandow BH (1967) Rev Mod Phys 39:771

5. Kvasnička V, Hubač I (1974) *J Chem Phys* 60:4483
6. Kvasnička V, Holubec A, Hubač I (1974) *Chem Phys Lett* 24:361
7. Lindgren I (1978) *Int J Quantum Chem, Quantum Chem Symp* 12:33
8. Hose G, Kaldor U (1979) *J Phys B* 12:3827
9. Shavitt I, Redmon LT (1981) *J Chem Phys* 73:5711
10. Bartlett RJ (1981) *Ann Rev Phys Chem* 32:359
11. Redmon LT, Bartlett RJ (1982) *J Chem Phys* 76:1938
12. Kutzelnigg W (1982) *J Chem Phys* 77:3081; Kutzelnigg W, Koch S (1983) *J Chem Phys* 79:4315; Kutzelnigg W (1984) *J Chem Phys* 80:822
13. Frisch MJ, Head-Gordon M, Trucks GW, Foreman JB, Schlegel HB, Raghavachari K, Robb M, Binkley JS, Gonzalez C, Defrees DJ, Fox DJ, Whiteside RA, Seeger R, Melius CF, Baker J, Martin RL, Kahn LR, Stewart JJP, Topiol S, Pople JA, GAUSSIAN 90, Gaussian Inc, Pittsburgh, PA
14. Knowles PJ, Somasundram K, Handy NC, Hirao K (1985) *Chem Phys Lett* 113:8
15. Handy NC, Knowles PJ, Somasundram K (1985) *Theor Chim Acta* 68:87
16. Hose G (1986) *J Chem Phys* 84:4505
17. Zarrabian S, Paldus J (1990) *Int J Quantum Chem* 38:761
18. Zarrabian S, Laidig WD, Bartlett RJ (1990) *Phys Rev A* 41:4711
19. Kirtman B (1968) *J Chem Phys* 49:3890
20. Certain PR, Hirschfelder JO (1970) *J Chem Phys* 52:5977; (1970) *J Chem Phys* 53:2992
21. Hirschfelder JO (1978) *Chem Phys Lett* 54:1
22. Primas H (1961) *Helv Phys Acta* 34:331; Primas H (1963) *Rev Mod Phys* 35:710
23. Jankowski K, Paldus J (1980) *Int J Quantum Chem* 18:1243
24. Tatewaki H, Huzinaga S (1980) *J Comput Chem* 1:205
25. Epstein PS (1926) *Phys Rev* 28:695; Nesbet RD (1955) *Proc Roy Soc (London)* A230:312
26. Møller C, Plesset S (1934) *Phys Rev* 46:618
27. Kaldor U (1973) *Phys Rev A* 7:427
28. Salomonson S, Lindgren I, Martensson AM (1980) *Phys Scr* 21:1351
29. COLOMBUS, the graphical unitary group approach CI by I. Shavitt, IMS Computer Center Library Program No. 0444. The programs have been adapted for IBM by F. Brown and for the IMS HITAC by I. Shavitt. Other changes and additions have been made by P. Laidig, B. Saxe, F. Brown and I. Shavitt, and the registration on IMS by S. Obara and S. Yamamoto. COLUMBUS contains Ohio state SCF/CI programs based on the R. M. Pitzer integrals/SCF programs and the Berkeley (B.R. Brooks and H. F. Schaefer) graphical unitary group approach CI programs