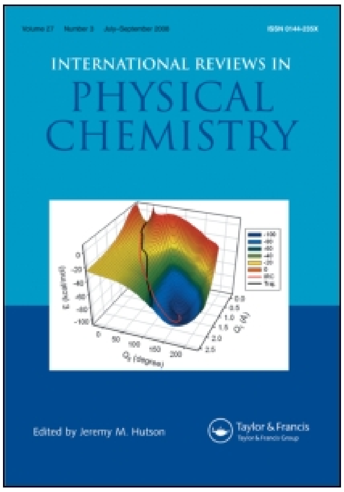

International Reviews in Physical Chemistry

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713724383

Iterative techniques for molecular CI wavefunctions

G. L. Bendazzoli[a]; S. Evangelisti[a]; P. Palmieri[a]; S. Rettrup[b]
[a] Dipartimento di Chimica Fisica ed Inorganica, Viale Risorgimento, Bologna, Italy [b] Department of Physical Chemistry, H.C. Ørsted Institute, University of Copenhagen, Copenhagen Ø, Denmark

To cite this Article Bendazzoli, G. L. , Evangelisti, S. , Palmieri, P. and Rettrup, S.(1987) 'Iterative techniques for molecular CI wavefunctions', International Reviews in Physical Chemistry, 6: 4, 385 — 392

To link to this Article: DOI: 10.1080/01442358709353201

URL: http://dx.doi.org/10.1080/01442358709353201

INTERNATIONAL REVIEWS IN PHYSICAL CHEMISTRY, 1987, VOL. 6, NO. 4, 385–392

Iterative techniques for molecular CI wavefunctions

by G. L. BENDAZZOLI, S. EVANGELISTI and P. PALMIERI
Dipartimento di Chimica Fisica ed Inorganica,
Viale Risorgimento 4, I-40136 Bologna, Italy

and S. RETTRUP
Department of Physical Chemistry, H.C. Ørsted Institute,
University of Copenhagen, Universitetsparken 5,
DK-2100 Copenhagen Ø, Denmark

Some iterative methods for eigenvalue problems and systems of linear equations involving large matrices are reviewed. Applications to the computation of molecular eigenstates and properties are presented and discussed.

1. Introduction

Iterative methods are routinely used for solving the eigenvalue problem

$$Hx = Ex \tag{1}$$

where the Hamiltonian H is a large symmetric matrix of dimension N. Here N is of the order of several thousand and therefore the matrix cannot be kept in fast storage, while single rows or columns are supposed to be.

A similar problem occurs in the direct calculation of transition energies in the random phase approximation (RPA), that leads to a non-Hermitian eigenvalue problem which can be written in the form

$$Rx = \omega^2 x \tag{2}$$

R is a non-Hermitian matrix, $R = (A - B)(A + B)$, where A and B are Hermitian submatrices, and ω represents the transition energy.

Iterative methods are also used to solve large systems of linear equations of the form

$$Ax = b \tag{3}$$

Such systems of equations are involved in many problems of interest in quantum chemistry, such as in perturbation theory and Newton–Raphson methods for solving nonlinear problems such as MCSCF and Coupled Cluster approaches. In this paper we review some known iterative methods for solving equations (1)–(3) and illustrate applications to molecular problems.

2. Eigenvalue problems

A common feature of the methods we are considering is that we approximate the solution of the original problem (in the large space L of dimension N) with a sequence of solutions of 'reduced' problems in smaller subspaces L_n of dimension $n+1$. This value n increases at each iteration and its maximum value must be such that the reduced problem can be solved in L_n by standard methods for small matrices. The effectiveness of the method depends upon the way this subspace is generated step after step. Usually an approximation of the resolvent of H is involved in some explicit or implicit way. One of the first contributions is that of Bartlett and Brändas (1972, 1973). They developed a variational-perturbation theory for solving the eigenvalue problem (1) based on a general splitting of the Hamiltonian $H=H_0+V$ and applied it in the case where H_0 is the outer projection of H in a subspace of small dimension. The best known method for symmetric matrices is due to Davidson (1975, 1983); his algorithm originated a family of methods (see, e.g. Hinze 1982). An important feature of Davidson's technique is that it can be programmed in a direct way, i.e. without explicit construction of the Hamiltonian matrix H.

We begin by considering the simplest problem, i.e. finding the lowest eigenstate of a symmetric matrix H. The approximate eigenvector x_n at the nth iteration is a linear combination of vectors h_k which constitute a basis of L_n:

$$x_n=\sum_{i=0}^{n} h_i c_{in} \tag{4}$$

The coefficients (c_{kn}) are obtained by solving the eigenvalue problem for the matrix M_n (dimension $n+1$)

$$M_n c_n = S_n c_n E_n \tag{5}$$

The elements of the Hamiltonian and overlap matrices M_n and S_n are defined as follows:

$$(M_n)_{ij}=h_i^\dagger H h_j \tag{6a}$$

$$(S_n)_{ij}=h_i^\dagger h_j \tag{6b}$$

The sequence starts with a guess h_0 and a new h_{n+1} is added at each iteration. We can associate to each x_n a residual r_n defined as

$$r_n=(H-E_n)x_n \tag{7}$$

This is used to test convergence and to generate h_{n+1}. The latter is given according to

$$h_{n+1}=Z_n^{-1}r_n \tag{8}$$

Different choices of the linear operator or matrix Z_n generate different methods. From a purely mathematical point of view one should use the reduced resolvent, $Z_n^{-1}=(E_n-P_nH)^{-1}P_n$ (Löwdin 1968) where P_n is the projector in the orthogonal complement of x_n. Usually, however, such a Z_n is difficult to invert and approximations which are more easily invertable are required. Davidson's method is obtained when Z_n is chosen to be the diagonal of $(H-E_n)$. Gradient methods are included in this scheme for $Z_n=I$; they have been used by Bartlett and Brändas and are closely related to Lanczos' method (see, e.g. Golub and van Loan 1983). Other choices to be considered are block-diagonal approximations of H. Another possibility has been pointed out (Bendazzoli *et al.* 1987 a): one can use the lower or upper triangle of $(H-E_n)$ since h_{n+1} can then be obtained by solving a triangular system of linear equations at each iteration. Such systems are easily programmed and the only requirement is to access to the elements of the matrix row-wise. This seems to destroy the possibility of a direct-CI type of implementation of the algorithm. Numerical experiments indicate, however, that the triangle is a good choice (Bendazzoli *et al.* 1987 a). Also band matrices can easily be inverted as is well known in the numerical solution of partial differential equations;

unfortunately CI matrices do not show any band structure, even in some approximate way. The different choices of Z_n can be intermixed and the possibility to switch from one another is trivially implemented.

Once the new vector h_{n+1} is generated, we have two alternatives. We can orthogonalize it to the previous h_n by the Schmidt process and get an orthonormal basis for L_n. In such a case S_n of equation (5) is the identity. The second possibility is to keep the basis non-orthogonal and solve equation (5) by a routine for generalized eigenvalue problems. Our numerical experiences seem to indicate a better stability of the first procedure, which involves some more computational work.

We have implemented these procedures in a conventional configuration-driven CI program based on the spin bonded formalism (Cooper and McWeeny 1966). For determination of the ground state, convergence is achieved typically in less than ten iterations and Davidson's choice of Z_n is the most convenient because it requires less computer time than the triangle and is more effective than the gradient. This program has been applied to compute wavefunctions at the SDCI level to test many-body methods for simple molecules (Bendazzoli *et al.* 1982 a).

The non-Hermitian problem in equation (2) can be treated similarly (Rettrup 1982). Considering specifically the application to the RPA method, it is known that equation (2) is equivalent to a generalized Hermitian eigenvalue problem with eigenvalues ω. It means that the solutions to equation (2) can be chosen to be real. In practical applications, however, we find that spurious complex contributions arise during the iterations in the solution of the small non-Hermitian eigenvalue problem (5). Computationally this difficulty is handled by using, instead of M_n, the modified matrix (Bouman *et al.* 1983)

$$M'_n = (M_n + aM_n^\dagger)/(1+a) \tag{9}$$

Gradually in each iteration we increase a (e.g.: 0·0, 0·1, 0·2,...) in solving (5) until the complex components disappear. At present we are also using the algorithm in a direct CI program for correlation studies based on the biorthogonal valence bond approach.

Davidson's algorithm is easily modified for excited states. We start the iterative process with a set of guess vectors spanning the 'dominant components' of the states lying below the particular eigenvector we are looking for. This idea has been refined in several ways, and an alternative has been explored (Bendazzoli *et al.* 1987). The idea, which goes back to Feler (1974), is to minimize the quadratic form

$$(Hx - wx)^\dagger (Hx - wx) \tag{10}$$

with the normalization constraint on x, and where w is a given number. This is equivalent to computing the ground eigenvector of $(H-w)^\dagger(H-w)$. The latter is nothing but the eigenvector of H closest to w in energy; only one guess vector is needed. All the process is programmed in a way very similar to the normal Davidson algorithm. At each step we solve equation (5) as before, where the matrix M_n is now given by

$$(M_n)_{ij} = ((H-w)h_i)^\dagger (H-w)h_j \tag{11}$$

while the overlap S_n is still given by equation (6 *b*). Therefore we do not need to compute the square of the matrix H. The new vector h_{n+1} is generated using the same choices of Z_n operators as in the ground state case. We have again two variants, one with Schmidt orthogonalization and one without. The method proved to be effective for CI matrices of the order of 2–3000, where it converged in less than 40 iterations for states 0·9 a.u. above the ground state. It should be pointed out that the computation of eigenvectors in a region of very high density of states like this one is a difficult task. Neither the diagonal nor simple gradient method was able to achieve convergence in this case; the only successful choice of Z_n was the triangle. As far as the numerical problems are concerned they turn up in the form of approximate linear dependencies. These lead to nearly singular overlap matrices S_n, the eigenvalue problem (5) becomes unstable and requires special treatment. In the other method we have the following situation. When we Schmidt-orthogonalize the vector h_{n+1} given by equation (8) to the previous vectors, its norm $\|h_{n+1}\|$ will become almost zero. Thereafter the normalization of this vector will destroy orthogonality because of round-off errors. We check the norm of the vector after orthogonalization and if it is smaller than a given threshold, we perform the Schmidt orthogonalization again.

This technique could be useful for studying X-ray transitions, where a hole state with very high energy can interact with a manifold of single and double excitations. The method can also be adapted to the computation of excited potential energy curves of molecules. It is convenient to compute the ground-state energy E_0 at the required geometries and to express w as $E_0+\delta$, where δ is a constant for all geometries. An example of this procedure is given in table 1 for an excited state of $^2\Sigma^+$ symmetry of OH. The dimension of the matrix is 2794 and the value of δ is 0·379 68.

Table 1. Ground and excited state energies E_0, E_i of $^2\Sigma^+$ symmetry for OH for a number of internuclear distances R. The dimension of the CI matrix is 2794 and the convergence threshold for residual norm is 10. All quantities are in a.u.

R	E_0	E_i	E_i-E_0	Number of iterations
1·50	−75·323 737	−74·944 06	0·374 248	11
2·00	−75·374 260	−74·994 58	0·356 212	13
2·50	−75·344 057	−74·964 38	0·337 167	15
4·00	−75·301 895	−74·922 22	0·399 831	26

3. Systems of linear equations

We write the system in the form

$$(H-w)x=b \tag{12}$$

where w is a given constant. This problem has remarkable similarities with the previous one. We consider methods where the solution x of equation (12) is approximated by an expansion similar to equation (4); we have again a subspace L_n growing with the number of iterations, and a residual vector r_n given by

$$r_n=(H-w)x_n-b \tag{13}$$

This residual is used to generate the new basis vector h_{n+1} by means of equation (8), where the operator Z_n is defined as before, but with E_n replaced by the constant w. Therefore Z does not depend on the iteration number n; in this case Z^{-1} should be an approximation not to the reduced, but to the full resolvent of H, without projectors in the orthogonal complement to x_n. The coefficients c_{in} are now obtained as solutions of a reduced system of linear equations

$$M_n c_n=b_n \tag{14}$$

There are essentially two ways to define the reduced equation (14). The first, introduced in quantum chemistry by Pople *et al.* (1980), can be described as a projection method (Purvis and Bartlett 1981), since the matrix M_n is the outer projection of H–w in the reduced subspace L_n. In this case

$$(M_n)_{ij} = h_i^\dagger (H-w) h_j \tag{15}$$

$$(b_n)_i = h_i^\dagger b \qquad i,j = 0, 1, \ldots, n \tag{16}$$

The problem can be viewed as the unconstrained minimization of the quadratic form

$$x_n^\dagger (H\quad w) x_n + b_n^\dagger x_n + x_n^\dagger b_n$$

in the subspace L. The second method has been considered by Purvis and Bartlett (1981) and given the name ‘least squares’. It can be described as the minimization of another quadratic form, i.e.

$$((H-w)x_n)^\dagger (H-w) x_n + ((H-w)b_n)^\dagger x_n + x_n^\dagger (H-w) b_n$$

corresponding to the square norm of the residual vector. The new definitions of matrix M_n and vector b_n in reduced space L_n are now

$$(M_n)_{ij} = ((H-w)h_i)^\dagger (H-w) h_j \tag{17}$$

$$(b_n)_i = ((H-w)h_i)^\dagger b_j \tag{18}$$

Equation (14) can be solved by standard methods for small matrices. The components c_{in} of the vector c_n will in general change from iteration to iteration. One can generate the basis vectors h_n in such a way that the matrix M_n is diagonal. This is obtained by a Schmidt-like orthogonalization in a suitable pseudo metric A ($A = (H-w)$ in the projection case, and $A = (H-w)^2$ in the least-square case). As soon as a h_n is generated, we replace it by

$$h_n - \sum_{i=0}^{n-1} [h_i^\dagger (H-w) h_n] h_i \tag{19}$$

Equation (14) is now trivially solved and the components c_{in}, $i = 1, 2, \ldots n$, computed at iteration n will not change in iteration $n+1$. This is strongly reminiscent of the conjugate gradient algorithm (5) and, as Wormer *et al.* (1982) have shown, both projection and least-squares procedures are nothing but particular cases of the preconditioned conjugate gradient algorithm, in which h_{n+1} is generated by choosing $Z = I$ and it is sufficient to perform the A-orthogonalization to h_{n-1} with considerable saving of computer time. In the preconditioned version of the method, one introduces a matrix C called the preconditioner and considers the modified system

$$C^{-1/2}(H-w)C^{-1/2} y = C^{-1/2} b \tag{20}$$

and tries to solve it by conjugate gradient. The preconditioner C is chosen to improve the convergence; in our case, the role of preconditioner is played by the matrix Z. In the present version of our program we use equation (19).

Perturbation theory provides good examples of systems of linear equations when the wavefunction is expanded in a CI basis. For a time-dependent Hamiltonian $H = H_0 + V\exp(\mathrm{i}\omega t) + V\exp(-\mathrm{i}\omega t)$ with unperturbed eigenstates Ψ_n such that $H_0\Psi_n = E_n\Psi_n$, the first order equations are (in atomic units)

$$(H_0 - E_n + \omega)\phi_n^+ = -V\Psi_n \tag{21\,a}$$

$$(H_0 - E_n - \omega)\phi_n^- = -V\Psi_n \tag{21\,b}$$

The quantity of physical interest is usually given by the sum

$$\phi_n^{+\dagger}\Psi_n+\phi_n^{-\dagger}\Psi_n \tag{22}$$

The time-independent case is obtained as the limit for $\omega=0$ of equation (22). It should be noticed that the individual equations (21 *a*, *b*) are singular for $\omega=0$; in equation (22) there is cancellation of a divergent contribution from the two terms. This contribution comes from the component of $V\Psi_n$ along the eigenvector Ψ_n of (H_0-E_n). Thus the first order time-independent equation differs from (21 *a*, *b*) in the right-hand side

$$(H_0-E_n)\Psi_n=(E_1-V)\Psi_n \tag{23}$$

which is now orthogonalized to Ψ_n, a necessary condition for the existence of solutions.

Expansion of Ψ_n and ϕ_n in a CI basis leads directly to systems of linear equations. A typical example is the computation of dynamic polarizability, where V is the usual dipole interaction with the field. An example concerns the OH radical (Bendazzoli *et al.* 1987 b). Equations (21 *a*, *b*) are solved for a range of values of ω and the solutions combined according to equation (22). Both projection and least-squares methods converge in seven to ten iterations when the system is well conditioned. In the case of equation (23), the iterative process can be carried out in the orthogonal complement to Ψ_n, which is known in advance. Accidental singularities (resonances) may be met for non-zero values of ω, when $E_0+\omega$ or $E_0-\omega$ is close to another eigenvalue E_r of H_0. This case could be treated by computing separately the corresponding eigenvector Ψ_r, and working in the orthogonal complement. This procedure is inconvenient, because one should previously test the neighbourhood of ω for eigenvectors, i.e. using the method for excited eigenvalues described in section 2. For this reason a method was developed which simultaneously computes the resonant, if any, eigenvector Ψ_r and the solution in the orthogonal complement to Ψ_r (Bendazzoli *et al.* 1987 b). Equation (14) is solved by diagonalizing the matrix M_n with a unitary transformation, in order to find its eigenvalues. These are related by the variational principle to those of $(H_0-E_n+\omega)$ if we are in the projection method, or to those of $(H_0-E_n+\omega)^2$ in least-squares. A near-resonant eigenstate corresponds to a small eigenvalue of M_n. In the projection method the matrix M_n is not positive definite and this brings about a rather disturbing feature, i.e. the appearance of temporary small eigenvalues. Because of the variational principle, the eigenvalues of M_n decrease as the number n of iterations increases. Therefore an initially positive eigenvalue can be positive at the beginning of the cycle, very close to zero at some intermediate stage, and finally assume a negative value at the end. In such a situation it is difficult to decide whether we have a genuine resonance or not. This is not the case for the least-squares method, for the M_n matrix is always positive definite. We have implemented a method called 'alternate iterations', in which equation (14) is solved by diagonalizing the M_n matrix defined as in equation (17) and a check is performed on the smallest eigenvalue of M_n. When this is below a given threshold, we split the problem in two parts, called singular and regular. The former is the resonant eigenvector Ψ_r, the latter the solution of the system in the orthogonal complement to Ψ_r. The iterative process continues with two residuals, a singular (r_{sn}) defined as in equation (7) and a regular (r_{rn}), defined as in equation (13). The vector h_{n+1} is generated using equation (8) with a Z operator acting alternatively on r_{sn} and r_{rn} (Bendazzoli *et al.* 1987 b).

Another field of application of the first-order inhomogeneous equations of perturbation theory is the computation of transition probabilities in spectroscopic processes. A two-photon transition amplitude from a molecular eigenstate Ψ_n to

Table 2. xz components of the two-photon tensor $\langle A^2\Sigma^+|x(H-\hbar\omega)^{-1}z|X^2\Pi\rangle$ for OH for different values of the internuclear separation R and photon energy $\hbar\omega$. All quantities are in a.u. The convergence threshold is 10^{-4}. The numbers in parentheses give the number of iterations.

$\hbar\omega$	$R=1{\cdot}50$	$R=2{\cdot}00$	$R=2{\cdot}50$	$R=4{\cdot}00$
−0·16	−1·007 46 (5)	−0·718 93 (5)	−0·382 52 (7)	0·068 60 (7)
−0·12	−1·309 78 (5)	−0·903 43 (6)	−0·449 41 (7)	0·083 47 (8)
−0·08	−1·903 32 (5)	−1·253 52 (6)	−0·559 51 (8)	0·106 78 (9)
−0·04	−3·655 12 (5)	−2·252 87 (6)	−0·830 51 (8)	0·157 06 (10)
0·04	3·211 94 (6)	1·519 32 (8)	−0·044 88 (9)	0·082 67 (12)
0·08	1·466 08 (7)	0·494 69 (8)	−0·363 72 (9)	0·254 57 (13)
0·12	0·850 74 (7)	0·093 30 (8)	−0·568 02 (9)	0·060 52 (12)
0·16	0·512 42 (7)	−0·174 62 (8)	−0·801 19 (10)	0·215 33 (12)

another state Ψ_m can be computed as the scalar product of the solution of equation (21) with $V\Psi_m$. Both quantities are in principle representable with CI expansions, if suitable approximations are introduced for the nuclear degrees of freedom (Bendazzoli *et al.* 1983). We attempted the computation of absolute values of absorptivities for two-photon transitions in OH and H_2 using symmetry-adapted CSF expansions for a number of internuclear distances. In table 2 we show some values for the xz component of the two-photon tensor for the $A^2\Sigma^+ \leftarrow X^2\Pi$ transition of OH. Only order-of-magnitude agreement with experimental data is obtained with our CI expansion at the multi-reference SDCI level of dimension 2185. Another example is the study of the $A_{1g}-B_{2u}$ transition in benzene (Bendazzoli *et al.* 1986). Two-photon absorption is forbidden in the dipole approximation and a vibronic mechanism is advocated. As an example we may consider the xx component of the transition tensor, where x is the axis bisecting two opposite C–C bonds. The CI computation has been performed at a geometry distorted according to a well defined prescription to account for vibronic interactions. Using a minimal basis set of STO-3G AOs, we generated a list of 4392 CSF selected from a larger set of configurations. The selection criteria take into account both energy and dipole matrix elements. The solution of the linear system (21) typically requires eight to ten iterations to get a residual norm less than 10^{-4} in the least-squares method. When comparison with experiment is possible, intensities and polarization ratios of various vibronic components are qualitatively reproduced.

As a last example we quote the computation of magnetic circular dichroism (MCD) constants for 9,9′-spirobifluorene (SBF) (Bendazzoli *et al.* 1982 b). The MCD spectrum originates when a transition to a degenerate excited state takes place in an external magnetic field. Our procedure applies in particular to the computation of the B constant. The latter contains matrix elements of operators between molecular states and the solution of equation (23), where V is the perturbation due to the magnetic field and Ψ_n is one of the degenerate components of the excited state. This technique more accurately described the dichroism of SBF, confirming the degeneracy of the lowest excited state of this molecule.

References

Bartlett, R. J., and Brändas, E. J., 1972, *J. chem. Phys.*, **56**, 5467.

Bartlett, R. J., and Brändas, E. J., 1973, *J. chem. Phys.*, **59**, 2032.

Bendazzoli, G. L., Fano, G., Ortolani, F., and Palmieri, P., 1982 a, *J. chem. Phys.*, **76**, 2498.

BENDAZZOLI, G. L., DEGLI ESPOSTI, A., PALMIERI, P., MARCONI, G., and SAMORI, B., 1982 b, *J. chem. Soc. Faraday Trans.*, **2** (78), 1623.

BENDAZZOLI, G. L., DEGLI ESPOSTI, A., EVANGELISTI, S., and PALMIERI, P., 1983, *Chem. Phys. Lett.*, **99**, 266.

BENDAZZOLI, G. L., EVANGELISTI, S., PALMIERI, P., and RETTRUP, S., 1986, *J. chem. Phys.*, **85**, 2105.

BENDAZZOLI, G. L., EVANGELISTI, S., and PALMIERI, P., 1987 a, *Int. J. quant. Chem.*, **31**, 663.

BENDAZZOLI, G. L., EVANGELISTI, S., and PALMIERI, P., 1987 b, *Int. J. quant. Chem.*, **31**, 673.

BOUMAN, T. D., HANSEN, A. E., VOIGT, B., and RETTRUP, S., 1983, *Int. J. quant. Chem.*, **23**, 595.

COOPER, I. L., and MCWEENY, R., 1966, *J. chem. Phys.*, **44**, 226; *ibid.*, **45**, 3484.

DAVIDSON, E. R., 1975, *J. comp. Phys.*, **17**, 87.

DAVIDSON, E. R., 1983, *Methods in Computational Molecular Physics*, NATO ASI Series C, 113, edited by G. H. F. Diercksen and S. Wilson (Dordrecht: Reidel), and references therein.

FELER, M. G., 1974, *J. comp. Phys.*, **14**, 341.

GOLUB, G. H., and VAN LOAN, C. F., 1983, *Matrix Computations* (Oxford: North Oxford Academic).

HINZE, J., 1982, *Lecture Notes in Mathematics* no. 968 (Berlin, Heidelberg, New York: Springer-Verlag), and references therein.

LÖWDIN, P.-O., 1968, *Int. J. quant. Chem.*, **2**, 867.

POPLE, J. A., KRISHNAN, R., SCHLEGEL, H. B., and BINKLEY, J. S., 1980, *Int. J. quant. Chem. Symp.*, **13**, 312.

PURVIS, G. D., and BARTLETT, R. J., 1981, *J. chem. Phys.*, **75**, 1284.

RETTRUP, S., 1982, *J. comp. Phys.*, **45**, 100.

WORMER, P. E. S., VISSER, F., and PALDUS, J., 1982, *J. comp. Phys.*, **48**, 23.