Intermediate Hamiltonians as a new class of effective Hamiltonians

1985 J. Phys. A: Math. Gen. 18 809

(http://iopscience.iop.org/0305-4470/18/5/014)

J. Phys. A: Math. Gen. 18 (1985) 809-826. Printed in Great Britain

Intermediate Hamiltonians as a new class of effective Hamiltonians

J P Malrieu, Ph Durand and J P Daudey

Laboratoire de Physique Quantique†, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France

Received 16 July 1984, in final form 12 October 1984

Abstract. The theory of effective Hamiltonians is well established. However, limitations appear in its applicability for many problems in molecular physics and quantum chemistry. The standard effective Hamiltonians may become strongly non-Hermitian when there is a large coupling between the model space, in which they are defined, and the outer space. Moreover, in the presence of intruder states, discontinuities appear in the matrix elements of these effective Hamiltonians as a function of the internuclear distances. To solve these difficulties, a new class of effective Hamiltonians (called intermediate Hamiltonians) is presented; only one part of their roots are exact eigen energies of the full Hamiltonian. The theory of these intermediate Hamiltonians is presented by means of a new wave-operator R which is the analogue of the wave-operator Ω in the theory of effective Hamiltonians. Solutions are obtained by a generalised degenerate perturbation theory (GDPT) and by iterative procedures. Two model systems are numerically solved which demonstrate the good convergence properties of GDPT with respect to standard degenerate perturbation theory (DPT). Continuity of the solutions is also checked in the presence of an intruder state.

1. Introduction

The effective Hamiltonian concept is a powerful tool for obtaining rigorous modelisation of complex problems in physics and quantum chemistry (see e.g. Brandow 1977, Kvasnička 1977, Freed 1977, Maynau *et al* 1983, Malrieu *et al* 1984). An effective Hamiltonian (Bloch 1958, Lindgren and Morrison 1982) acts in a low N_m-dimensional model space spanned by an *a priori* selected set of useful configurations or determinants K. The projection operator on the model space is

$$P_0 = \sum_K |K\rangle\langle K|$$

where the set $|K\rangle$ is an orthonormal basis.

The Bloch effective Hamiltonian H_{eff}, which is defined in the model space must achieve two tasks:

(i) the N_m roots of H_{eff} must be eigenvalues of the exact Hamiltonian H;

(ii) the corresponding eigenvectors must be the projections on the model space of the corresponding eigenvectors of H:

$$H_{\text{eff}} = \sum_{m=1}^{N_m} E_m |\tilde{\psi}_m\rangle\langle\tilde{\psi}_m^{\perp}|$$

† Unite de recherche Associée au CNRS no 505.

0305-4470/85/050809+18$02.25

while

$$H|\psi_m\rangle = E_m|\psi_m\rangle \qquad \text{and} \qquad |\bar{\psi}_m\rangle = P_0|\psi_m\rangle.$$

The states $|\psi_m\rangle$ are those having the largest components in the model space. The $|\bar{\psi}_m\rangle$ are the projections of $|\psi_m\rangle$ in the model space and $|\bar{\psi}_m^{\perp}\rangle$ is the biorthogonal state of $|\bar{\psi}_m\rangle$. The projected states $|\bar{\psi}_m\rangle$ are usually non-orthogonal and the corresponding effective Hamiltonian is non-Hermitian (Bloch 1958). Hermitian effective Hamiltonians can be obtained by orthogonalising the projected states, for instance by the symmetric $S^{-1/2}$ transformation (des Cloizeaux 1960).

These two conditions fulfilled by the effective Hamiltonian are severe: once the model space has been chosen, the effective Hamiltonian carries the useful information without any significant degree of freedom. Various methods are available for building effective Hamiltonians. The most explicit ones are various versions of the well known degenerate or quasi degenerate perturbation theory: (DPT and QDPT, respectively) (Bloch 1958, Brandow 1967, Lindgren 1974, Jørgensen 1975, Shavitt and Redmon 1980). The effective Hamiltonian is obtained through recurrence equations and order-by-order expansions (see § 3). These perturbative algorithms face difficult convergence problems as soon as there is not a well defined energy gap between the zeroth-order energies of the model space and those of the complementary outer space. The coupling between the model space and the outer space may be too large with respect to the energy denominators and the series will diverge. Frequently this problem is linked to the appearance of an intruder state which turns out to be nearly degenerate with some states of the model space.

These convergence difficulties may be overcome by using non-perturbative techniques (Suzuki and Lee 1980, Durand 1982, 1983). Starting from the Bloch equation Durand (1983) proposed a variational iterative scheme. Numerical tests (Maynau *et al* 1983) have demonstrated the high convergence performances of this algorithm which is able to trace the roots having the largest components in the model space around a curve crossing associated with an intruder state. However, this success made visible a qualitative defect of the effective Hamiltonians for modelisation purposes where the transferability of effective interactions is looked for as will be exemplified in § 1; the main difficulty consists in a discontinuity of the standard effective Hamiltonians around curve crossings which will make their use difficult for transfering effective interactions from small systems (diatomic molecules, for example) to larger systems (clusters of atoms, solids).

The above discussion leads to a global criticism of the excessive ambition of the usual effective Hamiltonians which are supposed to bring exact information on all their eigenvectors and eigenvalues, while some basic vectors may be introduced in the model space only because of their large coupling with the most important vectors or for logical consistency reasons in view of transferability to larger systems. The eigenvectors having large components in the model space are not necessarily interesting from a physical point of view and § 2 will discuss an alternative strategy which consists in building new effective Hamiltonians (hereafter called intermediate Hamiltonians) which are spanned by a $(N_m + N_i)$-dimensional subspace but which only delivers N_m exact eigenvalues and the corresponding exact projections in the $(N_m + N_i)$-dimensional model subspace. The N_i remaining solutions of the intermediate Hamiltonian will not be directly related to exact solutions of the total Hamiltonian. The basic theory of intermediate Hamiltonians will be presented in § 4. The whole theory will be based on a new R wave-operator which plays the same role in the theory of intermediate

Hamiltonians as the usual wave-operator Ω does in the theory of effective Hamiltonians. From the basic equation the derivation of the intermediate Hamiltonian will be given either by a GDPT or by efficient iterative procedures.

Section 5 gives two numerical tests of the method. The convergence properties of GDPT with respect to standard DPT will be demonstrated on a meaningful model. A model of an avoided crossing between a vector of the intermediate space and a vector of the outer space will show that the matrix elements of the intermediate Hamiltonians are continuous functions of the interatomic distances while those of the effective Hamiltonians are discontinuous. A final discussion will provide a connection between the concept of intermediate Hamiltonians and previous attempts for building 'dressed' Hamiltonians.

2. Defects of effective Hamiltonians or why does one need intermediate Hamiltonians

Let us assume that one wants to derive from accurate *ab initio* calculations (for example on the molecule Li_2) an effective Hamiltonian restricted to the valence model space. This model space will be defined from two orthogonal atomic orbitals a and b, one on each atom and looking like the 2s atomic orbitals of separated atoms and spanned by four determinants, two neutral determinants $|a\bar{b}|$, $|b\bar{a}|$ and two ionic ones $|a\bar{a}|$ and $|b\bar{b}|$. The orbitals 1s are kept frozen or treated through a pseudopotential. The use of either the QDPT or an iterative procedure (Durand 1983) leads to a minimal basis set valence effective Hamiltonian, slightly non-Hermitian, the upper matrix elements of which are

$\lvert a\bar{b}\rvert$	$\lvert b\bar{a}\rvert$	$\lvert a\bar{a}\rvert$	$\lvert b\bar{b}\rvert$
E_0	K_{ab}	F_{ab}	F_{ab}
	E_0	F_{ab}	F_{ab}
		$E_0+\Delta E$	K'_{ab}
			$E_0+\Delta E.$

The matrix elements of this four-dimensional effective Hamiltonian closely resemble those of the corresponding Hubbard (1963) and CNDO (Pople 1953, Pariser and Parr 1963) semi-empirical Hamiltonians:

$$H^{\mathrm{CNDO}}=\sum_{pq}(a_p^+a_q+a_{\bar{p}}^+a_{\bar{q}})\beta_{pq}+\tfrac{1}{2}\sum_{pq}(\hat{n}_p-1)(\hat{n}_q-1)\gamma_{pq}, \qquad \hat{n}_p=a_p^+a_p+a_{\bar{p}}^+a_{\bar{p}}$$

$$H^{\mathrm{Hubbard}}=\sum_{pq}(a_p^+a_q+a_{\bar{p}}^+a_{\bar{q}})\beta_{pq}+\sum_p Ua_p^+a_{\bar{p}}^+a_{\bar{p}}a_p.$$

One may be tempted to define the semi-empirical hopping integrals β_{pq} as the Fock-like F_{ab} terms of the aforementioned matrix and the bielectronic term from the energy difference ΔE between the neutral and ionic situations:

$$U=\Delta E \qquad \gamma_{pp}-\gamma_{pq}=\Delta E.$$

Freed and coworkers (Freed 1974, Iwata and Freed 1974) have followed this promising way in trying to establish the theoretical foundations of semi-empirical methods in quantum chemistry, especially for π systems of conjugated molecules and for deriving parameters such as β_{pq}, γ_{pq}, U from *ab initio* calculations (Iwata and Freed 1974, Lee and Freed 1983, Lee *et al* 1983).

The diagonal term E may be defined from the energy of the unique triplet state $^3\Sigma_u^+$ spanned by the model space

$$E(^3\Sigma_u^+) = E_0 - K$$

while the term ΔE can easily be defined from the energy of the $^1\Sigma_u^+$ singlet $2^{-1/2}[|a\bar{a}| - |b\bar{b}|]$ which is ionic in nature

$$E(^1\Sigma_u^+) = E_0 + \Delta E - K'.$$

The exchange integrals K and K' being rather small and almost equal one has

$$\Delta E \simeq E(^1\Sigma_u^+) - E(^3\Sigma_u^+).$$

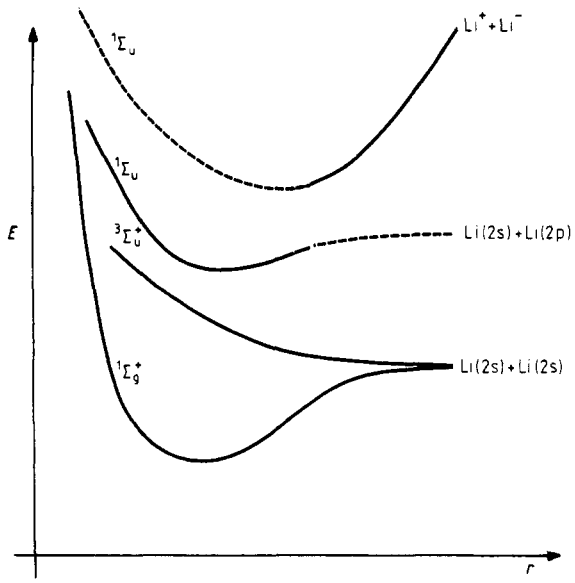

Figure 1. Low-lying energies of the lowest Σ states of Li_2 as a function of the internuclear distance r (schematic). Full lines represent the three eigen energies of the valence effective Hamiltonian.

Now if one examines the physical content of the interactions between the model space and the outer space one may notice that while at short interatomic distances (near the ground state equilibrium distance) the lowest $^1\Sigma_u^+$ state is actually dominated by the valence ionic situations, for large interatomic distances, the lowest $^1\Sigma_u^+$ is mainly neutral and involves one of the $2p_z$ atomic orbitals z_a and z_b:

$$^1\Sigma_u^+ \simeq \tfrac{1}{2}[|z_a\bar{b}| - |\bar{z}_a b| - |\bar{a}z_b| + |a\bar{z}_b|].$$

A crossing occurs (figure 1) between these two $^1\Sigma_u^+$ configurations since the ionic one dissociates into $Li^+ + Li^-$ at the energy IP − EA (ionisation potential minus the electroaffinity of the atoms) while the other dissociate into Li(2p) + Li(2s) which is much lower in energy. In the region of the curve crossing (Olson and Konowalow 1977, Konowalow and Olson 1979):

(i) the QDPT diverges, the neutral state made up of 2s and 2p orbitals on sites a and b acting as an intruder state;

(ii) the use of an iterative procedure (Durand 1983) would lead to a discontinuous Hamiltonian (see figure 1) since one of its roots would jump in the region of the curve

crossing from the lowest $^1\Sigma_u^+$ eigen energy of the exact problem at short distances to the second lowest eigenvalue $^1\Sigma_u^+$ at larger distances. The corresponding effective operators especially the ΔE term (U in the Hubbard formalism) or the hopping integral F_{ab} (or β) would be discontinuous functions of the interatomic distance r. This discontinuity is not unique since many other curve crossings occur between the ionic configuration $^1\Sigma_u^+$ and other Rydberg configurations and between the ionic $^1\Sigma_g^+$ state and Rydberg configurations of the same symmetry, each of them bringing a new discontinuity in the r dependence of the effective operators. These discontinuities are meaningless if one wants to investigate more complex systems, for example Li_n, since in Li_n the 2p or Rydberg-ionic crossing have no reason to take place at the same interatomic distance. Let us also note that the Rydberg states disappear in the central part of the cluster. Finally all this information is irrelevant for studying the ground state of a cluster and the related properties such as conformations and energies. One should be satisfied with some sort of effective (intermediate) Hamiltonian built on both neutral and ionic valence states which would give accurately the energy of the ground state and of the 'neutral' states in the sense of the valence bond theory while the ionic excited states could be approximately described.

The preceding discussion shows that the usual effective Hamiltonians face two types of problems both qualitative (divergence of the perturbation expansion) and quantitative (discontinuities for a diatomic molecule in the r dependence of the effective operator) due to curve crossings between the model space configurations with intruder states belonging to the complementary or outer space. These defects are consequences of the effective Hamiltonians which requires that all their roots be exact eigenvalues of the full problem. This formidable task can only be achieved when the model space is energetically well separated from the outer space, in other cases the effective Hamiltonians appear as strongly non-Hermitian with awkward and meaningless discontinuities resulting in a poor transferability. These difficulties will be solved by the introduction of the new concept of intermediate Hamiltonian. The theory of the usual effective Hamiltonians will be recalled in § 2 and the theory of the new intermediate Hamiltonians will be presented in § 3.

3. Main results in the theory of effective Hamiltonians

The vectorial space associated with the exact Hamiltonian H is partitioned into a N_m-dimensional model subspace S_0 and its orthogonal complement, the outer subspace $S_0^\perp$. The projectors associated with S_0 and $S_0^\perp$ are P_0 and Q_0, respectively

$$P_0 = \sum_{m=1}^{N_m} |m\rangle\langle m|, \qquad Q_0 = \sum_{\alpha} |\alpha\rangle\langle\alpha|, \qquad P_0 + Q_0 = 1. \tag{1}$$

Hereafter Latin and Greek letters will refer to the model and to the outer space, respectively.

The most basic non-Hermitian effective Hamiltonian (figure 2) (Bloch 1958) is defined by

$$H_{\text{eff}} = P_0 H \Omega. \tag{2}$$

Ω is the wave-operator which obeys the intermediate normalisation:

$$\Omega = \Omega P_0 = P_0 + Q_0 \Omega. \tag{3}$$

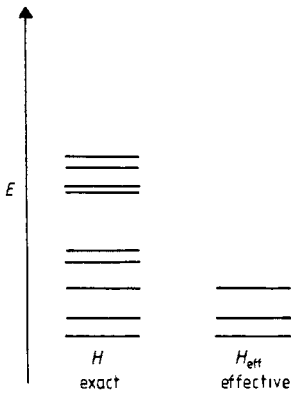

Figure 2. The exact Hamiltonian H and the effective Hamiltonian H_{eff} have the same eigen energies in a stable subspace S and in the model space S_0, respectively.

The wave-operator Ω transforms the projected states $P_0\psi_m$ which are solutions of H_{eff} back into the corresponding exact eigenstates ψ_m of H:

$$\Omega P_0|\psi_m\rangle = \Omega|\psi_m\rangle = |\psi_m\rangle. \tag{4}$$

The wave-operator obeys the operator equation

$$Q_0 H\Omega = Q_0\Omega H\Omega. \tag{5}$$

Equation (5) can be solved by various iterative schemes (Durand 1983) or by the quasi degenerate perturbation theory (QDPT). In the latter case the exact Hamiltonian H is split into an unperturbed zeroth-order Hamiltonian H_0 and a perturbation V

$$H_0 = \sum_{m=1}^{N_m} E_0|m\rangle\langle m| + \sum_{\alpha} E_\alpha^0|\alpha\rangle\langle\alpha| \tag{6}$$

where for sake of simplicity, we decided to choose H_0 as degenerate in the model space, limiting the further expansion to the DPT. This is not an important restriction since it is always possible to partition H into H_0 which is degenerate in the model space and the perturbation

$$V = H - H_0. \tag{7}$$

This definition of H_0 implies that if $\Sigma_{m=1}^{N_m}|m\rangle\langle m|H|m\rangle\langle m|$ is not strictly degenerate then there will be non-zero diagonal matrix elements of V. This choice of a degenerate unperturbed Hamiltonian in the model space greatly simplifies the perturbation expansions of Ω and H_{eff}. These simplifications will also remain valid for the intermediate Hamiltonians.

The resolution of equation (5) by the DPT leads to the expansion

$$\begin{aligned}
&\Omega = \sum_{n=0}^{\infty} \Omega^{(n)} \\
&\Omega^{(0)} = P_0 \\
&\cdots \\
&\Omega^{(1)} = gVP_0 \\
&\Omega^{(n)} = g\left(V\Omega^{(n-1)} - \sum_{k=1}^{n-1} \Omega^{(k)} V\Omega^{(n-k-1)} \right)
\end{aligned} \tag{8}$$

where

$$g = \frac{Q_0}{E_0 - H_0} = \sum_{\alpha} \frac{|\alpha\rangle\langle\alpha|}{E_0 - E_\alpha^0}. \tag{9}$$

The perturbation expansion of the effective Hamiltonian corresponding to Ω is

$$\begin{aligned} H_{\text{eff}} &= \sum_{n=0}^{\infty} H_{\text{eff}}^{(n)} \\ H_{\text{eff}}^{(0)} &= P_0 H_0 \\ H_{\text{eff}}^{(1)} &= P_0 V P_0 \\ &\cdots \\ H_{\text{eff}}^{(n)} &= P_0 V \Omega^{(n-1)}. \\ &\cdots \end{aligned} \tag{10}$$

Expansion of H_{eff} up to third order can be found in Lindgren (1982).

4. Definition and theory of intermediate Hamiltonians

4.1. General equations

Figure 3 gives an illustration of the main characteristics of an intermediate Hamiltonian. In contrast with usual effective Hamiltonians, an intermediate Hamiltonian denoted as $\tilde{H}$ has only N_m exact eigenenergies and projected eigenvectors in the $(N_m + N_i)$-dimensional model space. In other words we accept that a non-zero part of the spectrum of $\tilde{H}$ be only approximate solutions of H (see broken lines on figure 3).

Let us first split the full basis set in which the exact Hamiltonian is initially given in three classes.

The first class of vectors noted $|m\rangle$ the number of which is N_m defines the main model (sub)space S_m. The projector on this space is

$$P_m = \sum_{m=1}^{N_m} |m\rangle\langle m|. \tag{11}$$

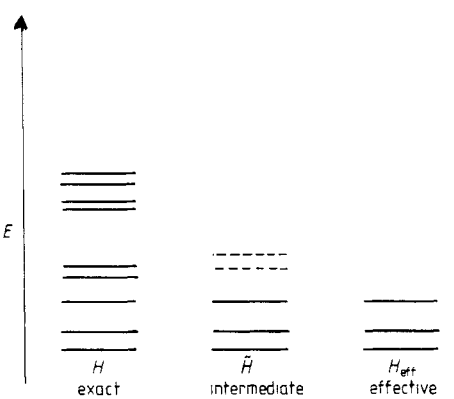

Figure 3. The exact Hamiltonian, the intermediate Hamiltonian $\tilde{H}$ and the effective Hamiltonian have the same eigen energies in a stable subspace S and in the main model subspace, respectively. Moreover $\tilde{H}$ has also approximate eigenvalues (broken lines) in the intermediate subspace.

A second class of N_i intermediate vectors denoted by $|i\rangle$ span the intermediate (sub)space S_i and the associated projector is

$$P_i = \sum_{i=1}^{N_i} |i\rangle\langle i|. \tag{12}$$

The direct sum of the main model space and of the intermediate space characterises the full model space $S_0 = S_m \oplus S_i$. In the following this space will be called the model space. The intermediate Hamiltonian will also be defined in this model space corresponding to the projector

$$P_0 = P_m + P_i. \tag{13}$$

The other vectors, denoted $|\alpha\rangle$, span the complementary or outer (sub)spaces $S_0^\perp$. The projector associated with $S_0^\perp$ is

$$Q_0 = \sum_\alpha |\alpha\rangle\langle\alpha|, \qquad P_0 + Q_0 = 1. \tag{14}$$

We want to build an intermediate Hamiltonian $\tilde{H}$ restricted to the model space (the direct sum of the main model subspace and of the intermediate subspace). This $(N_m + N_i)$-dimensional Hamiltonian should provide N_m significant roots

$$\tilde{H}|\tilde{\psi}_m\rangle = E_m|\tilde{\psi}_m\rangle, \qquad H|\psi_m\rangle = E_m|\psi_m\rangle, \qquad m = 1, 2 \ldots N_m \tag{15}$$

which should be exact eigenvalues of the exact Hamiltonian H. The N_i other eigenvalues of $\tilde{H}$ should only be approximate eigenvalues of H. There is obviously a lot of freedom in this definition of the intermediate eigenfunctions $\tilde{\psi}_m$. It seems obvious that their components on the main model subspace should be identical to the projections $P_m\psi_m$ of the exact solutions ψ_m. But one can ask more, namely that the components on the whole model space be also correct

$$|\tilde{\psi}_m\rangle = (P_m + P_i)|\psi_m\rangle. \tag{16}$$

This means that N_m solutions of $\tilde{H}$ are the projection of N_m exact solutions of H. This requirement is rather severe and was not compulsory.

Moving back to the spectral decomposition of $\tilde{H}$ it is clear that the above conditions are not sufficient for defining it in a unique way

$$\tilde{H} = \sum_{m=1}^{N_m} |\tilde{\psi}_m\rangle E_m \langle\tilde{\psi}_m^\perp| + \sum_{i=1}^{N_i} |\tilde{\psi}_i\rangle \tilde{E}_i \langle\tilde{\psi}_i^\perp| \tag{17}$$

since we have characterised neither the eigenstates $\tilde{\psi}_i$ which should have their largest components in the intermediate subspace nor their eigenvalues $\tilde{E}_i$. In (17) $\tilde{\psi}_m^\perp$ and $\tilde{\psi}_i^\perp$ are the biorthogonal vectors of $\tilde{\psi}_m$ and $\tilde{\psi}_i$ in the model space. It is clear that in taking into account the various degrees of freedom for defining $\tilde{H}$ one can produce many varieties of intermediate Hamiltonians. In the present paper we will restrict ourselves to those intermediate Hamiltonians which are direct generalisations of the standard Bloch Hamiltonian. They can be written in the form:

$$\tilde{H} = P_0 H R, \qquad R = R P_0. \tag{18}$$

Let us first recall that $P_0 = P_m + P_i$. The operator R looks like a wave-operator acting in the model space. The operator R is a nonorthogonal projector which obeys the

intermediate normalisation

$$R^2 = R, \qquad R = P_0 + Q_0 R. \tag{19}$$

It can easily be shown from the above conditions that $\tilde{H}$ defined by (18) can be considered as an intermediate Hamiltonian if and only if the following conditions are fulfilled:

$$RP_0|\psi_m\rangle = |\psi_m\rangle, \qquad m = 1, 2 \ldots N_m. \tag{20}$$

In (20) ψ_m is an exact solution of H. These conditions mean that R transforms the projected states $P_0\psi_m$ in the model space back to the exact solutions of H. Moreover it can immediately be checked that

$$\tilde{H}P_0|\psi_m\rangle = P_0HRP_0|\psi_m\rangle = P_0H|\psi_m\rangle = E_mP_0|\psi_m\rangle \tag{21}$$

which means that the $P_0\psi_m$'s are eigensolutions of $\tilde{H}$ with the exact energies E_m.

Conditions (20) can easily be transformed in a unique operator equation

$$R\Omega = \Omega \tag{22}$$

where Ω is the wave-operator associated with the Bloch effective Hamiltonian acting in the main model space

$$H_{\text{eff}} = P_mH\Omega \tag{23}$$

where

$$\Omega = \Omega P_m = P_m + (P_i + Q_0)\Omega \tag{24}$$

obeys the wave-operator equation (Durand 1983)

$$(P_i + Q_0)H\Omega = (P_i + Q_0)\Omega H\Omega. \tag{25}$$

Equation (22) is the basic equation of the theory of intermediate Hamiltonians. Multiplying both sides of (22) from the left by Q_0 and by taking into account the intermediate normalisation of R and Ω ((19) and (24)) gives

$$Q_0R(P_m + P_i)(P_m + P_i\Omega + Q_0\Omega) = Q_0\Omega. \tag{26}$$

We finally obtain

$$Q_0R(P_m + P_i\Omega) = Q_0\Omega. \tag{27}$$

Figure 4 gives a matrix representation of equation (27) which clearly indicates the operators which are effectively involved in the operator equation $R\Omega = \Omega$ (22).

Equation (27) can also be written in the form

$$Q_0RP_m = Q_0(1 - RP_i)\Omega. \tag{28}$$

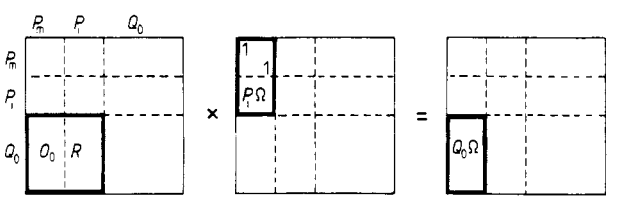

Figure 4. Matrix representation of the operators involved in the operator equation $Q_0R \times (P_0 + P_i\Omega) = Q_0\Omega$ (see (22), (27), (28)).

This expression indicates that the projection Q_0RP_i of Q_0R on the right by P_i is completely arbitrary and that once Q_0RP_i has been chosen, Q_0RP_m is determined in a unique way from (28). The simplest choice is obviously

$$Q_0RP_i = 0 \qquad \Rightarrow \qquad Q_0RP_m = Q_0\Omega. \tag{29}$$

This choice provides a very simple intermediate Hamiltonian which has, however, the drawback of being non-Hermitian at second order in a perturbation series. It is more useful in view of building more Hermitian $\tilde{H}$'s and transferable effective interactions to derive R from wave-operator-like equations. The simplest generalisation of the wave-operator equation (5) appears to be

$$Q_0HR = Q_0\Omega HR. \tag{30}$$

In contrast with the wave-operator equation (25), equation (30) is linear in the unknown operator R. This linearity in R will be maintained for other equations that will be given below. As required, it can easily be shown (see appendix 2) that R as a solution of (30) fulfils the basic condition $R\Omega = \Omega$. It is useful to generalise (30) by introducing an arbitrary energy parameter E_0:

$$Q_0(E_0 - H)R = Q_0\Omega(E_0 - H)R. \tag{31}$$

As previously one can show that R defined by (31) fulfils the condition $R\Omega = \Omega$. The operator R is now weakly E_0 dependent and the value to be given to the parameter E_0 will be discussed below. Multiplying both sides of (31) on the right by Ω, and using $\Omega^2 = \Omega$ and $R\Omega = \Omega$, leads also to

$$Q_0H\Omega = Q_0\Omega H\Omega \tag{32}$$

which is a projection on the left by Q_0 of the wave-operator equation (25). By taking into account the intermediate normalisation of R (19), equation (31) can be exactly solved:

$$Q_0R = \frac{Q_0}{E_0 - (1-\Omega)H}[H + \Omega(E_0 - H)]P_0. \tag{33}$$

$Q_0/E_0 - (1-\Omega)H$ is the inverse of the operator $E_0 - (1-\Omega)H$ restricted to the outer space. For actual situations this operator cannot be directly inverted. Otherwise one does not need Q_0R but only its projection Q_0RP_i on the right by P_i since the complementary operator Q_0RP_m is then determined by (28). Multiplying both sides of (31) on the right by P_i provides an equation for Q_0RP_i:

$$Q_0(E_0 - H)RP_i = Q_0\Omega(E_0 - H)RP_i. \tag{34}$$

The linear equations (31) and (34) can easily be solved either by generalised Newton-Raphson methods similar to those that were previously derived for the wave-operator in the theory of effective Hamiltonians or by a direct generalisation of the quasi degenerate perturbation expansion.

4.2. Generalised degenerate perturbation theory (GDPT)

The exact Hamiltonian is split into an unperturbed zero-order Hamiltonian H_0 and a perturbation $V = H - H_0$:

$$H_0 = \sum_{m=1}^{N_m} E_0|m\rangle\langle m| + \sum_{i=1}^{N_i} E_i^0|i\rangle\langle i| + \sum_{\alpha} E_\alpha^0|\alpha\rangle\langle\alpha|. \tag{35}$$

As previously in § 2, it is assumed that H_0 is degenerate in the main model subspace. For deriving compact expansions of R in powers of V which do not imply H_0, E_0 in (31) has to be identical to the E_0 which appears in the expression (35) of H_0. Then (31) becomes

$$Q_0(E_0 - H)R = Q_0 \Omega VR. \tag{36}$$

This equation has a nice structure since H_0 has disappeared on its right-hand side. It can also be written in the form

$$Q_0 R = \frac{Q_0}{E_0 - H_0}(1 - \Omega) VR \tag{37}$$

which is suitable for a perturbation expansion:

$$\begin{aligned} R &= \sum_{n=0}^{\infty} R^{(n)} \\ R^{(0)} &= P_0 \\ R^{(1)} &= gVP_0 \\ &\cdots \\ R^{(n)} &= g\left(VR^{(n-1)} - \sum_{k=1}^{n-1} \Omega^{(k)} VR^{(n-k-1)} \right) \end{aligned} \tag{38}$$

where

$$g = \frac{Q_0}{E_0 - H_0} = \sum_{\alpha} \frac{|\alpha\rangle\langle\alpha|}{E_0 - E_\alpha^0}. \tag{39}$$

The expansion of the intermediate Hamiltonian (18) is obtained order-by-order from the expansion of R:

$$\begin{aligned} \tilde{H} &= \sum_{n=0}^{\infty} \tilde{H}^{(n)} \\ \tilde{H}^{(0)} &= H_0 P_0 \\ \tilde{H}^{(1)} &= P_0 V P_0 \\ \tilde{H}^{(2)} &= P_0 V g V P_0 \\ &\cdots \\ \tilde{H}^{(n)} &= P_0 V R^{(n-1)} \\ &\cdots . \end{aligned} \tag{40}$$

Expansions of $\tilde{H}$ up to the fifth order can be found in appendix 1.

4.3. Relationships between H, $\tilde{H}$ and H_{eff}

Figure 5 shows how one can pass from the exact Hamiltonian H which is defined in the full vectorial space to the intermediate Hamiltonian defined in the model space and finally to the Bloch effective Hamiltonian which is restricted to the main model subspace. H_{eff} can be derived either from H or from $\tilde{H}$ by means of the wave-operators

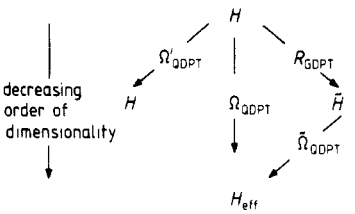

Figure 5. The intermediate Hamiltonian $\tilde{H}$ is derived from the exact Hamiltonian H by mean of the GDPT. The Bloch effective Hamiltonian can be determined either directly from H or from $\tilde{H}$ by the standard QDPT. This scheme clearly indicates the intermediate character of $\tilde{H}$ between and H_{eff}. H' is the exact effective Hamiltonian that could be defined in the full model space.

Ω and $\tilde{\Omega}$, respectively.

$$H_{eff} = P_m H \Omega = P_m \tilde{H} \tilde{\Omega}. \tag{41}$$

The new wave-operator $\tilde{\Omega}$ is solution of the equation

$$P_i \tilde{H} \tilde{\Omega} = P_i \tilde{\Omega} \tilde{H} \tilde{\Omega}. \tag{42}$$

The operator equation

$$\Omega = R \tilde{\Omega} \tag{43}$$

also illustrates how one can pass directly from the projected wavefunctions $P_m \psi_m$ in the main model subspace back to exact solutions either directly by Ω or indirectly through an intermediate step corresponding to the product $R\tilde{\Omega}$.

5. Numerical applications

The first comparative test between DPT and GDPT was made on a real symmetric matrix with five vectors in the main model space ($N_m = 5$), five vectors in the intermediate space ($N_i = 5$) and ten in the outer space ($N_\alpha = 10$) for GDPT and five vectors in the model space for DPT. The matrix elements h_{ij} are defined by

$$h_{ii} = 1, \qquad (1 \leqslant i \leqslant 5), \qquad h_{i-1,i} = 0.1, \qquad (2 \leqslant i \leqslant 5)$$

$$h_{ii} = 2 + \tfrac{1}{10}(i-6), \qquad h_{i-5,i} = 0.1 \qquad (6 \leqslant i \leqslant 10)$$

$$h_{ii} = i - 8, \qquad h_{i-10,i} = 0.1, \qquad h_{i-5,i} = h_{i-1,i} = 0.5 \qquad (11 \leqslant i \leqslant 20).$$

The results in table 1 illustrate the better convergence of GDPT. The origin of this improvement is evident by looking at the perturbation expressions given in appendix 1. The convergence rate is mainly determined by g_i which involves the smallest denominators $E_0 - E_i^0$ (equation (A1.4)). In GDPT, terms implying g_i appear at fourth order only. The convergence radius for GDPT must however be the same as for DPT, since the main equation determining the wave operator R implies the wave-operator Ω of DPT. Nevertheless this advantage of better results at low orders is not negligible for the practical use of the theory.

Table 1. Comparison of convergence between GDPT (left) and DPT (right) for the 20×20 test matrix defined in the text. At each order n, significant figures are given.

n	λ_1		λ_2		λ_3		λ_4		λ_5	
2	0.818	0.81	0.890	0.89	0.991	0.99	1.089	1.089	1.1628	1.162
3	—	0.82	—	0.891	0.9909	0.99	—	1.090	—	1.164
4	0.8182	0.818	0.890 04	0.890	0.990 940	0.990	1.0892	1.0892	1.16249	1.162
5	0.818 19	—	0.890 03	—	0.990 94	0.991	1.089 25	1.0893	1.162 87	1.163
6	—	0.8181	—	0.8900	—	—	—	—	—	1.1628
7	0.818 186	0.8182	—	—	0.990 9395	—	—	—	1.162 8691	1.1629
8	0.818 185	—	0.890 034	0.890 04	0.990 9394	0.9909	—	1.089 25	—	1.162 87
9	0.818 1853	0.818 19	0.890 035	—	—	—	1.089 248	—	—	—
10		0.818 18	—	0.890 036	—	0.990 93	—	—	1.162 8692	—
11		—	—	—	—	0.990 94	—	—		—
12		—	0.890 0353	—	0.990 39	—	—	—		—
13		0.818 186		0.890 035		0.990 940	—	1.089.248		—
14		0.818 185		—		—	1.089 2475	—		—
15		—		—		—		—		
16		—		0.890 0354		0.990 0395		—		1.162 869
17		—		—		—		—		—
18		0.818 1853		—		—		1.089 2476		1.162 8693
19				—		0.990 9394		1.089 2475		1.162 8692
20				0.890 0353		0.990 9394				
exact	0.818 1853		0.890 035 37		0.990 9394		1.089 2475		1.162 8692	

In the second test, we used a model 3 by 3 matrix of the type:

	m	i	α
m	0	λ	μ
i	λ	1	γ
α	μ	γ	x.

In this example $N_m = N_i = N_\alpha = 1$ and the convergence rate is directed by the ratio $\lambda/1$ or μ/x, both for DPT (which reduces in that case to standard Rayleigh Schrödinger perturbation theory) and for GDPT. The convergence curves obtained with $0.1 \leqslant \lambda \leqslant 0.6$, $\mu = \gamma = 0.1$, and $x = 1.1$ are displayed in figure 6.

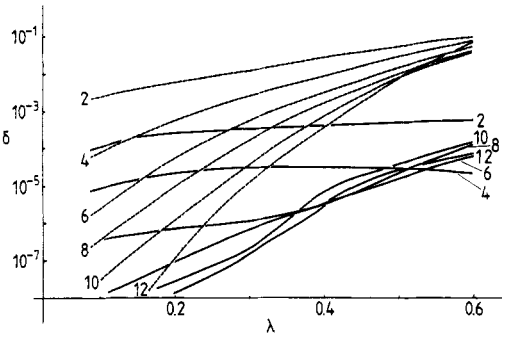

Figure 6. Absolute value of error $\delta = |E(\lambda) - E^{(p)}(\lambda)|$ where $E(\lambda)$ is the exact eigenvalue and $E^{(p)}(\lambda)$ the value obtained at order p either with DPT (broken lines) or with GPDT (full lines). λ is the strength of the coupling between main model space and intermediate space.

At a given order of perturbation, the error increases as a function of λ are much more slower in GDPT. For instance at $\lambda = 0.3$ the fourth order GDPT result is equivalent to the tenth order of DPT. Even for large λ ($0.5 \leqslant \lambda \leqslant 0.6$) for which divergences appear in both treatments (as expected) see preceeding paragraph) low orders of GDPT results are quite good, far better than DPT. The behaviour of the different orders of perturbation in GDPT is no longer regular for $\lambda > 0.35$ where there is a crossing between the curves corresponding to orders 6 and 8.

By varying x in the vicinity of 1, the same matrix may be used as a model for the intruder state situation discussed in detail in § 2. Let us suppose that for physical reasons we want to build a 2×2 effective Hamiltonian on the first two basic states. The second root of the effective Hamiltonian corresponds to the second root of the 3×3 matrix for $x > 1$ and to the third root for $x < 1$. The 2×2 effective Hamiltonian is discontinuous, and this discontinuity in the second root is reflected by a discontinuity in the extradiagonal elements and the second diagonal element. If we are interested in the first eigenstate and eigenvalue, built on two basic states, this discontinuity of the second root gives no useful information and introduces highly undesirable variations in the effective matrix elements. With GDPT algorithm, precise information on the second root is lost but no discontinuity appears in the matrix elements (see figure 7). All the elements of $\tilde{H}$ are smooth functions of x, well adapted to interpolation purposes.

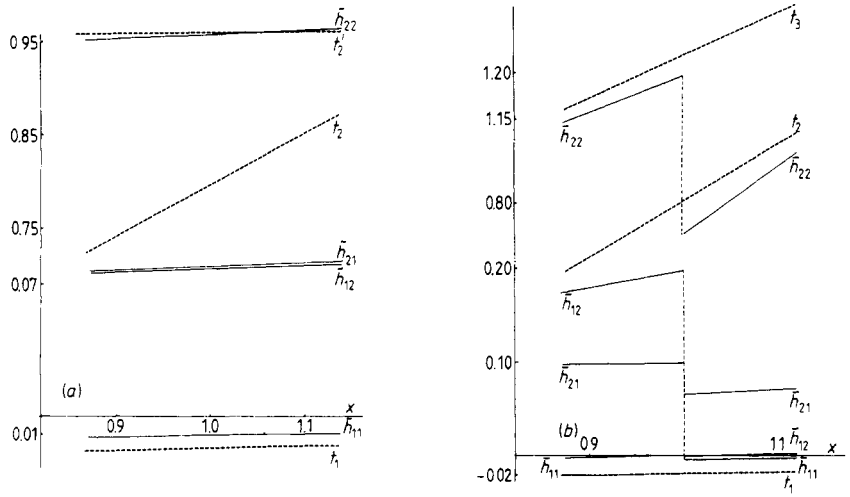

Figure 7. (*a*) Variation of the intermediate Hamiltonian matrix elements ($\tilde{h}_{ij}$), exact eigenvalues (t_1, t_2) of H exact and second approximate eigenvalue of $\tilde{H}(t_2')$. (*b*) Variations of the effective Hamiltonian matrix elements ($\bar{h}_{ij}$) and exact eigenvalues (t_1, t_2, t_3) of H exact.

6. Conclusion

The present work proposes a new class of effective operators which contains the usual effective operators as a special case. General wave operator equations have been derived, together with an order-by-order perturbative expansion, which contains the usual Rayleigh Schrödinger (quasi) degenerate perturbation theory as a particular case. The basic idea of our intermediate Hamiltonians is to be less ambitious than the traditional effective Hamiltonians which are asked to possess a maximal information on all their roots and eigenvectors. The present method rests on a partition of the model space into a main subspace (of dimension N_m), bearing the larger components of the eigenvectors of physical interest, and an intermediate subspace (of dimension N_i), which is introduced for consistency reasons, or for its large interaction with the main subspace. The resulting intermediate Hamiltonians deliver N_m exact roots and the corresponding eigenvectors are the projections of exact eigenvectors into the whole model space, the components on both the main and intermediate subspaces being exact. The N_i other roots, although they are not meaningless in most cases, are distorted with respect to the eigensolutions having their largest components into the intermediate subspace.

Speaking in terms of dressed Hamiltonians, one may say that the intermediate vectors are dressed by their interaction with the outer space, but this dressing is not intended to give the corresponding eigenvectors, their dressing is simply performed to give the correct components of the N_m vectors essentially spanned by the main model subspace on the intermediate subspace, as evident from the fact that if i and j are two vectors of the intermediate subspaces in the second order expression

$$\langle i|\tilde{H}|j\rangle = \langle i|H|j\rangle + \sum_{\alpha} \frac{\langle i|V|\alpha\rangle\langle\alpha|V|j\rangle}{E_0 - E_\alpha},$$

the energy denominators refer to the mean zeroth-order energy of the main vectors.

One may notice that this second-order expression is identical to the effective operator introduced by Davidson *et al* (Nitzsche *et al* 1978, Davidson *et al* 1981) under the name of 'shifted B_k approximation'. (This terminology refers to an early work by Gershgorn and Shavitt (1968), which proposes a specific approximation of the partitioning technique; one should also mentioned Meyer (1971, 1973).) The shifted B_k approximation was limited to the *second* order and to a main subspace reduced to *one* vector. It has been introduced to maintain a good dependence of the ground state correlation energy on the number of particles, which is insured by the non-degenerate Rayleigh-Schrödinger perturbation theory, and which disappears in both the original B_k approximation (due to its Brillouin-Wigner structure) and in the usual QDPT at the same order, as demonstrated by Malrieu (1982) in a convincing model problem (N non-interacting H_2 molecules, in a minimal basis set, the model space consisting of the ground state configuration plus all doubly excited determinants). The present work extends Davidson's treatment to higher orders and to the simultaneous treatment of several states. Its relevance for CI problems and for the search of the lowest solutions of Heisenberg Hamiltonians will be illustrated in the future.

Its main goal is however different since it concerns the rigorous modelisation and the construction of *transferable* effective Hamiltonians though a careful extraction of information from high accuracy calculations on small molecules. This philosophy is well illustrated by the work of Freed and coworkers (π Hamiltonians, valence electron effective Hamiltonians) and by the work of our group on effective spin Hamiltonians; the expected conceptual advantages of our intermediate Hamiltonians are:

(i) a better continuity of the effective operators around the curve crossings between intermediate and outer vectors, as illustrated in a model calculation in the last section;

(ii) a better hermiticity which will be documented in details on the derivation of effective magnetic Hamiltonians for (poly)acetylenic systems.

On a fundamental level, the field open by the present work deserves supplementary efforts; the above-proposed intermediate Hamiltonians are weakly energy dependent, since the N_i roots associated with the intermediate subspace depend on the chosen E_0 value. This dependence will be studied in a further paper. The main problem concerns the possible variety of the intermediate Hamiltonians; one type of intermediate Hamiltonian has been proposed above, but other versions may certainly be derived, the relative advantages of which should be explored.

As another extension, one should generalise our approaches to the cases where H_0 is not taken as degenerate in the main model space, in order to build a generalised QDPT. This extension would lead to more complex formulae, but one should recall that the present algorithms are already able to treat *quasi* degenerate problems.

Appendix 1. Generalised degenerate perturbation expansion (GDPT) of intermediate Hamiltonians

The entire vectorial space is split into three parts: the main model subspace, the intermediate space and the outer space the projectors of which are P_m, P_i, Q_0, respectively

$$P_m = \sum_{m=1}^{N_m} |m\rangle\langle m| \qquad P_i = \sum_{i=1}^{N_i} |i\rangle\langle i| \qquad Q_0 = \sum_{\alpha} |\alpha\rangle\langle\alpha| \tag{A1.1}$$

$$P_m + P_i + Q_0 = P_0 + Q_0 = 1. \tag{A1.2}$$

The unperturbed Hamiltonian, degenerate in the main model subspace, is denoted

$$H_0 = \sum_{m=1}^{N_m} E_0 |m\rangle\langle m| + \sum_{i=1}^{N_i} E_i^0 |i\rangle\langle i| + \sum_{\alpha} E_\alpha^0 |\alpha\rangle\langle\alpha|. \tag{A1.3}$$

Two Green operators will appear in the expansions:

$$g = \frac{Q_0}{E_0 - H_0} = \sum_{\alpha} \frac{|\alpha\rangle\langle\alpha|}{E_0 - E_\alpha^0} \qquad \text{and} \qquad g_i = \frac{P_i}{E_0 - H_0} = \sum_{i} \frac{|i\rangle\langle i|}{E_0 - E_i^0}. \tag{A1.4}$$

$$\begin{aligned}
\tilde{H} &= \sum_{n=0}^{\infty} \tilde{H}^{(n)} \\
\tilde{H}^{(0)} &= P_0 H_0 \\
\tilde{H}^{(1)} &= P_0 V P_0 \\
\tilde{H}^{(2)} &= P_0 V g V P_0 \\
\tilde{H}^{(3)} &= P_0 [V g V g V - V g^2 V P_m V] P_0 \\
\tilde{H}^{(4)} &= P_0 [V g V g V g V - V g V g^2 V P_m V - V g^2 V (g + g_i) V P_m V - V g^2 V P_m V g V \\
&\quad + V g^3 V P_m V P_m V] P_0 \\
\tilde{H}^{(5)} &= P_0 [V g V g V g V g V - V g V g V g^2 V P_m V - V g V g^2 V (g + g_i) V P_m V \\
&\quad - V g^2 V (g + g_i) V (g + g_i) V P_m V - V g V g^2 V P_m V g V \\
&\quad - V g^2 V (g + g_i) V P_m V g V - V g^2 V P_m V g V g V + V g V g^3 V P_m V P_m V \\
&\quad + V g^2 V (g + g_i)^2 V P_m V P_m V + V g^3 V (g + g_i) V P_m V P_m V \\
&\quad + V g^2 V P_m V g^2 V P_m V + V g^3 V P_m V (g + g_i) V P_m V \\
&\quad + V g^3 V P_m V P_m V g V - V g^4 V P_m V P_m V P_m V] P_0.
\end{aligned} \tag{A1.5}$$

The DPT and the GDPT expansions have similar stuctures. Note that the reduced Green operator acting in the intermediate space (g_i) only appears at the fourth-order of perturbation.

Appendix 2. Properties of the operator R

R is assumed to be the solution of (30):

$$Q_0 H R = Q_0 \Omega H R. \tag{A2.1}$$

By assuming the intermediate normalisation (19)

$$R = P_m + P_i + Q_0 R, \tag{A2.2}$$

we will demonstrate that R as a solution of (A2.1) fulfils the basic relation $R\Omega = \Omega$ (22). Multiplying both sides of (A2.1) on the right by Ω gives

$$Q_0 H R \Omega = Q_0 \Omega H R \Omega. \tag{A2.3}$$

(A2.2) and $R\Omega = \Omega$ allows us to put $R\Omega$ as

$$R\Omega = (P_m + P_i + Q_0 R)(P_m + P_i\Omega + Q_0\Omega) = P_m + P_i\Omega + Q_0 R\Omega. \tag{A2.4}$$

Then (A2.3) becomes

$$Q_0 H(P_m + P_i\Omega + Q_0 R\Omega) = Q_0\Omega H(P_m + P_i\Omega + Q_0 R\Omega). \tag{A2.5}$$

Equation (A2.5) shows that $Q_0 R\Omega$ is unique. The wave operator equation for Ω

$$Q_0 H\Omega = Q_0\Omega H\Omega \tag{A2.6}$$

can also be written in the form:

$$Q_0 H(P_m + P_i\Omega + Q_0\Omega) = Q_0\Omega H(P_m + P_i\Omega + Q_0\Omega). \tag{A2.7}$$

The comparison of (A2.5) and (A2.7) finally leads to

$$Q_0 R\Omega = Q_0\Omega \tag{A2.8}$$

$$R\Omega = P_m + P_i\Omega + Q_0 R\Omega = P_m + P_i\Omega + Q_0\Omega = \Omega. \tag{A2.9}$$

It can also be shown by a similar demonstration that R as a solution of (32) also fulfills the condition $R\Omega = \Omega$.

References

Bloch C 1958 *Nucl. Phys.* **6** 329

Brandow B H 1967 *Rev. Mod. Phys.* **39** 771

—— 1977 *Advances in Quantum Chemistry* vol 10, ed P O Löwdin (New York: Academic) p 187

Davidson E R, McMurchie L B and Day S J 1981 *J. Chem. Phys.* **74** 5491

des Cloizeaux J 1960 *Nucl. Phys.* **20** 321

Durand Ph 1982 *J. Physique Lett.* **43** L461

—— 1983 *Phys. Rev.* A **28** 3184

Freed K F 1977 *Semi-empirical methods of electronic structure calculation, Part A: Techniques* ed G A Segal (New York: Plenum)

Freed K 1974 *J. Chem. Phys.* **60** 1765

Gershgorn Z and Shavitt I 1968 *Int. J. Quant. Chem.* **2** 751

Hubbard J 1963 *Proc. R. Soc.* A **276** 238

Iwata S and Freed K F 1974 *J. Chem. Phys.* **61** 1500

Jørgensen F 1975 *Mol. Phys.* **29** 1137

Konowalow D D and Olson M L 1979 *J. Chem. Phys.* **71** 450

Kvasnička V 1977 *Advances in Chemical Physics* vol XXXVI, ed I Prigogine and S A Rice (New York: Wiley) p 345

Lee Y S and Freed K F 1983 *Chem. Phys. Lett.* **94** 202

Lee Y S, Freed K F, Sun H and Yeager D L 1983 *J. Chem. Phys.* **79** 3862

Lindgren I 1974 *J. Phys. B: At. Mol. Phys.* **7** 2441

Lindgren I and Morrison J 1982 *Atomic many-body theory* (Berlin: Springer) ch 9

Malrieu J P 1982 *Theor. Chim. Acta* **62** 163

Malrieu J P, Maynau D and Daudey J P 1984 *Phys. Rev.* B **30** 1817

Maynau D, Saïd M and Malrieu J P 1983 *J. Am. Chem. Soc.* **105** 5244

Maynau D, Durand Ph, Daudey J P and Malrieu J P 1983 *Phys. Rev.* A **28** 3193

Meyer W 1971 *Int. J. Quant. Chem.* **5** 341

—— 1973 *J. Chem. Phys.* **58** 1017

Nitzsche L E and Davidson E R 1978a *J. Chem. Phys.* **68** 3103

—— 1978b *J. Am. Chem. Soc.* **100** 7201

Olson M L and Konowalow D D 1977 *Chem. Phys.* **21** 393

Pariser R and Parr R G 1963 *J. Chem. Phys.* **21** 466, 767

Pople J 1953 *Trans. Faraday Soc.* **49** 1375

Shavitt I and Redmon L T 1980 *J. Chem. Phys.* **73** 5711

Suzuki K and Lee S Y 1980 *Prog. Theor. Phys.* **64** 2091