Rayleigh–Schrödinger perturbation theory: Practical implementation of matrix and vector formalisms and description of an heuristic sufficiency convergence criterion

Emili Besalú and Ramon Carbó-Dorca

Institute of Computational Chemistry, University of Girona, 17071 Girona, Catalunya, Spain

Received 25 July 1996; revised 5 August 1997

Starting from previous work, where Rayleigh–Schrödinger perturbation theory has been reformulated in matrix form, a practical algorithm implementation is described using both full matrix and vector alternatives. An heuristic convergence sufficiency criterion based on Gershgorin discs is also presented. Some numerical examples related to atomic CI computations are reported to illustrate the theoretical framework.

1. Introduction

The relevance of perturbation theory (PT) into the treatment in some procedures related to CI computational schemes [31,33,34]. A great deal of purely formal PT related work can be found in [1,2,10,11,13–21,23–27,29,35–37,39–43,47] is well known. Thus, it seems interesting to develop such practical facets of the theory which may be useful to wavefunction calculation. Recently, the present authors have published some papers related to the Rayleigh-Schrödinger PT (RS-PT) [4,6]. In [4], the matrix notation of the RS framework was described. There, the PT treatment is general enough to study several perturbation matrices or operators at the same time. Here we focus our work on the *practical implementation* of the formulation given previously and the discussion is circumscribed to the most habitual case where only one perturbative term is present. First, the PT problem will be presented from the matrix point of view. It will be followed by a brief description of the algorithm which implements the matrix notation of [4]. From the previous algorithm it is straightforward to describe the related vector notation dealing with the perturbative resolution of a particular state or eigenvalue-eigenvector pair. A heuristic convergence criterion will be described afterwards and some calculation examples will be finally given.

© J.C. Baltzer AG, Science Publishers

2. Perturbation theory: a matrix point of view

As it is usual, the PT scheme is formulated in order to solve an eigenvalue problem which can be finally written as the matrix equation

$$HU = UE, \tag{1}$$

where the *perturbed matrix* H is expressed in terms of an *unperturbed matrix* $H^{(0)}$ plus a *perturbation* V:

$$H = H^{(0)} + V. (2)$$

The solution of equation (1) is found, from the PT point of view, starting from the *full knowledge* of the solution of the unperturbed system secular equation

$$H^{(0)}U^{(0)} = U^{(0)}E^{(0)},$$
(3)

where the matrices $U^{(0)}$ and $E^{(0)}$ contain in their columns and diagonal elements the so-called unperturbed eigenvectors and eigenvalues, respectively.

The unperturbed eigenvectors can be used as a working basis set, expanding all the space of the same dimension as the order of the involved matrices. As the unperturbed vectors can be considered orthonormalized, it is easy to realize how they can act as a canonical basis set if the proper definitions are taken into account. Expressing $H^{(0)}$ by means of the matrix product

$$H^{(0)} = U^{(0)} E^{(0)} U^{(0)+},$$
(4)

and defining the following similarity transformation of the perturbation matrix V:

$$P = U^{(0)+} V U^{(0)}, (5)$$

it is straightforward to express the perturbed matrix to be diagonalized as

$$H = U^{(0)} R U^{(0)+}, (6)$$

the matrix \boldsymbol{R} being defined in turn as

$$R = E^{(0)} + P. (7)$$

That is, the eigenvalue problem (1) has been transformed into the one associated to diagonalization of the matrix R. Once this diagonalization is performed, by solving the secular equation

$$RZ = ZE, (8)$$

the obtained eigenvalues are the same as those in equation (1) and the eigenvectors contained in the columns of the matrix Z express the change of basis, which has to be applied over the unperturbed basis set $U^{(0)}$, in order to obtain the eigenvectors of the original secular equation (1), that is,

$$\boldsymbol{U} = \boldsymbol{U}^{(0)}\boldsymbol{Z}.\tag{9}$$

86

Because of these previous considerations, from now on, the *perturbational solution* of the system (8) will be the main concern of the present study. From the PT point of view, in equation (8), the matrix $E^{(0)}$ can be interpreted as the unperturbed system matrix and P acts as a perturbative term. Within this problem framework, as $E^{(0)}$ matrix is diagonal, the same matrix $E^{(0)}$ and the unit matrix I define the transformed problem unperturbed eigenvalues and eigenvectors, respectively. Moreover, as it is well known, within the RS-PT, the diagonal elements of the matrix P are the first order eigenvalue corrections

$$p_{ii} = e_{ii}^{(1)} \quad \forall i, \tag{10}$$

being defined as

$$e_{ii}^{(1)} = \left[\boldsymbol{U}^{(0)} \right]_{i}^{+} \boldsymbol{V} \left[\boldsymbol{U}^{(0)} \right]_{i} \quad \forall i.$$
 (11)

Here, $[U^{(0)}]_i$ means the column attached to the *i*th state of matrix $U^{(0)}$ and $[U^{(0)}]_i^+$ stands for the conjugate of $[U^{(0)}]_i$ in a row form. When speaking of the *i*th state we make simultaneous reference to both the *i*th eigenvalue and the related eigenvector, following the notation usually encountered in the quantum mechanical framework.

3. Optional unperturbed eigenvalues redefinition

The particular point of view of the preceding section, concerning the diagonalization of matrix \mathbf{R} , reflects only a possible particular PT problem choice. In general, there are infinite continuous ways to split the matrix \mathbf{R} as a diagonal matrix plus a perturbation. For instance, the zeroth and first eigenvalue corrections can be mixed redefining an equivalent matrix \mathbf{R} . For every diagonal element r_{ii} of \mathbf{R} , one can choose a parameter $\gamma_i \in (-\infty, +\infty)$ which produces the particular splitting for this element:

$$r_{ii} = e_{ii}^{(0)} + e_{ii}^{(1)} = r_{ii}(1 - \gamma_i) + r_{ii}\gamma_i \quad \forall i,$$
(12)

leading easily towards a parametrized redefinition of the values contained in the diagonals of $E^{(0)}$ and P matrices:

$$e_{ii}^{(0)}(\gamma_i) = r_{ii}(1 - \gamma_i) \quad \land \quad p_{ii}(\gamma_i) = r_{ii}\gamma_i \quad \forall i, \tag{13}$$

and leaving the secular equation (8) invariant. In this way, another diagonal unperturbed system and a related perturbation matrix are defined. Despite that the whole matrix \boldsymbol{R} remains unaltered but, from the numerical point of view, the convergence of the PT is modified. This possibility can be used as an attempt to speed up the convergence of the numerical method or to remove the degeneracy or quasi-degeneracy problems, if present, in the system under study. The classical PT scheme is equivalent to using a set of $\{\gamma_i\}$ parameters which leaves invariant the original definitions for the $\boldsymbol{E}^{(0)}$ and \boldsymbol{P} matrices, provided in equations (3) and (5), respectively.

4. Matrix and vector algorithms

Here, two algorithms will be described. They are related to the full matrix and vector notations of the RS-PT. As PT, formulated by means of the considerations of section 2 above, can be considered a diagonalization procedure, the algorithms defined below correspond to a full and a partial search of the system states. With respect to the vector notation, it can be related to the method to find a unique set of eigenparameters using Jacobi rotations [9].

4.1. Matrix algorithm A1

From the general matrix PT scheme, developed in [4], and the practical considerations outlined above, an algorithm can be described implementing the RS-PT to solve the secular equation (1). Such a matrix algorithm, if convergent, may find the solution for all the system's states. It is described in algorithm A1.

The perturbed eigenvalues and the eigenvectors of matrix R will be calculated as a truncation of the corresponding RS-PT series

$$\boldsymbol{E} = \sum_{p=0}^{\infty} \boldsymbol{E}^{(p)} \tag{14}$$

and

88

$$\boldsymbol{Z} = \sum_{p=0}^{\infty} \boldsymbol{Z}^{(p)},\tag{15}$$

respectively, the terms $E^{(p)}$ and $Z^{(p)}$ being the so-called *p*th eigenvalue and eigenvectors corrections. As it was signaled above, the eigenvectors of matrix H will be obtained at the end of the procedure when the change of basis described in equation (9) is performed.

A particular case is obtained when all the $\{\gamma_i\}$ parameters appearing in equation (13) are set to zero. Then, the diagonal of P is null and the unperturbed diagonal matrix $E^{(0)}$ now contains the sum of the zeroth and first order eigenvalue corrections. For this particular case, the following notation is used in order to avoid misinterpretations:

$$\boldsymbol{E}^{[0]} = \boldsymbol{E}^{(0)} + \operatorname{diag}\{\boldsymbol{P}\} = \boldsymbol{E}^{(0)} + \boldsymbol{E}^{(1)}$$
(16)

and

$$\boldsymbol{P}_0 = \text{outdiag}\{\boldsymbol{P}\},\tag{17}$$

where outdiag $\{P\}$ stands for the zero-diagonal matrix containing all the elements $p_{ij} \in P$ for which $i \neq j$.

At the beginning of the algorithm A1, if all the $\{\gamma_i\}$ coefficients are set to zero, the matrices $E^{(0)}$ and P become

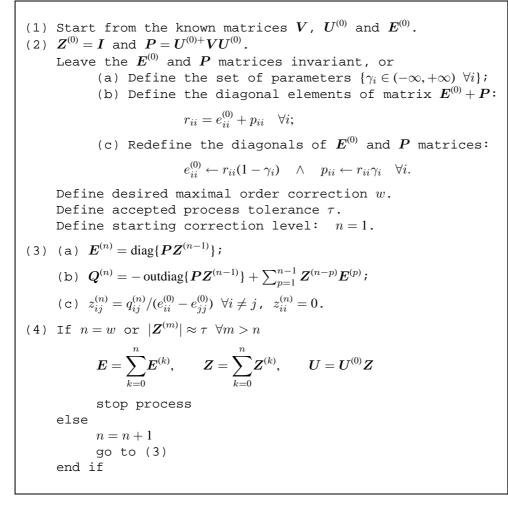
$$\boldsymbol{E}^{(0)} \leftarrow \boldsymbol{E}^{[0]} \tag{18}$$

and

$$P \leftarrow P_0,$$
 (19)

respectively. The whole procedure is slightly simplified because it behaves as if the first order energy corrections are null, $E^{(1)} = 0$. In this case the summation of equation (3b) of the algorithm A1 starts now from p = 2.

Equation (3c) in algorithm A1 can be the origin of divergence problems due to the implied division. The comments outlined above with respect to the choice of a good set of $\{\gamma_i\}$ parameters can help to overcome this division by zero drawback, inherent of RS-PT if present, because the divergence or convergence features of the algorithm A1 can be significantly altered in this way. A simple numerical example will be given below.



Algorithm A1. Full matrix RS-PT algorithm.

4.2. Vector algorithm A2

On the other hand, the study of the particular expressions, attached to a specific state, s, may also lead into the formulation of a vector RS-PT scheme. Algorithm A2 constitutes a computational procedure which can be followed to implement a vectorial PT scheme. The main difference with respect to algorithm A1 can be found in the particular matrix algebraic operation codified in equation (3c) of the mentioned algorithm: this operation is replaced by a multiplication by a predefined diagonal matrix. The vector e_s stands for the canonical vector attached to the *s*th state. The symbol $\delta(\mathcal{E})$ is called a logical Kronecker delta [5,7] and has a unit value if the expression \mathcal{E} is true, being zero otherwise. The same considerations as in the matrix case apply here.

```
(1) Start from the known matrices V , U^{(0)} and E^{(0)} .
(2) Define the state s to be studied.
      oldsymbol{z}^{(0)}=oldsymbol{e}_s and oldsymbol{P}=oldsymbol{U}^{(0)+}oldsymbol{V}oldsymbol{U}^{(0)} .
      Leave the E^{\left(0
ight)} and P matrices invariant, or
                (a) Define the set of parameters \{\gamma_i \in (-\infty, +\infty) \ \forall i\};
                (b) Define the diagonal elements of matrix m{E}^{(0)}+m{P}:
                                   r_{ii} = e_{ii}^{(0)} + p_{ii} \quad \forall i;
                (c) Redefine the diagonals of E^{(0)} and P matrices:
                                    e_{ii}^{(0)} \leftarrow r_{ii}(1-\gamma_i) \quad \land \quad p_{ii} \leftarrow r_{ii}\gamma_i \quad \forall i.
      Define diagonal matrix \Delta = \{d_{ii} = \delta(i=s) + \delta(i\neq s)(e_{ii}^{(0)} - e_{ss}^{(0)})^{-1} \ \forall i\}.
      Define desired maximal order correction w.
      Define accepted process tolerance \tau.
      Define starting correction level: n = 1.
(3) (a) e_{ss}^{(n)} = [\boldsymbol{P}]_s^+ \boldsymbol{z}^{(n-1)};
      (b) \boldsymbol{z}^{(n)} = -\Delta \boldsymbol{P} \boldsymbol{z}^{(n-1)} + \sum_{p=1}^{n-1} e_{ss}^{(p)} \Delta \boldsymbol{z}^{(n-p)}.
(4) If n = w or |\boldsymbol{z}^{(m)}| \approx \tau \ \forall m > n
               e_{ss} = \sum_{k=0}^{n} e_{ss}^{(k)}, \qquad \boldsymbol{z} = \sum_{k=0}^{n} \boldsymbol{z}^{(k)}, \qquad [\boldsymbol{U}]_{s} = \boldsymbol{U}^{(0)} \boldsymbol{z}
                stop process
      else
                n = n + 1
                go to (3)
      end if
```

Algorithm A2. Vector RS-PT algorithm for a particular state s.

90

5. The Gershgorin-like convergence criterion

When working with the P_0 and $E^{[0]}$ matrices as defined in the previous sections, it is possible to establish an empirical rule to be known beforehand while the RS-PT numerical series will converge, when applied to evaluate the *s*th state eigenparameters. This criterion cannot be taken as a theorem, because there is not given here any proof, but provides a sufficient heuristic condition for the PT series convergence. As its formulation is related to the Gershgorin discs theorem [12,22,44,45] it can be called *Gershgorin-like convergence sufficiency criterion*.

For a hermitian matrix \mathbf{R} written as the sum of a diagonal matrix $\mathbf{E}^{[0]}$ plus a zero diagonal matrix \mathbf{P}_0 ,

$$R = E^{[0]} + P_0, (20)$$

the Gershgorin theorem states that every eigenvalue λ_i of R will lie in the following interval, called a Gershgorin disc:

$$\lambda_i \in e_{ii}^{[0]} \pm \sum_{j=1}^n |[P_0]_{ij}|,$$
(21)

 $\{[P_0]_{ij}\}$ being the elements of the matrix P_0 .

Gershgorin-like convergence sufficiency criterion may now be described as follows: The RS-PT series is convergent for the *s*th state if the following condition is fulfilled:

$$\left| e_{ss}^{[0]} - e_{ii}^{[0]} \right| > \sum_{j=1}^{n} \left(\left| [\boldsymbol{P}_0]_{sj} \right| + \left| [\boldsymbol{P}_0]_{ij} \right| \right) \quad \forall i \neq s.$$
(22)

This is equivalent to saying that the RS-PT series converges for the *s*th state if the corresponding Gershgorin disc does not overlap with any of the remaining ones.

A computational test of the trueness of the Gershgorin-like convergence criterion has been performed in a large number (> 10^6) of randomly constructed matrices of varied dimensions ($2 \le N \le 200$). No failures have been encountered.

The Gershgorin convergence criteria has more practical features than the wellknown Kato–Rellich theorem [3,28,32,46]. This last theorem is based on the *existence* of a couple of scalars involved in an inequality which must be fulfilled in all the operators domain. It is a sort of existence theorem and cannot be implemented in a program code in an immediate form.

6. Application examples

6.1. A naïve illustration of algorithm A2

As a first numerical example, figure 1 summarizes the results obtained from algorithm A2 for the following 2×2 matrix expressed in the form of equation (7):

E. Besalú, R. Carbó-Dorca / Rayleigh-Schrödinger perturbation theory

$$\boldsymbol{R} = \begin{pmatrix} 1 & 0\\ 0 & 3 \end{pmatrix} + 0.7 \begin{pmatrix} -1 & 1\\ 1 & 1 \end{pmatrix}.$$
(23)

The segments connecting squares represent the first 10 successive partial series terms for the first eigenvalue when the problem is treated as in the classical PT. It can be seen how the convergence is improved when the matrix to diagonalize is treated defining $\gamma_1 = \gamma_2 = 0$ in equation (13). The present algorithm results are represented in the same figure by the segments connecting the circles. Of course the result is due to the fact the PT problem to solve now becomes

$$\boldsymbol{R} = \begin{pmatrix} 0.3 & 0\\ 0 & 3.7 \end{pmatrix} + 0.7 \begin{pmatrix} 0 & 1\\ 1 & 0 \end{pmatrix}, \tag{24}$$

and the separation between the unperturbed values has increased. The purpose of this simple example is to show the matrix manipulation described in section 3 and how the convergence characteristics can be improved in this way by means of the definition of the $\{\gamma_i\}$ parameter set.

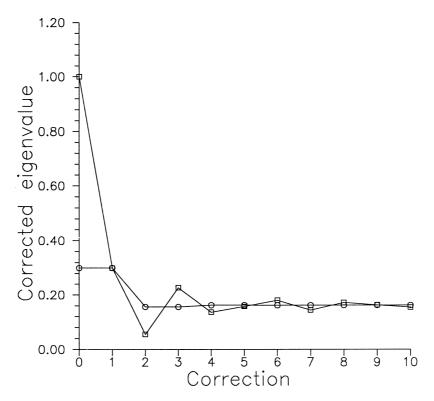


Figure 1. Graphical demonstration of how the convergence features of the RS-PT series can be improved by means of the definition of the $\{\gamma_i\}$ parameters. See text for meaning.

92

Table 1

Obtained results for several He-like atoms. Tabulated data are: Atom symbol, atomic number (Z), number of 1s STO basis functions (n), α and β well-tempered parameters taken from [30], full CI energy in atomic units, number of states entering into the calculation, MOC: maximal order correction needed to reproduce the full CI energy using the vectorial PT algorithm A2. The last column indicates if the convergence is ensured following the Gershgorin-like criterion. Pd result was obtained using a modified α value.

Atom	Z	n	α	eta	Full CI	States	MOC	EC
He	2	4	0.853931	1.660897	-2.878841857	10	9	Ν
С	6	7	0.705082	1.669878	-32.37619653	28	6	Y
Ne	10	7	1.187022	1.683000	-93.87577833	28	6	Y
S	16	10	1.187022	1.683000	-246.1255179	55	4	Y
Ti	22	11	0.420420	1.481960	-470.3747763	66	4	Y
Br	35	11	0.980206	1.392420	-1203.249445	66	3	Y
Pd	46	13	1.678062	1.339474	-2087.375226	91	4	Y
Ι	53	13	0.854821	1.409715	-2775.999807	91	3	Y

6.2. Two-electron full CI calculations

In [8,38] full CI calculations of a He isoelectronic family using spherical functions are reported. The one-electron basis set was taken as a sequence of well-tempered 1s AO STO functions described by Koga et al. [30]. The number of functions was also kept and were taken as defined in the full electron atoms.

Here, the same full CI energy and wavefunctions are obtained using the vector algorithm A2 for the fundamental state. The matrix to diagonalize has been split directly into the diagonal part and the off-diagonal one, following a numerical scheme attached to the form which appears in equation (20). The iterative procedure was stopped when the reported full CI energy (see [8,30,38]) was reproduced. Table 1 summarizes the obtained results together with the maximal order correction needed to achieve the energy value and a label with signals if the Gershgorin-like criterion did *a priori* assure convergence.

7. Conclusions

After a matrix formulation of the PT problem, two RS-PT algorithms have been described and it has been shown how they act in a general way as a *variational-like diagonalization procedure when completed*. A heuristic sufficiency convergence criterion has been described and numerical results related to the two-electron atomic full CI calculations have been given.

Note: The FORTRAN 90 codes implementing the algorithms described here are available upon request to the authors.

Acknowledgements

This work has been partially supported by a grant of the Spanish CICYT: #SAF 96-0158. The authors wish to thank referees' comments which have helped to improve the initial manuscript version.

References

- [1] J.P. Ader, Phys. Lett. A 97 (1983) 178.
- [2] V.C. Aguilera-Navarro, F.M. Fernández, R. Guardiola and J. Ros, J. Phys. A 25 (1992) 6379– 6392.
- [3] G.A. Arteca, F.M. Fernández and E.A. Castro, Large Order Perturbation Theory and Summation Methods in Quantum Mechanics (Springer, Berlin, 1990).
- [4] E. Besalú and R. Carbó, J. Math. Chem. 15 (1994) 397-406.
- [5] E. Besalú and R. Carbó, J. Math. Chem. 18 (1995) 37–72.
- [6] R. Carbó and E. Besalú, J. Math. Chem. 13 (1993) 331–342.
- [7] R. Carbó and E. Besalú, Comput. Chem. 18(2) (1994) 117–126.
- [8] R. Carbó-Dorca and E. Besalú, J. Math. Chem. 20 (1996) 263–271.
- [9] R. Carbó, L. Molino and B. Calabuig, J. Comput. Chem. 13 (1992) 155.
- [10] M. Dunn, T.C. Germann, D.Z. Goodson, C.A. Traynor, J.D. Morgan III, D.K. Watson and D.R. Herschbach, J. Chem. Phys. 101(7) (1994) 5987.
- [11] M. Dunn, D.K. Watson, J.R. Walkup and T.C. Germann, J. Chem. Phys. 104(24) (1996) 9870.
- [12] E. Durand, Solutions Numériques des Équations Algébriques, Vol. II (Masson & Cie, Paris, 1961).
- [13] F.M. Fernández, Phys. Rev. A 48(1) (1993) 189.
- [14] F.M. Fernández, Phys. Lett. A 194 (1994) 343-347.
- [15] F.M. Fernández, Phys. Rev. 50(4) (1994) 2953.
- [16] F.M. Fernández, J. Chem. Phys. 103(15) (1995) 6581.
- [17] F.M. Fernández, J. Math. Phys. 36(8) (1995) 3922.
- [18] F.M. Fernández and E.A. Castro, J. Chem. Phys. 98(8) (1993) 6392.
- [19] F.M. Fernández and E.A. Castro, J. Math. Phys. 34(8) (1993) 3670.
- [20] F.M. Fernández and R. Guardiola, J. Phys. A 26 (1993) 7169-7180.
- [21] F.M. Fernández and J.F. Ogilvie, Phys. Lett. A 178 (1993) 11-16.
- [22] G.H. Golub and Ch.F. van Loan, *Matrix Computations* (The Johns Hopkins University Press, London, 1993).
- [23] R. Guida, K. Konishi and H. Suzuki, Ann. Phys. 241 (1995) 152-184.
- [24] W. Janke and H. Kleinert, Phys. Lett. 199 (1995) 287-290.
- [25] W. Janke and H. Kleinert, Phys. Rev. Lett. 75(15) (1995) 2787.
- [26] G. Jolicard, A. Grosjean and J.P. Killingbeck, J. Chem. Phys. 105(14) (1996) 5359.
- [27] G. Jolicard and J.P. Killingbeck, J. Chem. Phys. 103(5) (1995) 1846.
- [28] T. Kato, Perturbation Theory for Linear Operators (Springer, Berlin, 1966).
- [29] J.P. Killingbeck and G. Jolicard, Chem. Phys. Lett. 255 (1996) 79-82.
- [30] T. Koga, H. Tatewaki and A.J. Thakkar, Theor. Chim. Acta 88 (1994) 273.
- [31] C. Møller and M.S. Plesset, Phys. Rev. 46 (1934) 618.
- [32] F. Rellich, Perturbation Theory of Eigenvalue Problems (Gordon and Breach, New York, 1969).
- [33] B.O. Roos and P.E.M. Siegbahn, in: *Methods of Electronic Structure Theory*, Vol. 3, ed. H.F. Shaefer (Plenum Press, New York, 1977).
- [34] I. Shavitt, in: *Methods of Electronic Structure Theory*, Vol. 3, ed. H.F. Shaefer (Plenum Press, New York, 1977).

- [35] E.L. Sibert III, J. Chem. Phys. 88(7) (1988) 4378.
- [36] L. Skála, J. Čížek, J. Dvořák and V. Špirko, Phys. Rev. A 53(4) (1996) 2009.
- [37] A.A. Suvernev and D.Z. Goodson, J. Chem. Phys. 106(7) (1997) 2681.
- [38] TWOEL-96 program, Institute of Computational Chemistry, University of Girona, Girona, Spain (1996).
- [39] F. Vinette and J. Čížek, Comput. Phys. Comm. 52 (1988) 35-41.
- [40] E.J. Weniger, Ann. Phys. 246 (1996) 133-165.
- [41] E.J. Weniger, Int. J. Quant. Chem. 57 (1996) 265-280.
- [42] E.J. Weniger, Phys. Rev. Lett. 77(14) (1996) 2859.
- [43] E.J. Weniger, J. Čížek and F. Vinette, J. Math. Phys. 34(2) (1993) 571.
- [44] J.H. Wilkinson, The Algebraic Eigenvalue Problem (Clarendon Press, Oxford, 1965).
- [45] J.H. Wilkinson and C. Reinsch, Linear Algebra (Springer, Berlin, 1971).
- [46] C.H. Wilsox, ed., *Perturbation Theory and Its Applications in Quantum Mechanics* (Wiley, New York, 1966).
- [47] M.R.M. Witwit and J.P. Killingbeck, J. Phys. A 26 (1993) 3659–3662.