

Residue-squaring iterative diagonalization method for perturbation problems

P M MATHEWS

Department of Theoretical Physics, University of Madras. Madras 600025

MS received 6 December 1974

Abstract. We present a new method for the evaluation of the change in eigenvalues due to a perturbation of strength λ . It is a fast converging iterative method which, at the n th step, gives results accurate to order $(2^{n+1} - 1)$ in λ . Unlike the Rayleigh-Schrödinger perturbation theory in quantum mechanics, which becomes prohibitively cumbersome when carried to higher orders, the present method involves a routine which remains straightforward at all stages.

Keywords. Perturbation theory; residue-squaring.

The determination of the changes in eigenvalues and eigenfunctions of the Hamiltonian of a system under the action of a perturbation is one of the basic problems of quantum mechanics. The purpose of this note is to present an iterative scheme for the solution of the problem, which is straightforward in principle and is a vast improvement over the conventional perturbation theory in the rapidity of convergence.

We view the problem as one of diagonalizing a *matrix* $\mathbf{H} \equiv \mathbf{D} + \mathbf{R}$ whose diagonal part \mathbf{D} is dominant compared to the residual (off-diagonal) part \mathbf{R} , the latter being characterized by a small parameter λ . The basic step in our procedure is a similarity transformation which takes \mathbf{H} over into an equivalent matrix $\mathbf{H}' \equiv \mathbf{D}' + \mathbf{R}'$ wherein \mathbf{R}' as well as $(\mathbf{D}' - \mathbf{D})$ is $O(\lambda^2)$. By iterating this step one gets a succession of matrices \mathbf{H}'' , \mathbf{H}''' , ..., all equivalent to \mathbf{H} , whose off-diagonal parts are of leading orders λ^4 , λ^8 , The *changes* in the diagonal parts, *i.e.*, $\mathbf{D}'' - \mathbf{D}'$, $\mathbf{D}''' - \mathbf{D}''$, ... are also of orders λ^4 , λ^8 , It is clear then that in taking the diagonal elements at any particular stage of iteration to be the eigenvalues of \mathbf{H} , the residual error is of the order of the *square* of the residual error at the previous stage: What we have is a "residue-squaring" process. The elements of $\mathbf{D}^{(n)}$ give the eigenvalues of \mathbf{H} correct to the $(2^{n+1} - 1)$ th order in λ .

Construction of the residue-squaring transformation is straightforward. Let

$$\mathbf{H}' = (\mathbf{I} - \mathbf{F})^{-1} \mathbf{H} (\mathbf{I} - \mathbf{F}) \quad (1)$$

where the matrix \mathbf{F} is taken to be $O(\lambda)$. Introducing $\mathbf{H} = \mathbf{D} + \mathbf{R}$, we reduce \mathbf{H}' to

$$\mathbf{D} + (\mathbf{I} - \mathbf{F})^{-1} (\mathbf{R} - \mathbf{D}\mathbf{F} + \mathbf{F}\mathbf{D} - \mathbf{R}\mathbf{F}). \quad (2)$$

If we now choose \mathbf{F} such that

$$[\mathbf{D}, \mathbf{F}] = \mathbf{R} \quad (3)$$

\mathbf{H}' becomes

$$\mathbf{H}' = \mathbf{D} - (\mathbf{I} - \mathbf{F})^{-1} \mathbf{R} \mathbf{F} \quad (4)$$

Note that eq. (3) is consistent with \mathbf{F} being $O(\lambda)$. The last term in (4) is now clearly $O(\lambda^2)$. Thus we have succeeded in finding \mathbf{H}' , equivalent to \mathbf{H} , whose off-diagonal part is $O(\lambda^2)$. Rewriting \mathbf{H}' as $\mathbf{H}' = \mathbf{D}' + \mathbf{R}'$ with

$$\mathbf{D}' = \mathbf{D} - [(\mathbf{I} - \mathbf{F})^{-1} \mathbf{R} \mathbf{F}]_d \quad (5 a)$$

$$\mathbf{R}' = - [(\mathbf{I} - \mathbf{F})^{-1} \mathbf{R} \mathbf{F}]_{o,d} \quad (5 b)$$

(where the subscripts *d* and *o,d* stand for the diagonal and off-diagonal parts respectively) we can repeat the above process to obtain \mathbf{H}'' , and so on. The diagonalised form of \mathbf{H} is then

$$\begin{aligned} \mathbf{D}^{(\infty)} = \mathbf{D} - [(\mathbf{I} - \mathbf{F})^{-1} \mathbf{R} \mathbf{F}]_d - [(\mathbf{I} - \mathbf{F}')^{-1} \mathbf{R}' \mathbf{F}']_d - \\ - [(\mathbf{I} - \mathbf{F}'')^{-1} \mathbf{R}'' \mathbf{F}'']_d + \dots \end{aligned} \quad (6)$$

and the eigenvectors of \mathbf{H} are given by

$$u_{(r)} = [\dots (\mathbf{I} - \mathbf{F}'')^{-1} (\mathbf{I} - \mathbf{F}')^{-1} (\mathbf{I} - \mathbf{F})^{-1}] e_{(r)} \quad (7)$$

where $e_{(r)}$, the r th eigenvector of $\mathbf{D}^{(\infty)}$, has elements $(e_{(r)})_s = \delta_{rs}$.

Returning to eq. (3) and writing it in terms of matrix elements:

$$(d_m - d_n) F_{mn} = R_{mn} \quad (8)$$

we observe that off-diagonality of \mathbf{R} is indeed required by this equation. It follows that \mathbf{F} also can be consistently chosen to be off-diagonal. Incidentally, it is important to note that in the quantum mechanical context where \mathbf{H} is the matrix of the total Hamiltonian in the unperturbed representation, the diagonal elements $d_m \equiv H_{mm}$ include the diagonal part of the perturbation: $d_m = H_{mm}^0 + \lambda V_{mm}$. The elements of the residue \mathbf{R} are $\lambda V_{mn} (\mathbf{I} - \delta_{mn})$.

If it so happens that r of the d_m are equal or differ from each other by amounts of the same order as the corresponding R_{mn} , we shall speak of an r -fold degeneracy occurring. The elements F_{mn} pertaining to these rows and columns would in general become infinite or large (instead of being of order λ). But this difficulty can be circumvented simply by "preconditioning" \mathbf{H} by a similarity transformation which diagonalizes the $r \times r$ submatrix involving the "degenerate" diagonal elements. It will be assumed that this preconditioning is done before commencing the residue-squaring process, so that all the R_{mn} in this $r \times r$ block are zero and the corresponding F_{mn} too are chosen zero.

Finally it may be noted that if $(\mathbf{I} - \mathbf{F})^{-1}$ itself is to be evaluated perturbatively, the following rearrangement of the binomial series is most convenient:

$$(\mathbf{I} - \mathbf{F})^{-1} = (\mathbf{I} + \mathbf{F})(\mathbf{I} + \mathbf{F}^2)(\mathbf{I} + \mathbf{F}^4)(\mathbf{I} + \mathbf{F}^8) \dots \quad (9)$$

If, for instance the diagonalisation process is to be carried up to \mathbf{D}'' (residual error of order λ^8), the first three factors in eq. (9) would give $(\mathbf{I} - \mathbf{F})^{-1}$ to sufficient accuracy, and at the second stage one can take $(\mathbf{I} - \mathbf{F}')^{-1} \approx (\mathbf{I} + \mathbf{F}')$.

A detailed account of the method and its variations and extensions (*e.g.*, to matrices which are not diagonal-dominant), together with a discussion in comparison with other methods, will be published separately.