Direct determination of effective Hamiltonians by wave-operator methods. I. General formalism

Philippe Durand
Laboratoire de Physique Quantique (Equipe de Recherche No. 821 associée au Centre National de la Recherche Scientifique), Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse Cedex, France
(Received 29 March 1983)

The determination of the most standard effective Hamiltonians is obtained by means of a simple general similarity transformation. The wave operator is shown to be a solution of an operator equation which is the analog of the Møller equations of scattering theory and which generalizes those previously established by Bloch, Löwdin, Jørgensen, and Lindgren. The wave-operator equation is solved by efficient iteration or perturbation-iteration methods which exhibit good convergence properties for degenerate systems and/or in presence of intruder states. In the following paper the method is applied to the theoretical determination of transferable effective-spin interactions.

INTRODUCTION

The concept of effective interaction and effective Hamiltonian plays a central role in the field of nuclear structure and for studying the electronic and magnetic structure of atoms, molecules, and solids.[1-5] Despite much progress that has been made in the last twenty years, difficulties remain both at the theoretical and computational levels.

Although there is a general agreement that the most fundamental effective Hamiltonians are those that were previously found by Bloch[6] and des Cloizeaux,[7] there is apparently no direct elementary way for deriving these effective Hamiltonians. The most transparent derivation seems to have been given by Brandow[2] in the framework of the quasidegenerate perturbation theory. Brandow emphasizes the central position played by the wave operator from which all effective Hamiltonians can easily be found. Up to now various operator equations have been proposed for determining this wave operator. The original Bloch's equation[6] for exact degenerate states was first generalized by Jørgensen for quasidegenerate systems. A more general equation has been obtained by Lindgren.[8] However, this equation cannot be considered as the most general one since it is based explicitly on an unnecessary partition of the exact Hamiltonian into an unperturbed Hamiltonian and a perturbation. An important contribution of this paper will be to show that there is in fact a very general canonical equation for the wave operator which is the analog of the Møller equations of scattering theory.

Another problem in the field of effective Hamiltonians is to find efficient procedures for calculating realistic effective interactions. Within the framework of one-electron methods progress has already been made in that direction in determining transferable atomic potentials in molecules and polymers.[9] Beyond one-electron methods, two- and three-body electronic interactions in molecules have, for example, been investigated with success by Freed and co-workers.[10] However, many important problems, generally outside the capabilities of one-electron methods, remain unsolved. For transition elements there is no direct way for computing accurately the multiplets arising from the quasidegenerate $3d^n4s^1$ and $3d^{n-1}4s^2$ electronic configurations. In the field of magnetism there are no standard methods for deriving from first principles the various effective spin Hamiltonians in molecules and solids. Perhaps the most difficult problems with which we have to deal with are those arising in studies of metals, surfaces, atoms adsorbed on surfaces, etc. For these very degenerate systems there are no standard accurate methods for studying the main geometrical and energetic properties associated with the ground and excited states. For most of the realistic applications the standard Rayleigh-Schrödinger perturbation method diverges. Although powerful methods for improving the convergence were previously derived such as partial summation techniques and Padé approximants, the problem of how to calculate realistic effective Hamiltonians is still an open problem. The most typical situation can be found in many-body-perturbation theory where diagrammatic and partial summations to all orders cannot be considered as the final development of the theory. Thus in the present paper we will present systematic algebraic iterative methods for determining wave operators. These methods will be developed in close relationship with the usual linear and quadratic (Newton-Raphson) methods of numerical analysis. These procedures will then be applied to the theoretical determination of spin effective interactions in the following paper, paper II.

The first section will introduce the notations and give the general similarity transformation for obtaining the standard effective Hamiltonians. Section II is devoted to the determination of a general equation for the wave operator Ω. Efficient iterative schemes for determining Ω will be presented in Sec. III.

I. A GENERAL SIMILARITY TRANSFORMATION FOR DERIVING EFFECTIVE HAMILTONIANS

The solutions of an exact Hamiltonian H define a Hilbert space which can be split into the subspace S of exact solutions in which we are interested in and the orthogonal

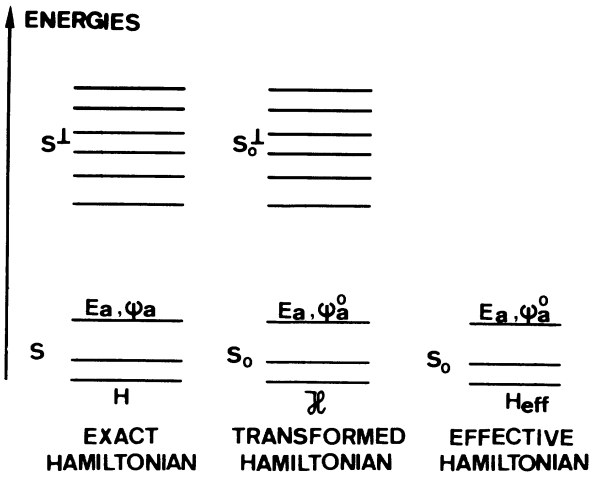

FIG. 1. Exact Hamiltonian H, the transformed Hamiltonian $\mathcal{H}=U^{-1}HU$, and the effective Hamiltonian $H_{\text{eff}}=P_0\mathcal{H}P_0$ have the same energies in the subspace S and in the model space S_0, respectively.

subspace $S^\perp$. The projection operators associated with S and $S^\perp$ are P and Q, respectively,

$$P^2=P,\quad Q^2=Q,\quad P+Q=1\ . \tag{1}$$

The Hilbert space can also be split into a model subspace S_0 spanned by approximate solutions of H and the orthogonal subspace $S_0^\perp$. The projection operators associated with S_0 and $S_0^\perp$ are P_0 and Q_0, respectively,

$$P_0^2=P_0,\quad Q_0^2=Q_0,\quad P_0+Q_0=1\ . \tag{2}$$

In order to derive effective Hamiltonians, H is first transformed by means of a similarity transformation:

$$\mathcal{H}=U^{-1}HU\ . \tag{3}$$

U has an inverse but it is not required to be unitary. It is also assumed that the tranformation U decouples $\mathcal{H}$ within the subspaces S_0 and $S_0^\perp$ (see Figs. 1 and 2):

$$\mathcal{H}=P_0\mathcal{H}P_0+Q_0\mathcal{H}Q_0\ . \tag{4}$$

It follows from (4) that an effective or model Hamiltonian can be defined by

$$H_{\text{eff}}=P_0\mathcal{H}P_0\ . \tag{5}$$

In order to go a step further, it is useful to consider the transformation

$$U=P(P_0PP_0)^{-\nu}+Q(Q_0QQ_0)^{-\nu}\ . \tag{6}$$

U depends on a non-negative index ν, the useful values of which will be given below. Operators $(P_0PP_0)^{-\nu}$ and $(Q_0QQ_0)^{-\nu}$ are defined in the entire Hilbert space. The nonzero matrix elements of $(P_0PP_0)^{-\nu}$ and $(Q_0QQ_0)^{-\nu}$ are defined from the matrix elements of P_0PP_0 and Q_0QQ_0 in the subspaces S_0 and $S_0^\perp$, respectively (Fig. 3). This definition assumes that P_0PP_0 and Q_0QQ_0 have inverse operators in S_0 and $S_0^\perp$. It can equivalently be stated that if the vectors ϕ_a and ϕ_i span S and $S^\perp$, the vectors $\phi_a^0=P_0\phi_a$ and $\phi_i^0=Q_0\phi_i$ which span S_0 and $S_0^\perp$ must be independent vectors. These conditions are not too severe even for strong perturbations.

It can easily be checked that the inverse of U is

$$U^{-1}=(P_0PP_0)^{\nu-1}P+(Q_0QQ_0)^{\nu-1}Q\ . \tag{7}$$

The main interest of the transformation U is that it leads immediately to most basic effective Hamiltonians. The value $\nu=0$ leads to

$$H_{\text{eff},\hat{O}}=(P_0PP_0)^{-1}PHP_0\ . \tag{8}$$

Letter $\hat{O}$ in $H_{\text{eff},\hat{O}}$ is the first letter of Ôkubo. It has been given in tribute to his original contribution in the field.[11] The value $\nu=1$ leads to the Bloch[6] effective Hamiltonian:

$$H_{\text{eff},B}=P_0HP(P_0PP_0)^{-1}\ . \tag{9}$$

$H_{\text{eff},\hat{O}}$ and $H_{\text{eff},B}$ are related to each other by a Hermitian transformation but are not in themselves Hermitian,

$$H_{\text{eff},B}=H_{\text{eff},\hat{O}}^\dagger\ . \tag{10}$$

The canonical Hermitian effective Hamiltonian first given by des Cloizeaux[7] corresponds to the intermediate value $\nu=\frac{1}{2}$:

$$H_{\text{eff,dC}}=(P_0PP_0)^{-1/2}PHP(P_0PP_0)^{-1/2}\ . \tag{11}$$

In our opinion the use of the operator U is the most direct and transparent way for deriving the three basic effective Hamiltonians. It does not seem that the importance of this transformation was previously recognized although

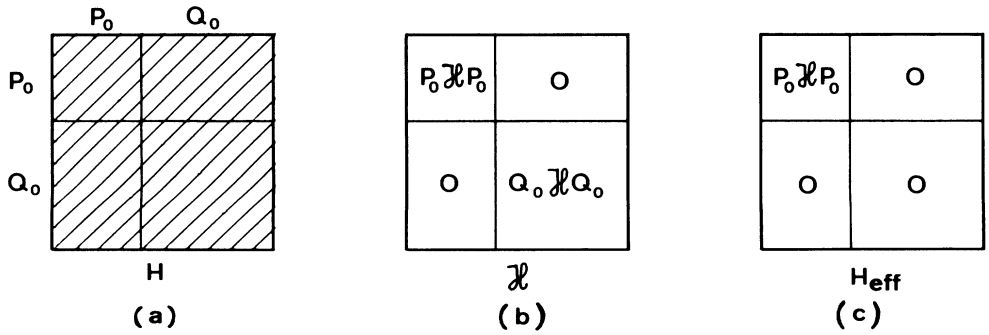

FIG. 2. Matrix representation of H (a), $\mathcal{H}$ (b), and H_{eff} (c) in the model space S_0 (projector P_0) and the orthogonal subspace (projector Q_0). Hatched part of (a) represents nonzero matrix elements.

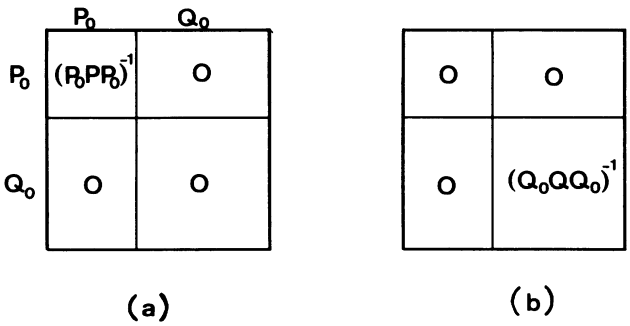

FIG. 3. Matrix representation of $(P_0PP_0)^{-1}$ (a) and $(Q_0QQ_0)^{-1}$ (b) in the model space (projector P_0) and in the orthogonal subspace (projector Q_0).

Jørgensen[12] used it for $\nu=\frac{1}{2}$ in conjunction with a unitary Van Vleck transformation.

Note that there is not always an obvious correspondance between the model subspace S_0 and the exact subspace S. This occurs, for example, in the presence of intruder states. This phenomenon is illustrated in Fig. 4, where one considers two possible ways for defining S from S_0 for a two-dimension effective Hamiltonian. In case (a), the two eigenvalues of H_{eff} are continuous but there is a discontinuity of the physical content of the eigensolutions. On the contrary, in case (b), the eigenvalues are discontinuous whereas the physical content of S remains stable. This means that for all values of parameter λ which is linearly related to the energy of the intruder state, the subspace S remains as similar as possible to the model subspace S_0. The second definition of S is obviously the most interesting one, especially for deriving transferable effective interactions. In fact, it will be seen in Sec. III of this paper and in paper II that the perturbation-iteration procedure presented below leads to an effective Hamiltonian corresponding to stable exact subspaces.

II. CANONICAL EQUATION FOR THE WAVE OPERATOR

A. The wave operator

The basic effective Hamiltonians given in Sec. I depend in a nonlinear and complicated way upon the projectors P_0 and P. However, it has been found useful from theoretical arguments and also for practical applications to introduce a wave operator, usually designated as Ω, which transforms the wave functions belonging to the model space S_0 back into the subspace S of the exact solutions. This wave operator was first introduced for the continuum part of the spectrum [Møller (Ref. 13)] and then for describing bound states [Eden and Francis (Ref. 14)]. In this paper we use the following definition:

$$\Omega=P(P_0PP_0)^{-1}\,. \tag{12}$$

This expression was previously given by Jørgensen (see also Ref. 15, p. 220). In order to get a general equation for Ω we first recall and specify some of its main properties. Multiplying both sides of Eq. (12) at left either by P_0 or by P gives the usual projection properties:

$$P_0\Omega=P_0,\quad P\Omega=\Omega\,. \tag{13}$$

The first relation corresponds to the so-called intermediate

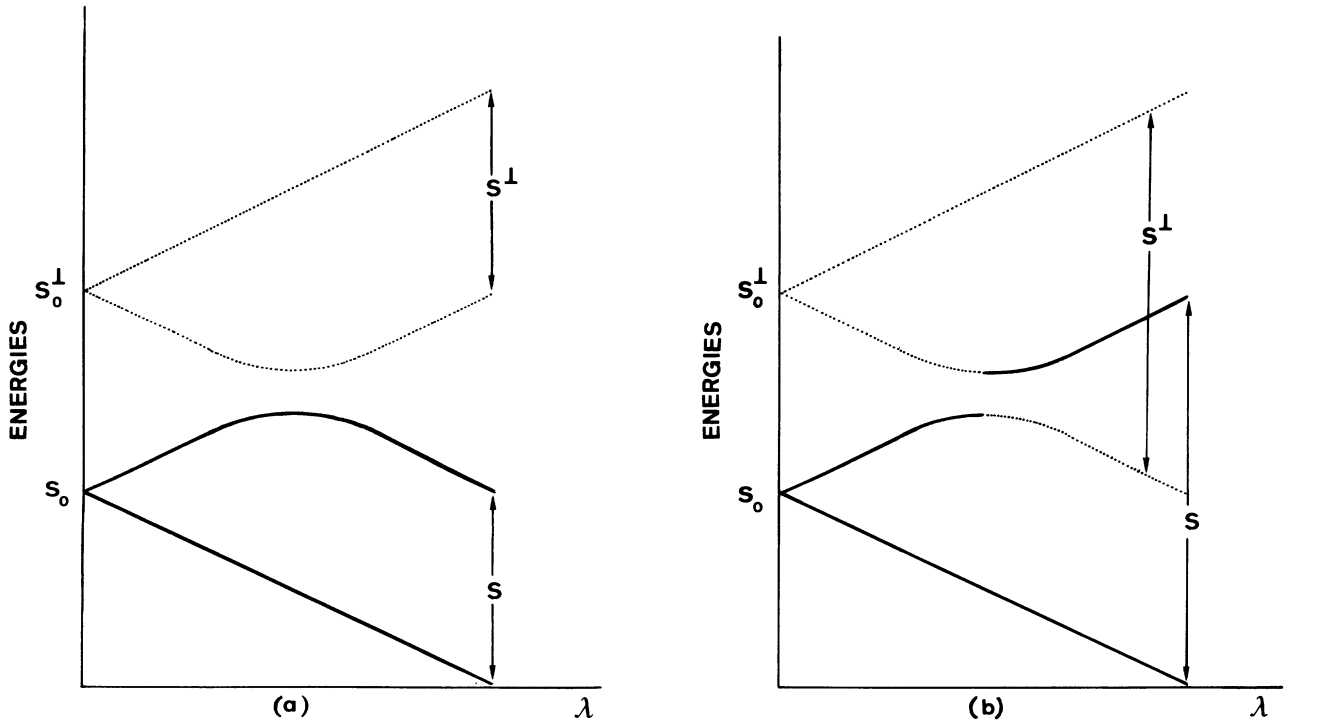

FIG. 4. Eigenvalues of a two-dimensional Hamiltonian in presence of an intruder state. The two eigenvalues of H_{eff} (solid lines) are represented as a function of an arbitrary parameter λ. (a) and (b) correspond to two different possible definitions of the exact subspace S.

normalization. Multiplying both sides of Eq. (12) at the right by P_0 and P leads similarly to

$$\Omega P_0=\Omega, \quad \Omega P=P \ . \tag{14}$$

The relation $\Omega P=P$ is established from the definition of Ω [Eq. (12)] and by noting that ΩP is a normal projector. A different demonstration was previously given by Jørgensen.[12] Other useful relations are

$$\Omega^{\dagger}\Omega=(P_0PP_0)^{-1}, \quad \Omega^2=\Omega \ . \tag{15}$$

Expressions (15) can immediately be established from Eq. (12):

$$\Omega^{\dagger}\Omega=(P_0PP_0)^{-1}P(P_0PP_0)^{-1}=(P_0PP_0)^{-1} \ , \tag{16}$$

$$\Omega^2=P(P_0PP_0)^{-1}P(P_0PP_0)^{-1}=P(P_0PP_0)^{-1}=\Omega \ .$$

$\Omega^2=\Omega$ and $\Omega^{\dagger}\neq\Omega$ mean that Ω is a nonorthogonal projector. By introducing Ω the three basic effective Hamiltonians can be written

$$H_{\mathrm{eff},B}=P_0H\Omega, \quad H_{\mathrm{eff},\hat{O}}=\Omega^{\dagger}HP_0 \ , \tag{17}$$

$$H_{\mathrm{eff,dC}}=(\Omega^{\dagger}\Omega)^{-1/2}\Omega^{\dagger}H\Omega(\Omega^{\dagger}\Omega)^{-1/2} \ . \tag{18}$$

The non-Hermitian Bloch and Ôkubo Hamiltonians depend linearly on Ω whereas the Hermitian des Cloizeaux Hamiltonian has a much more complicated dependence upon Ω. Since there is a direct way for passing from Ω to the above effective Hamiltonians we consider that our main objective is to find out a general equation for Ω.

B. Canonical equation for the wave operator

The commutation relation $[H,P]=0$ and Eqs. (13) and (14) allows us to write:

$$H\Omega=HP\Omega=PH\Omega=\Omega PH\Omega=\Omega HP\Omega=\Omega H\Omega \ . \tag{19}$$

Thus we have established the general operator equation

$$H\Omega=\Omega H\Omega \ . \tag{20}$$

By using $\Omega^2=\Omega$, this equation can also be written in the form

$$[H,\Omega]\Omega=0 \ . \tag{21}$$

Equation (20) generalizes wave-operator equations previously established by Bloch,[6] Löwdin,[16] Jørgensen,[12] and Lindgren.[8] For an exactly degenerate system and assuming that the exact energy E is known, Eq. (20) gives immediately Löwdin's[16] expression:

$$H\Omega=E\Omega \ . \tag{22}$$

Splitting up the exact Hamiltonian H into H_0 and V allows us to write Eq. (20) in the form

$$\begin{aligned}(H_0+V)\Omega&=\Omega(H_0+V)\Omega\\&=\Omega H_0(P_0+Q_0)\Omega+\Omega V\Omega\\&=\Omega H_0P_0\Omega+\Omega V\Omega\\&=\Omega H_0+\Omega V\Omega\end{aligned} \tag{23}$$

which can be cast into the form of an operator equation which was first established by Lindgren,[8]

$$[H_0,\Omega]=-V\Omega+\Omega V\Omega \ . \tag{24}$$

In contrast with Eqs. (23) and (24) which depend explicitly on the partition of H into H_0 and V, Eq. (20) does not depend on this partition. Equation (20) is uniquely determined from the knowledge of the model subspace S_0 and from S through the projectors P_0 and P. It results that Eq. (20) can be considered as a canonical equation for Ω. In view of its simplicity and compactness we consider that it is the best starting point for general developments and for practical applications in the field of effective interactions.

C. Reduced wave operator

For further developments and applications it is useful to introduce[17] a reduced wave operator X (noted χ by Lindgren[18])

$$X=Q_0\Omega \ . \tag{25}$$

Then

$$\Omega=(P_0+Q_0)\Omega=P_0+X \ . \tag{26}$$

Multiplying both sides of Eq. (20) from the left by P_0 and using $P_0\Omega=P_0$ gives the identity

$$P_0H\Omega=P_0H\Omega\equiv H_{\mathrm{eff},B} \ .$$

Multiplying both sides of Eq. (20) from the left by Q_0 and using Eq. (25) leads to the compact operator equation

$$Q_0(1-X)H(1+X)P_0=0 \ . \tag{27}$$

In the framework of perturbation theory Eq. (27) can be transformed into

$$[H_0,X]=-Q_0(1-X)V(1+X)P_0 \ . \tag{28}$$

The resolution of this commutator equation leads to the implicit operator equation,

$$X=\sum_{a\in S_0}\frac{Q_0}{a}(1-X)V(1+X)P_a \ . \tag{29}$$

The sum runs over the model subspace. P_a is the projector corresponding to the eigensolution $|a\rangle$ of H_0

$$P_a=|a\rangle\langle a|, \quad H_0|a\rangle=E_a^0|a\rangle \tag{30}$$

and the reduced resolvent Q_0/a is given by

$$\frac{Q_0}{a}=\frac{Q_0}{E_a^0-H_0} \ . \tag{31}$$

The operator equations (27) and (29) will be the starting point of two classes of iteration and perturbation-iteration methods that will be given in the next section.

III. ITERATIVE METHODS FOR SOLVING THE WAVE-OPERATOR EQUATION

Iterative methods for determining X will now be derived from Eqs. (27) and (29). The former equation is at the ori-

gin of iteration methods whereas Eq. (29) will lead to perturbation-iteration methods which are in close relationship with the standard Rayleigh-Schrödinger method.

A. Iteration methods

By introducing

$$F(X)=Q_0(1-X)H(1+X)P_0 , \tag{32}$$

Eq. (27) can formally be written as

$$F(X)=0 . \tag{33}$$

Let us first note that Eq. (33) is quadratic in the unknown operator X. For determining X, $F(X)$ is first linearized by expanding $F(X+\Delta X)$ in the neighborhood of X:

$$F(X+\Delta X)=F(X)+A\,\Delta X+\cdots , \tag{34}$$

$$A\,\Delta X=\Delta X\,H(1+X)P_0+Q_0(1-X)H\,\Delta X . \tag{35}$$

ΔX denotes a small variation of X and A is an operator or superoperator[19] acting in the vectorial space $\mathcal{G}$ of all transition operators coupling S_0 and $S_0^{\perp}$. Notations and results concerning this vectorial space can be found in Appendix A. In the particular case of a finite Hilbert space associated with the solutions of H, the matrix elements of A are given in Appendix B.

If we were able to invert A in $\mathcal{G}$, the variation ΔX of X at each iteration would be given by

$$\Delta X=-A^{-1}F(X) . \tag{36}$$

Equation (36) defines a Newton-Raphson iterative procedure which converges quadratically (see Appendix C for a comparison with the ordinary Newton-Raphson method applied to an algebraic equation). Unfortunately expression (36) cannot be used directly. The calculation of ΔX for a d-dimensional model space in a finite N-dimensional Hilbert space would imply solving a dN-dimensional linear system of equations which would require a computational time proportional to N^3. For practical applications various approximations must be made on A^{-1} which lead to quasiquadratic procedures and computational times proportional to N^2 at each iteration. The easiest way is to partition A into two terms:

$$A=A_0+\Delta A . \tag{37}$$

In a matrix representation A_0 is assumed to contain the main part of A. The expansion of A^{-1} in powers of ΔA gives

$$A^{-1}=A_0^{-1}-A_0^{-1}\,\Delta A\,A_0^{-1}+\cdots . \tag{38}$$

By using (38) one can write

$$\Delta X=\Delta X_0+\Delta X_1+\cdots+\Delta X_k+\cdots \tag{39}$$

with

$$\begin{aligned}
&\Delta X_0=-A_0^{-1}F(X) ,\\
&\Delta X_1=-A_0^{-1}\,\Delta A\,\Delta X_0 ,\\
&\cdots ,\\
&\Delta X_k=-A_0^{-1}\,\Delta A\,\Delta X_{k-1} ,\\
&\cdots .
\end{aligned} \tag{40}$$

For most realistic applications expansion (38) diverges. Development (39) has to be rescaled and we choose it of the form

$$\Delta X=C_0\,\Delta X_0+C_1\,\Delta X_1+\cdots+C_k\,\Delta X_k+\cdots . \tag{41}$$

Let us first note that for the exact solution one has

$$(F(X)\,|\,F(X))=0 . \tag{42}$$

In (42) we have used the scalar product notation between operators for which more details are given in Appendix A. In our iterative procedure the best coefficients C_k of (41) are determined at each iteration by minimizing

$$(F(X+\Delta X)\,|\,F(X+\Delta X)) . \tag{43}$$

Assuming that ΔX is small with respect to X and using the linearization of $F(X)$ in the neighborhood of X according to Eq. (34), the minimization of (43) can be replaced by

$$(F(X)+A\,\Delta X\,|\,F(X)+A\,\Delta X)_{\min} . \tag{44}$$

Expression (44) is quadratic in ΔX and the coefficents C_k in (44) are best determined by the standard Fourier method[9] that leads to the resolution of a system of linear equations:

$$\sum_l (A\,\Delta X_k\,|\,A\,\Delta X_l)C_l=-(A\,\Delta X_k\,|\,F(X)) . \tag{45}$$

B. Perturbation-iteration methods

Most of the results of the preceding paragraph are still valid but the operator equation to be solved is now replaced by

$$\begin{aligned}
&F(X)=X-f(X)=0 ,\\
&f(X)=\sum_a \frac{Q_0}{a}(1-X)V(1+X)P_a .
\end{aligned} \tag{46}$$

With respect to (46) the operator A is now defined by

$$A\,\Delta X=\Delta X+\sum_a \frac{Q_0}{a}[\Delta X\,V(1+X)-(1-X)V\,\Delta X]P_a . \tag{47}$$

In order to establish a relation between our perturbation-iteration methods and the standard perturbation methods, it is useful to introduce an operator C by

$$C=A^{-1} . \tag{48}$$

C will be called the convergence operator. Expression (47) shows immediately that C tends to the identity operator

for small perturbations. A reasonable approximation for small perturbations is then to choose

$$C=1 . \tag{49}$$

The variation $\Delta X^{(n)}=X^{(n)}-X^{(n-1)}$ of X at the nth iteration is now given by (for the linear method)

$$\Delta X^{(n)}=-F(X^{(n-1)}) . \tag{50}$$

Equation (50) defines an iterative procedure for determining X, simpler than the Rayleigh-Schrödinger expansion, which converges linearly for small perturbations. This terminology is borrowed from the standard field of linear and quadratic methods for solving ordinary algebraic equations (see Appendix C). From this point of view, the Rayleigh-Schrödinger method appears as a rather complicated quasilinear method. For larger perturbations C is far from being the identity operator and one has to use expansion (41) with a linear optimization of the C_k's.

C. Practical aspects

The perturbation-iteration method has been implemented on a computer for determining effective Hamiltonians of dimension up to $d=20$ in a finite 1000-dimensional Hilbert space. In most cases we took $A=1$ and limited the expansion (38) to its first term. Thus the variation of X at the nth iteration is

$$\Delta X^{(n)}=C_0\,\Delta X_0^{(n)}=-C_0\,F(X^{(n-1)}) . \tag{51}$$

According to (45) the best coefficient C_0 is at each iteration given by

$$C_0=\frac{(AF(X^{(n-1)})\,|\,F(X^{(n-1)}))}{(AF(X^{(n-1)})\,|\,AF(X^{(n-1)}))} . \tag{52}$$

We also found it useful, as in the Chebyshev iterative procedure for solving linear equations, to mix in at the nth iteration $\Delta X^{(n)}$ given by (51), the variation $\Delta X^{(n-1)}$ obtained at the $(n-1)$th iteration. In all cases examined the convergence of the procedure was improved but at the expense of a larger memory for keeping the numerical information associated with three successive iterations. Our procedure requires about dN^2 multiplications at each iteration. This number is to be compared with the N^2 multiplications needed by most iterative methods for calculating a proper solution of an N-dimensional matrix. Finally for very strong perturbations we have used a two-term expansion in Eq. (38). Damping and level shift techniques could also be used for particular cases. All the computational details will be published elsewhere.

In conclusion we have succeeded in determining effective Hamiltonians by quasi-Newton procedures. The efficiency of these methods can also be understood by considering that the resolution of Bloch-type equations is obtained by the minimization of the norm of $F(X)$. Thus our procedures are implicitly based on a variational principle on the wave functions whereas most of the standard methods of resolution of the Schrödinger equation are based on the Rayleigh-Ritz variational principle. It results from the different nature of these minimum principles that our methods are stable with respect to the physical content of the effective Hamiltonian: We can easily pass from the unperturbed model subspace with projector P_0 to the nearest exact subspace with projector P even in the presence of intruder states, as will be shown in paper II.

CONCLUSIONS

It has first been shown in this paper that the most basic effective Hamiltonians can easily be derived from a simple similarity transformation. In a second step we have established a general compact equation for the wave operator. This equation is likely to be the most general one in this field. In contrast with other similar equations it does not require any partition of the exact Hamiltonian into an unperturbed Hamiltonian and a perturbation. For this reason we suggest that this equation should be called the *canonical* wave-operator equation. For quantum mechanics of bound states and the corresponding effective Hamiltonians, this equation is the analog of the Møller equations of scattering theory. In the third part of this paper we have proposed generalized Newton-Raphson iterative methods for determining the wave operators. These methods appear as the natural generalization of the Rayleigh-Schrödinger method. The relationship between these methods and the usual perturbation theory is established by means of a convergence operator which tends to the identity operator for small perturbations. It will be shown in paper II that for spin effective interactions these methods have good convergence properties for highly degenerate systems even in the presence of intruder states.

ACKNOWLEDGMENT

The author thanks Professor Ph. R. Wallace for kindly reading and revising the manuscript.

APPENDIX A: THE VECTORIAL SPACE OF OPERATORS ACTING IN THE HILBERT SPACE

The set of all operators $A,B,C,\ldots$, acting in the Hilbert space spanned by the solutions of H define a new Hilbert space in which the scalar product between two operators A and B is defined by

$$(A\,|\,B)=\mathrm{Tr}(A^{\dagger}B) . \tag{A1}$$

The above notation with parentheses is used in order to avoid confusion with the scalar product notation in the ordinary Hilbert space. The operators or superoperators[19] acting in the new Hilbert space will also be denoted by capital letters $A,B,C,\ldots$.

In this paper we are mainly interested with the determination of a reduced wave operator X coupling the model subspace S_0 and its orthogonal complement $S_0^{\perp}$, the projectors of which are P_0 and Q_0, respectively. X can be expanded in basic operators:

$$X=\sum_{i\in S_0^{\perp}}\sum_{a\in S_0} x_{ia}\,|\,i,a) , \tag{A2}$$

$$|\,i,a)\equiv|\,i\rangle\langle a\,| . \tag{A3}$$

The $|i,a)$'s define an orthonormal basis set of transition operators coupling S_0 and $S_0^{\perp}$:

$$(i,a\,|\,j,b)=\delta_{ij}\delta_{ab}\ . \tag{A4}$$

With this notation (A2) becomes

$$X=\sum_{i\in S_0^{\perp}}\sum_{a\in S_0}|i,a)(i,a\,|\,X)\ . \tag{A5}$$

$\sum_{i,a}|i,a)(i,a|$ can be considered as the closure relation with respect to all operators coupling S_0 and $S_0^{\perp}$. Any linear operator A acting in the new Hilbert space can be defined by its matrix elements between the $|i,a)$'s:

$$A=\sum_{i,j\in S_0^{\perp}}\sum_{a,b\in S_0}|i,a)(i,a\,|\,A\,|\,j,b)(j,b|\ . \tag{A6}$$

(Note. In Appendixes A—C the letters a,b,c will denote the states belonging to the model space S_0 and the letters $i,j,k,\ldots$ will be used for the states belonging to its orthogonal complement $S_0^{\perp}$.)

APPENDIX B: MATRIX ELEMENTS OF A

For practical applications, one needs the matrix elements of A between the orthonormal set of transition operators defined in Appendix A. We consider successively operators A corresponding to the iteration and perturbation-iteration methods.

1. Iteration methods

From definition (35) it can easily be checked that the only nonzero matrix elements are those given by

$$(i,a\,|\,A\,|\,i,a)=\langle a\,|\,H(1+X)\,|\,a\rangle-\langle i\,|\,(1-X)H\,|\,i\rangle\ , \tag{B1}$$

$$(i,a\,|\,A\,|\,i,b)=\langle b\,|\,H(1+X)\,|\,a\rangle\ , \tag{B2}$$

$$(i,a\,|\,A\,|\,j,a)=-\langle i\,|\,(1-X)H\,|\,j\rangle\ . \tag{B3}$$

In (B1)—(B3) the unperturbed basic vectors $|a\rangle$ and $|i\rangle$ span the model space S_0 and its orthogonal complement, respectively.

It is to be noted that the matrix elements given by (B1)—(B3) define a very sparse matrix. Figure 5 displays the nonzero matrix elements corresponding to a finite Hilbert space of dimension $N=8$ and to a two-dimensional model space. The matrix representation of A is made up of a diagonal of full 2×2 matrices and, outside this diagonal, there are lines of nonzero elements which couple the six vectors belonging to $S_0^{\perp}$.

Assuming that $H=H_0+V$, (A1), (A2), and (A3) can be expanded in first order in V:

$$(i,a\,|\,A\,|\,i,a)=E_a^0-E_i^0+\langle a\,|\,V\,|\,a\rangle-\langle i\,|\,V\,|\,i\rangle\ , \tag{B4}$$

$$(i,a\,|\,A\,|\,i,b)=\langle b\,|\,V\,|\,a\rangle\ , \tag{B5}$$

$$(i,a\,|\,A\,|\,j,a)=-\langle i\,|\,V\,|\,j\rangle\ . \tag{B6}$$

E_a^0 and E_i^0 are the unperturbed energies of the states $|a\rangle$ and $|i\rangle$ belonging to S_0 and $S_0^{\perp}$, respectively.

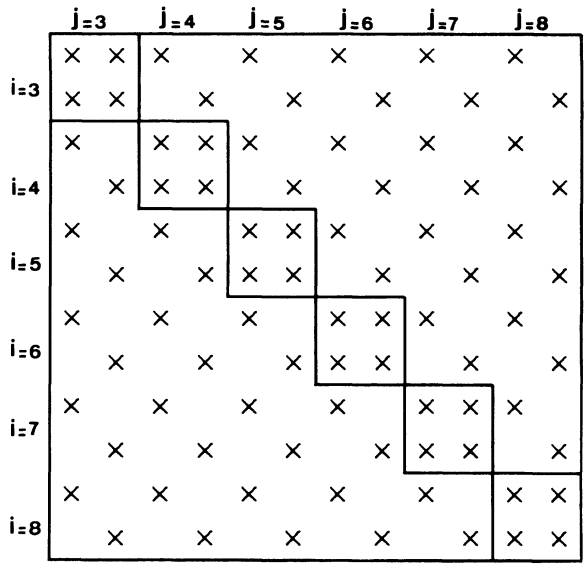

FIG. 5. Matrix representation of A in the basis set of the operators $|i,a)=|i\rangle\langle a|$; $a\in S_0$ and $i\in S_0^{\perp}$. Dimension of the Hilbert space, $N=8$. Dimension of the model space, $d=2$. Matrix representation of order $d(N-d)=12$. Nonzero matrix elements have been represented by crosses.

For small perturbations the matrices associated with A and A^{-1} are almost diagonal and the matrix elements of A^{-1} can be approximated by

$$(i,a\,|\,A^{-1}\,|\,j,b)=\frac{1}{E_a^0-E_i^0}\delta_{ab}\delta_{ij}\ . \tag{B7}$$

In the case of a one-dimensional model space Eqs. (37) and (B7) lead to variations of the wave function at each iteration which are quite similar to those used in the Davidson method.[20]

2. Perturbation-iteration methods

It results from (47) that the nonzero matrix elements of A are now given by

$$(i,a\,|\,A\,|\,i,a)=1+\frac{1}{E_a^0-E_i^0}[\langle a\,|\,V(1+X)\,|\,a\rangle-\langle i\,|\,(1-X)V\,|\,i\rangle]\ , \tag{B8}$$

$$(i,a\,|\,A\,|\,i,b)=\frac{1}{E_a^0-E_i^0}\langle b\,|\,V(1+X)\,|\,a\rangle\ , \tag{B9}$$

$$(i,a\,|\,A\,|\,j,a)=-\frac{1}{E_a^0-E_i^0}\langle i\,|\,(1-X)V\,|\,j\rangle\ . \tag{B10}$$

Up to the first order in V, expressions (B8), (B9), and (B10) become, respectively,

$$(i,a\,|\,A\,|\,i,a)=1+\frac{1}{E_a^0-E_i^0}[\langle a\,|\,V\,|\,a\rangle-\langle i\,|\,V\,|\,i\rangle]\ , \tag{B11}$$

$$(i,a \mid A \mid i,b)=\frac{1}{E_a^0-E_i^0}\langle b \mid V \mid a\rangle \ , \tag{B12}$$

$$(i,a \mid A \mid j,a)=-\frac{1}{E_a^0-E_i^0}\langle i \mid V \mid j\rangle \ . \tag{B13}$$

Expressions (B11)—(B13) clearly indicate that for small perturbations A and $C=A^{-1}$ tend towards the identity operator. Thus we have established the mathematical correspondence between our method based on the convergence operator C and the Rayleigh-Schrödinger methodology which is valid for small perturbations and corresponds within our formalism to $C=1$.

APPENDIX C: LINEAR AND QUADRATIC ITERATIVE METHODS

This appendix briefly recalls the main results concerning the standard linear and quadratic methods for solving ordinary algebraic equations. These results are then extended to operator equations such as those met in Sec. III.

1. Ordinary algebraic equations[21]

Let

$$F(x)=0 \tag{C1}$$

be an ordinary algebraic equation. One of the most effective ways for determining the roots of this equation is the iterative Newton-Raphson procedure. At the nth iteration the variation $\Delta x^{(n)}=x^{(n)}-x^{(n-1)}$ of x is given by (for the quadratic method)

$$\Delta x^{(n)}=-\frac{F(x^{(n-1)})}{F'(x^{(n-1)})} \ . \tag{C2}$$

$F'(x)$ is the derivative of $F(x)$ with respect to x. The geometrical interpretation of this procedure which converges quadratically is given in Fig. 6(a). Quasiquadratic methods can be considered as those for which $1/F'(x)$ is calculated in an approximate way. For characterizing the linear methods it is useful to introduce the notation

$$f(x)=x-F(x) \ . \tag{C3}$$

Equations (C1) and (C2) become

$$F(x)=x-f(x)=0 \ , \tag{C4}$$

$$\Delta x^{(n)}=\frac{F(x^{(n-1)})}{1-f'(x^{(n-1)})} \ . \tag{C5}$$

If $f(x)$ is a slowly varying function of x [$|f'(x)| \ll 1$], Eq. (C5) can be approximated by (for the linear method)

$$\Delta x^{(n)}=-F(x^{(n-1)}) \ , \tag{C6}$$

which by using (C3) can also be written (also for the linear method)

$$x^{(n)}=f(x^{(n-1)}) \ . \tag{C7}$$

(C6) or (C7) define iterative procedures which converge linearly, the geometrical interpretations of which are given in Figs. 6(b) and 6(c). Note the oscillatory convergence towards the solution in 6(c) when $f'(x)<0$.

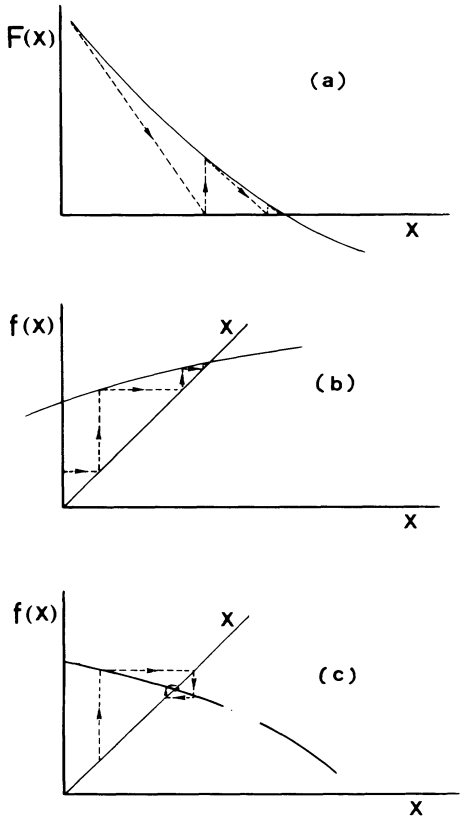

FIG. 6. (a) Newton-Raphson iterative method (quadratic convergence) for solving the algebraic equation $F(x)=0$. (b) and (c) Iterative method (linear convergence) for $F(x)=x-f(x)=0$.

2. Operator equation

Let us now consider the operator equation

$$F(X)=0 \ . \tag{C8}$$

[See, for example, Eqs. (32) and (46) of Sec. III.] The analog of (C2) for the variation $\Delta X^{(n)}$ of the operator X at the nth iteration is

$$\Delta X^{(n)}=-A^{-1}F(X^{(n-1)}) \ . \tag{C9}$$

(C9) defines a generalized Newton-Raphson procedure. The comparison between (C2) and (C9) shows that the operator or superoperator A for determining X plays the same role as the derivative $F'(x)$ for finding the roots of an ordinary algebraic equation.

When the operator equation is put in the form

$$X-f(X)=0 \ , \tag{C10}$$

it is useful to introduce the convergence operator

$$C=A^{-1} \tag{C11}$$

and (C9) becomes (for the quadratic method)

$$\Delta X^{(n)}=-CF(X^{(n-1)}) \ . \tag{C12}$$

The matrix elements of A given by (B8)—(B10) show that for small perturbation A and C converge to the identity

operator and (C12) becomes (for the linear method)

$$\Delta X^{(n)} = -F(X^{(n-1)}) \ . \tag{C13}$$

(C13) defines an iterative procedure which converges linearly and is the analog of (C6).

The results given in this appendix clearly indicate that all the operator iterative procedures presented in Sec. III must be replaced in the general framework of the standard linear and Newton-Raphson iterative methods for solving ordinary algebraic equations. We emphasize again the importance of the operator C on which are based all the convergence properties of these iterative methods.

[1]B. H. Brandow, Rev. Mod. Phys. 39, 771 (1967).

[2]*Effective Interactions and Operators in Nuclei,* Vol. 40 of *Lecture Notes in Physics,* edited by B. R. Barrett (Springer, Berlin, 1975).

[3]V. Kvasnička, in *Advances in Chemical Physics,* edited by I. Prigogine and S. A. Rice (Wiley, New York, 1977), Vol. XXXVI, p. 345.

[4]B. H. Brandow, in *Advances in Quantum Chemistry,* edited by P. O. Löwdin (Academic, New York, 1977), Vol. 10, p. 187.

[5]K. F. Freed, J. Chem. Phys. 60, 1765 (1974); in *Semiempirical Methods of Electronic Structure Calculation, Part A: Techniques,* edited by G. A. Segal (Plenum, New York, 1977).

[6]C. Bloch, Nucl. Phys. 6, 329 (1958).

[7]J. des Cloizeaux, Nucl. Phys. 20, 321 (1960).

[8]I. Lindgren, J. Phys. B 7, 2441 (1974). See also I. Lindgren and J. Morrison, *Atomic Many-Body Theory* (Springer, Berlin, 1982), Chap. 9.

[9]G. Nicolas and Ph. Durand, J. Chem. Phys. 72, 453 (1980); J. L. Brédas, R. R. Chance, R. Silbey, G. Nicolas, and Ph. Durand, *ibid.* 75, 255 (1981).

[10]M. G. Sheppard and K. F. Freed, J. Chem. Phys. 75, 4507 (1981); 75, 4525 (1977); H. Sun, M. G. Sheppard, K. F. Freed, and M. F. Herman, Chem. Phys. Lett. 77, 555 (1981); Y. S. Lee and K. F. Freed, J. Chem. Phys. 77, 1984 (1982).

[11]S. Ôkubo, Prog. Theor. Phys., 12, 603 (1954).

[12]F. Jørgensen, Mol. Phys. 29, 1137 (1975); see also C. E. Soliverez, Phys. Rev. A 24, 4 (1981).

[13]C. Møller, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. XXIII, No. 1 (1945); 19, No. 3 (1946). See also R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), Chap. 6, Sec. 3 and J. R. Taylor, *Scattering Theory* (Wiley, New York, 1972), Chap. 3, Sec. a.

[14]R. J. Eden and N. C. Francis, Phys. Rev. 97, 1366 (1955).

[15]B. H. Brandow, Int. J. Quantum Chem. XV, 207 (1979).

[16]P. O. Löwdin, J. Math. Phys. 3, 969 (1962).

[17]Ph. Durand, J. Phys. Lett. 43, L-461 (1982).

[18]I. Lindgren, Int. J. Quant. Chem. Symp. 12, 33 (1978).

[19]H. Primas, Rev. Mod. Phys. 35, 710 (1963).

[20]E. R. Davidson, J. Comput. Phys. 3, 969 (1975).

[21]E. Durand, *Solutions Numériques des Équations Algébriques* (Masson, Paris, 1960), Tome I.