Comparison between the projection operator and continued fraction approaches to perturbation theory

1981 J. Phys. A: Math. Gen. 14 3169

(http://iopscience.iop.org/0305-4470/14/12/014)

J. Phys. A: Math. Gen. **14** (1981) 3169–3179. Printed in Great Britain

Comparison between the projection operator and continued fraction approaches to perturbation theory

R I Jackson and S Swain

Department of Applied Mathematics and Theoretical Physics, The Queen's University of Belfast, Belfast BT7 1NN, Northern Ireland

Received 8 April 1981

Abstract. The connection between the projection operator approach and the continued fraction approach to perturbation theory is investigated. A concise solution to the linear operator equation $A|x\rangle = |b\rangle$ is found in terms of the level shift operator using projection operator techniques. The analogous process to the use of projection operators in the continued fraction method of solving the same problem is identified, and a parallel development performed. The connection between the two approaches is thereby established, and continued fraction expressions for the level shift operator obtained. The abstract equation is then specialised to deal with (i) the eigenvalue problem and (ii) the calculation of transition probabilities for quantum mechanical systems described by a time-independent Hamiltonian. Particular attention is paid to the problem of degeneracy and it is shown that the most convenient expressions are found by a hybrid of the two approaches.

1. Introduction

A standard way of developing perturbation theory utilises projection operators and the formal elimination of unwanted quantities by manipulation of operator equations. These techniques are discussed in standard textbooks such as Messiah (1962) and Goldberger and Watson (1964) (see also Cohen–Tannoudji 1968). The fundamental properties of the system, such as the shifts in the energy levels and the induced transition due to a perturbation V, are determined by the level shift operator R, which is usually expanded in a power series in V.

An alternative to the projection operator method is to use continued fraction expansions; perturbation theories in terms of such quantities were first obtained by Feenberg (1948a, b) (see also Richards 1948, Feshbach 1948) for stationary state problems. Recently, continued fraction expressions for transition probabilities have been obtained (Swain 1975), and the general method of obtaining continued fraction solutions to systems of linear equations has been considered as a problem in linear algebra, thereby emphasising its generality (Swain 1976, 1977).

High-order perturbation theory may be straightforwardly developed using either of these formalisms.

One of the advantages of the projection operator approach is the neatness with which it can tackle problems involving degeneracy. The projection operator can be defined to span only the degenerate states and one is left with a matrix equation whose dimension is equal to the degree of degeneracy, which is usually small. The continued

0305-4470/81/123169+11$01.50

fraction method can also be applied to degenerate problems (Swain 1977) but the resulting expressions are complicated.

The continued fraction approach has the advantage that self-energy effects are treated naturally in this formalism, so that the method is particularly useful for dealing with saturation and other nonlinear effects. In the projection operator approach the resolvent operator is usually developed in a power series expansion, which is not suitable for describing saturation. The level shift operator can be developed in a continued fraction representation using projection operator techniques quite generally (Mower 1980) but the derivation is complicated, making use of repeated partitioning of the projection operators, and the final result is complex. Löwdin (1962) has discussed the connection between projection operator and matrix partitioning techniques and has obtained some formal continued fraction expressions.

The object of the present investigations is to develop the connection between the projection operator and continued fraction approaches, thereby making it possible to take advantage of the strengths of both. The fundamental step is to identify the process analogous to the use of the projection operator in the continued fraction approach. By these means we are able to obtain simple continued fraction expressions for the level shift operator, and the method of derivation is also considerably more concise than those used previously.

To facilitate comparisons the projection operator approach is briefly described in § 2 and the level shift and projection operators introduced. The problem considered is the solution of a set of linear inhomogeneous equations which by appropriate specialisation may be used to describe the calculation of perturbed energy levels or transition probabilities. The same problem is considered in § 3 from the continued fraction viewpoint, the essential step being to determine the procedure which is analogous to the use of projection operators. In § 4 the connection between the two approaches is established, and continued fraction series for the level shift operator set down. In §§ 5 and 6 the formalism is applied to the eigenvalue problem and to the calculation of transition probabilities respectively, in both cases particular attention being paid to dealing with degenerate problems.

2. The projection operator technique

In this section we outline the projection operator method using a concise but general approach. Consider the equation

$$A|x\rangle = |b\rangle \tag{1}$$

which we wish to solve for $|x\rangle$. We introduce the projection operators P and $Q = 1 - P$ which satisfy

$$PQ = QP = 0 \tag{2a}$$

$$P^2 = P \qquad Q^2 = Q. \tag{2b}$$

For the moment we do not specify these quantities further. The next step is to separate the operator A into a part A^0 which commutes with P (and therefore with Q) and a remaining part A^1. That is, we set

$$A = A^0 + A^1 \tag{3a}$$

where

$$[A^0, P] = [A^0, Q] = 0 \tag{3b}$$

(in typical applications we take A^0 to be some unperturbed Hamiltonian and A^1 the perturbation). Since $P + Q = 1$ we may write equation (1) in the form

$$(A^0 + A^1)(P + Q)|x\rangle = |b\rangle. \tag{4}$$

Operating on equation (4) first with P, then with Q, gives us the two equations

$$(A^0 + PA^1)P|x\rangle + PA^1Q|x\rangle = P|b\rangle \tag{5a}$$

$$(A^0 + QA^1)Q|x\rangle + QA^1P|x\rangle = Q|b\rangle \tag{5b}$$

where we have made use of the properties (2a) and (2b). If the operator $(A^0 + QA^1)^{-1}$ exists, we may solve (5b) formally for $Q|x\rangle$:

$$Q|x\rangle = (A^0 + QA^1)^{-1}(Q|b\rangle - QA^1P|x\rangle). \tag{6}$$

This expression may be used to eliminate $Q|x\rangle$ from the left-hand side of equation (5a) to give

$$\{A^0 + PA^1P - PA^1(A^0 + QA^1)^{-1}QA^1P\}P|x\rangle = P\{1 - PA^1(A^0 + QA^1)^{-1}Q\}|b\rangle \tag{7}$$

an equation for $P|x\rangle$ only. In terms of the *level shift operator*, R,

$$R \equiv PA^1P - PA^1(A^0 + QA^1)^{-1}QA^1P \tag{8}$$

and the operator, S,

$$S \equiv P - PA^1(A^0 + QA^1)^{-1}Q \tag{9}$$

equation (7) becomes

$$(A^0 + R)P|x\rangle = S|b\rangle. \tag{10}$$

Equation (10) is the main result of the projection operator approach. The particular form to which it reduces depends upon the choice of the projection operator P. Thus if we take

$$P \equiv |1\rangle\langle 1| \tag{11}$$

we obtain

$$(A^0 + R)|1\rangle\langle 1|x\rangle = \sum_i S|i\rangle\langle i|b\rangle \tag{12}$$

where we have introduced a complete set of states $\Sigma_i |i\rangle\langle i| = 1$ on the right-hand side. Operating with $\langle 1|$ from the left on equation (12) gives

$$(A^0_{11} + R_{11})x_1 = \sum_i S_{1i}b_i \tag{13}$$

where we have used the notation $\langle 1|x\rangle = x_1$, $\langle i|b\rangle = b_i$, $\langle 1|S|i\rangle = S_{1i}$, etc. Thus, by using the projection operator (11), we obtain the component x_1 of the vector $|x\rangle$ in terms of the matrix elements R_{11} and S_{1i}.

A more general choice of projection operator is

$$P = \sum_{\alpha=1}^{n} |\alpha\rangle\langle\alpha|. \tag{14}$$

If the vectors $|\alpha\rangle$ are all eigenvectors of A^0 then the conditions (3*b*) are satisfied. Proceeding as before we obtain the equations

$$\sum_{\beta=1}^{n} (A^0_{\alpha\alpha}\delta_{\alpha\beta} + R_{\alpha\beta})x_\beta = \sum_i S_{\alpha i}b_i, \qquad \alpha = 1, 2, \ldots, n. \tag{15}$$

Thus the x_β are determined by the n equations (15). In the remainder of this paper we use Greek labels to indicate the n vectors singled out in equation (14).

In order to calculate x_1 or the x_β one needs explicit expressions for the matrix elements $R_{\alpha\beta}$ and $S_{\alpha i}$. The usual procedure is to develop these in a power series expansion.

Iterating the operator relation

$$(\lambda + \mu)^{-1} = \lambda^{-1} - \lambda^{-1}\mu(\lambda + \mu)^{-1} \tag{16}$$

one obtains the expansion

$$\frac{1}{A^0 + QA^1} = \frac{1}{A^0} - \frac{Q}{A^0}A^1\frac{1}{A^0} + \frac{Q}{A^0}A^1\frac{Q}{A^0}A^1\frac{1}{A^0} + \ldots \tag{17}$$

which, after substitution into equations (8) and (9), gives the power series expansions

$$R = P\sum_{k=0}^{\infty}\left(-A^1\frac{Q}{A^0}\right)^k A^1 P \tag{18}$$

$$S = P\sum_{k=0}^{\infty}\left(-A^1\frac{Q}{A^0}\right)^k. \tag{19}$$

It is clear from either expressions (8) and (9) or (18) and (19) that R and S are related:

$$R = SA^1P. \tag{20}$$

In terms of matrix elements

$$R_{\alpha\beta} = \sum_i S_{\alpha i}A^1_{i\beta}. \tag{21}$$

The present formalism is abstract but by choosing appropriate forms for A, $|x\rangle$ and $|b\rangle$ we describe either time-dependent or time-independent perturbation theory, as we demonstrate in §§ 5 and 6.

3. The continued fraction technique

We now investigate the type of problem posed in § 2 using continued fraction techniques.

Consider the set of linear equations

$$\sum_{j=1}^{N} a_{ij}x_j = b_i, \qquad i = 1, 2, \ldots, N. \tag{22}$$

We have shown that the solution of such a set in terms of continued fractions (Swain 1976) is given by

$$x_j = \frac{b_j}{\mathscr{D}_j} - \sum_k{}^*\frac{a_{jk}b_k}{\mathscr{D}_{jk}} + \sum_{k,l}{}^*\frac{a_{jk}a_{kl}b_l}{\mathscr{D}_{jkl}} + \ldots \tag{23}$$

where the notation of that paper is followed, and the $\mathscr{D}$ functions are defined by the relations

$$\mathscr{D}_{jk} \equiv \mathscr{D}_j \mathscr{D}_k^{(j)}, \qquad \mathscr{D}_{jkl} \equiv \mathscr{D}_j \mathscr{D}_k^{(j)} \mathscr{D}_l^{(jk)}, \ldots \text{etc} \tag{24a}$$

$$\mathscr{D}_j = a_{jj} - \sum_r{}^* \frac{a_{jr}a_{rj}}{\mathscr{D}_r^{(j)}} + \sum_{r,s}{}^* \frac{a_{jr}a_{rs}a_{sj}}{\mathscr{D}_{rs}^{(jr)}} + \ldots \tag{24b}$$

$$\mathscr{D}_m^{(ij\ldots n)} = a_{mm} - \sum_r{}^* \frac{a_{mr}a_{rm}}{\mathscr{D}_r^{(ij\ldots nm)}} + \sum_{r,s}{}^* \frac{a_{mr}a_{rs}a_{sm}}{\mathscr{D}_{rs}^{(ij\ldots nm)}} + \ldots . \tag{24c}$$

Thus the $\mathscr{D}$ are defined iteratively which results in their taking a continued fraction form (or more exactly, the form of a sum of series of products of continued fractions). Note that in expressions (23), (24*b*) and (24*c*) the sums over $j, k \ldots ; r, s \ldots$ etc are such that no diagonal matrix elements a_{jj} appear.

Here we wish to find a way of handling the set (22) which is analogous to the projection operator methods used in § 2. This can be done by splitting the x_i into two sets, $x_\alpha : \alpha = 1, 2, \ldots, n$ and $x_j : j = n+1, n+2, \ldots, N$. We use only Greek subscripts to describe members of the first set which involve the same states as those employed in the definition of the projection operator of equation (14). The procedure is to solve the second set amongst themselves, regarding the members of the first set as inhomogeneous terms wherever they appear in the equations. That is, we consider the set of equations

$$\sum_{j=n+1}^{N} a_{ij}x_j = b_i - \sum_{\alpha=1}^{n} a_{i\alpha}x_\alpha \equiv c_i, \qquad i = n+1, \ldots, N. \tag{25}$$

This set may be solved using equation (23), when we obtain

$$x_j = \frac{c_j}{\mathscr{D}_j^{(1\ldots n)}} - \sum_k{}^* \frac{a_{jk}c_k}{\mathscr{D}_{jk}^{(1\ldots n)}} + \sum_{k,l}{}^* \frac{a_{jk}a_{kl}c_l}{\mathscr{D}_{jkl}^{(1\ldots n)}} + \ldots . \tag{26}$$

The superscripts $(1, 2, \ldots, n)$ on the $\mathscr{D}$ functions in equation (26) indicate that j, and the variables summed over, $k, l, \ldots$, cannot take any of the values $1, 2, \ldots, n$, and that these states are also *excluded* in the calculation of the $\mathscr{D}$ functions. (This must clearly be so as i and j on the left-hand side of equation (25) do not take the values $1, 2, \ldots n$.)

Using $c_i = b_i - \Sigma_\alpha a_{i\alpha}x_\alpha$ in equation (26), the latter may be rewritten as

$$x_j = \frac{b_j}{\mathscr{D}_j^{(1\ldots n)}} - \sum_k{}^* \frac{a_{jk}b_k}{\mathscr{D}_{jk}^{(1\ldots n)}} + \sum_{k,l}{}^* \frac{a_{jk}a_{kl}b_l}{\mathscr{D}_{jkl}^{(1\ldots n)}} + \ldots$$

$$- \sum_\alpha \left(\frac{a_{j\alpha}}{\mathscr{D}_j^{(1\ldots n)}} - \sum_k{}^* \frac{a_{jk}a_{k\alpha}}{\mathscr{D}_{jk}^{(1\ldots n)}} + \sum_{k,l}{}^* \frac{a_{jk}a_{kl}a_{l\alpha}}{\mathscr{D}_{jkl}^{(1\ldots n)}} + \ldots \right) x_\alpha . \tag{27}$$

We may now substitute the above expressions for the x_j, $j = n+1, \ldots, N$, into the remaining set of equations for the x_α

$$\sum_{\beta=1}^{n} a_{\alpha\beta}x_\beta + \sum_{j=n+1}^{N} a_{\alpha j}x_j = b_\alpha, \qquad \alpha = 1, 2, \ldots, n \tag{28}$$

when, after a little simplification, one obtains

$$\sum_{\beta=1}^{n}\left(a_{\alpha\beta}-\sum_{j=n+1}^{N}{}^{*}\frac{a_{\alpha j}a_{j\beta}}{\mathscr{D}_{j}^{(1\ldots n)}}+\sum_{j,k=n+1}^{N}{}^{*}\frac{a_{\alpha j}a_{jk}a_{k\beta}}{\mathscr{D}_{jk}^{(1\ldots n)}}+\ldots\right)x_{\beta}$$

$$=b_{\alpha}-\sum_{j=n+1}^{N}{}^{*}\frac{a_{\alpha j}b_{j}}{\mathscr{D}_{j}^{(1\ldots n)}}+\sum_{j,k=n+1}^{N}{}^{*}\frac{a_{\alpha j}a_{jk}b_{k}}{\mathscr{D}_{jk}^{(1\ldots n)}}+\ldots,\qquad\alpha=1,2,\ldots,n.\tag{29}$$

Expression (29) is the main result of the continued fraction approach; it defines a set of equations connecting the x_{α} variables only. Thus it represents the partial solution of the set of equations (22) obtained by eliminating the subset of variables $x_{n+1}, x_{n+2}, \ldots, x_{N}$. (Although we have derived our results for finite N we assume that they hold also for N infinite.) Expression (29) is more general than expression (23) in that it includes the latter as a special case. Thus if we set $n=1$, $\alpha=\beta$ in equation (29) we recover equation (23) if we bear in mind that the expression in parentheses in the former is then just equal to $\mathscr{D}_{\alpha}$.

Equation (29) obtained using continued fraction methods corresponds to equation (15) obtained using projection operator techniques. The use of the projection operator (14) involving the n states $\alpha=1,2,\ldots,n$ is clearly equivalent to eliminating all variables except $x_1, x_2, \ldots, x_n$ from the linear equations (22).

4. Comparison of projection operator and continued fraction methods

Equations (29) and (15) must obviously be equivalent; by comparing the two we are led to the correspondences

$$R_{\alpha\beta}=a_{\alpha\beta}-\sum_{j=n+1}^{N}{}^{*}\frac{a_{\alpha j}a_{j\beta}}{\mathscr{D}_{j}^{(1\ldots n)}}+\sum_{j,k=n+1}^{N}\frac{a_{\alpha j}a_{jk}a_{k\beta}}{\mathscr{D}_{jk}^{(1\ldots n)}}+\ldots,\qquad\alpha\neq\beta\tag{30a}$$

$$R_{\alpha\alpha}=a_{\alpha\alpha}-A_{\alpha\alpha}^{0}-\sum_{j=n+1}^{N}{}^{*}\frac{a_{\alpha j}a_{j\alpha}}{\mathscr{D}_{j}^{(1\ldots n)}}+\sum_{j,k=n+1}^{N}\frac{a_{\alpha j}a_{jk}a_{k\alpha}}{\mathscr{D}_{jk}^{(1\ldots n)}}+\ldots=\mathscr{D}_{\alpha}^{(1\ldots n)}-A_{\alpha\alpha}^{0}\tag{30b}$$

$$S_{\alpha i}=-\frac{a_{\alpha i}}{\mathscr{D}_{i}^{(1\ldots n)}}+\sum_{j=n+1}^{N}\frac{a_{\alpha j}a_{ji}}{\mathscr{D}_{ij}^{(1\ldots n)}}-\sum_{j,k=n+1}^{N}\frac{a_{\alpha j}a_{jk}a_{ki}}{\mathscr{D}_{ijk}^{(1\ldots n)}}+\ldots,\qquad i=n+1,n+2,\ldots,N\tag{31a}$$

$$S_{\alpha\beta}=\delta_{\alpha\beta},\qquad\alpha,\beta=1,2,\ldots,n.\tag{31b}$$

Equation (30a) defines $R_{\alpha\beta}$ only for α, β members of the set $1,2,\ldots,n$; however, we are free to extend the definition by allowing α, β to range over the full set of values $1,2,\ldots,N$. Then we may write equation (31a) as

$$S_{\alpha i}=-R_{\alpha i}/\mathscr{D}_{i}^{(1\ldots n)},\qquad i\neq 1,2,\ldots,n.\tag{32}$$

The second line of equation (30b) follows from making use of equation (24). The fact that the $A_{\alpha\alpha}^{0}$ term need not cancel the $a_{\alpha\alpha}$ term in equation (30b) is a consequence of the fact that A^{1} may still have non-zero diagonal elements. If we write $A=A^{0}+A^{1}$ where A^{0} is diagonal we may write equations (30) as

$$R_{\alpha\beta}=A_{\alpha\beta}^{1}-\sum_{j=n+1}^{N}{}^{*}\frac{A_{\alpha j}^{1}A_{j\beta}^{1}}{\mathscr{D}_{j}^{(1\ldots n)}}+\sum_{j,k=n+1}^{N}\frac{A_{\alpha j}^{1}A_{jk}^{1}A_{k\beta}^{1}}{\mathscr{D}_{jk}^{(1\ldots n)}}+\ldots\tag{33}$$

thereby obtaining a continued fraction expression for the level shift operator.

By these equivalences we may write equations (29) in the form of equation (15) with the $R_{\alpha\beta}$ and $S_{\alpha i}$ given explicitly by the continued fraction expressions (30)–(33).

It will be observed that expressions (30)–(32) are consistent with equation (21).

Expression (33) provides a continued fraction alternative to the power series expansion (18). In addition to providing a natural way of describing saturation effects, continued fractions usually have much better convergence properties than power series.

Although this has not been made explicit, we should emphasise that the form of the level shift operator depends upon the projection operator being used. We may indicate this dependence by writing $R_{\alpha\beta}$ as given, for example, by equation (33), as $R_{\alpha\beta}^{(1\ldots n)}$ where the superscripts 1, 2, ..., n indicate the states employed in the definition of the projection operator. Making use of this notation and of equation (32) we may write equation (29) in the combined level shift operator/continued fraction form

$$\sum_{\beta}(A_{\alpha\alpha}^{0}\delta_{\alpha\beta}+R_{\alpha\beta}^{(1\ldots n)})x_{\beta}=b_{\alpha}-\sum_{i=n+1}^{N}\frac{R_{\alpha i}^{(1\ldots n)}}{\mathscr{D}_{i}^{(1\ldots n)}}b_{i}, \qquad \beta=1,2,\ldots,n. \tag{34}$$

Finally, we note that there is a connection between the level shift operators defined for different projection operators. For example, we have the relation

$$R_{\alpha\beta}^{(1\ldots n)}=R_{\alpha\beta}^{(1\ldots nn+1)}-R_{\alpha n+1}^{(1\ldots n+1)}\frac{1}{\mathscr{D}_{n+1}^{(1\ldots n)}}R_{n+1\beta}^{(1\ldots n+1)}. \tag{35}$$

This equation may be understood as follows. Looking at expression (30a) we see that $R_{\alpha\beta}^{(1\ldots n)}$ is the alternating sum of all processes which connect the states α and β, the intermediate states being all different from each other and from the projected states 1, 2, ..., n. This sum may be decomposed into the sum of all processes connecting α and β as before but excluding those which pass through the state $n+1$ plus the sum of all processes which do pass through the state $n+1$. These are just the first and second terms of equation (35) respectively. Operator equations analogous to equation (35) were used with a hierarchy of level shift operators by Mower (1980) to generate operator continued fraction expressions.

5. The eigenvalue problem

Let us now indicate some physical situations to which our formalism may be applied. First we consider the eigenvalue problem where we wish to determine the perturbed energy levels of some Hamiltonian operator $H=H^{0}+V$. Setting $|b\rangle\equiv 0$, $A^{0}\equiv\lambda-H^{0}$, $A^{1}=-V$, where λ is one of the eigenvalues to be determined, equation (1) becomes

$$(\lambda-H^{0}-V)|x\rangle=0. \tag{36}$$

If the states $|i\rangle$ are chosen to be eigenvectors of H^{0}

$$H^{0}|i\rangle=E_{i}^{0}|i\rangle, \qquad i=1,2,\ldots \tag{37}$$

and we take $P=|1\rangle\langle 1|$, then equation (13) gives

$$\lambda-E_{1}^{0}+R_{11}(\lambda)=0 \tag{38}$$

(assuming $x_1 \neq 0$). Using the power series expansion (18) for R, we obtain

$$\lambda = E_1^0 + V_{11} + \sum_{j\neq 1} V_{1j} \frac{1}{\lambda - E_j^0 - V_{jj}} V_{j1} + \sum_{j\neq 1} \sum_{k\neq 1} V_{1j} \frac{1}{\lambda - E_j^0 - V_{jj}} V_{jk} \frac{1}{\lambda - E_k^0 - V_{kk}} V_{k1} + \ldots \tag{39}$$

which is usually solved iteratively for λ, thereby giving the eigenvalue of H which reduces to E_1^0 when V tends to zero.

In the continued fraction approach we take $n = 1$, $b_i = 0$ in equation (29) and replace all the diagonal elements a_{ii} of the matrix A by $\lambda - E_i^0 - V_{ii}$ and the diagonal elements a_{ij} by $-V_{ij}$. Then we obtain

$$E_1^0 + V_{11} - \lambda + \sum_{j\neq 1}^{*} \frac{V_{1j}V_{j1}}{\lambda - E_j^0 - V_{jj} - \Sigma_{k\neq 1,j}^{*} V_{jk}V_{kj}/(\lambda - E_k^0 - V_{kk} - \ldots)} + \sum_{j,k\neq 1}^{*} \frac{V_{1j}V_{jk}V_{k1}}{(\lambda - E_j^0 - V_{jj} - \ldots)(\lambda - E_k^0 - V_{kk} - \ldots)} + \ldots = 0 \tag{40}$$

where we have written out explicitly the first few terms of the $\mathscr{D}$ functions using equation (24*b*). The power series expansion of expression (40) agrees with equation (29). For a finite number of levels, N, equation (40) is a finite series and it gives the exact Nth-order polynomial for the eigenvalues. In practice it is again usually solved iteratively for the particular eigenvalue which reduces to E_1^0 as V tends to zero.

When the levels are degenerate this approach needs to be modified. For example, if the states $|1\rangle$ and $|2\rangle$ are degenerate, $E_1^0 = E_2^0$, terms such as $E_1^0 - E_2^0 - V_{22}$ appear in the denominators of equation (39) in the course of the iterative solution. If $V_{jj} = 0$ (all j) (as is usually the case) this gives rise to a singularity. In equation (39) it means that the $\mathscr{D}$ functions become of order V (or of order V^2 if $V_{jj} = 0$) instead of order one as in the case of non-degeneracy, and the iteration scheme breaks down.

To illustrate the procedure in the case of degeneracy we consider the case where we have two levels, $|1\rangle$ and $|2\rangle$ degenerate. We take $b_i = 0$, and $P = \Sigma_{\alpha=1}^{2} |\alpha\rangle\langle\alpha|$ in equation (15), or equivalently $n = 2$ in equations (29) or (34). Using the fact that the determinant of the resulting set of equations must be zero for consistent solutions we obtain

$$\begin{vmatrix} \lambda - E_1^0 + R_{11}(\lambda) & R_{12}(\lambda) \\ R_{21}(\lambda) & \lambda - E_2^0 + R_{22}(\lambda) \end{vmatrix} = 0. \tag{41}$$

This equation treats the two states $|1\rangle$ and $|2\rangle$ on an equal footing and may be solved to lowest order as a quadratic equation in the λ's. An iterative scheme may now be safely used to obtain the higher-order solutions, because the unperturbed energies E_1^0 and E_2^0 *no longer appear explicitly in the denominators of the* R_{ij}. This may be seen from equation (30), where the sums over j and k on the right-hand side extend over 3, 4, . . . , or from equation (8) where the factors $Q = 1 - |1\rangle\langle 1| - |2\rangle\langle 2|$ prevent reference to the states $|1\rangle$ and $|2\rangle$. Equation (41) is equivalent to equation (23) of Swain (1977), or equation (86) of Mower (1980). The solution of this equation is discussed further in both papers, but we should point out that in the former there is a sign error in equation (28). (The right-hand side should be multiplied by minus one, which means that in equations (32),

(36) and (37) the factors $\bar{E}-E_\alpha$ and $\lambda^{(1)}-E_\alpha$ in the denominators should also be multiplied by minus one.)

Clearly if we had the three states $|1\rangle$, $|2\rangle$, $|3\rangle$ degenerate or nearly degenerate, we would take $P=\Sigma_{\alpha=1}^{3}|\alpha\rangle\langle\alpha|$ in equation (15), or $n=3$ in equations (37), which lead to

$$\begin{vmatrix} \lambda-E_1^0+R_{11} & R_{12} & R_{13} \\ R_{21} & \lambda-E_2^0+R_{22} & R_{23} \\ R_{31} & R_{32} & \lambda-E_3^0+R_{33} \end{vmatrix}=0 \tag{42}$$

and so on for higher-order degeneracies.

Again this is equivalent to equations (44) and (47) of Swain (1977) but the results are written in a more concise form here.

Let us now consider briefly the calculation of the eigenvectors. These may be written

$$|x\rangle=\sum_i |i\rangle\langle i|x\rangle\equiv\sum_i x_i|i\rangle. \tag{43}$$

For the non-degenerate case we may obtain the x_i by noting that a solution of the matrix equation $Ax=0$ is

$$x_j=cA^{kj} \tag{44}$$

where c is a constant, and A^{kj} is a cofactor of the matrix A (k indicates any row). Choosing $c=1/A^{kk}$ we may write this solution (cf Swain 1976) in continued fraction form as

$$x_j=\begin{cases} 1 & j=k \\ -\dfrac{a_{jk}}{\mathscr{D}_j^{(k)}}+\sum_l{}^{*}\dfrac{a_{jl}a_{lk}}{\mathscr{D}_{jl}^{(k)}}-\sum_{l,m}{}^{*}\dfrac{a_{jl}a_{lm}a_{mk}}{\mathscr{D}_{jlm}^{(k)}}+\ldots & j\neq k. \end{cases} \tag{45}$$

Making use of equation (31a) this may be written simply as

$$x_j=S_{jk}. \tag{46}$$

This gives us the eigenvector which reduces to $x_j=\delta_{jk}$ when the off-diagonal elements of the matrix are set to zero. Note that with this particular choice of the constant c in equation (44) the wavefunction is *not* normalised; if normalised wavefunctions are required we must use instead of (43) the expression

$$|x\rangle=\sum_i S_{ki}|i\rangle\Big(\sum_i |S_{ki}|^2\Big)^{-1}. \tag{47}$$

In the case of n-fold degeneracy we solve equation (34) with the right-hand side set to zero for the n degenerate components x_α. Again, these are only determined to within an arbitrary constant. When these are known, the remaining x_j may be found by substituting the x_α into equation (27) with all the b_i set to zero. The diagonal elements a_{jj} are of course functions of the perturbed eigenvalue λ and it is assumed that the appropriate value is used for this quantity. The wavefunction must again be normalised.

6. Transition probabilities

The wavefunction of a system described by a time-independent Hamiltonian H is given at time t in terms of its value at time $t = 0$ by the relation

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle = \frac{1}{2\pi i}\oint G(z)\, e^{-izt}|\psi(0)\rangle\, dz \tag{48}$$

where $G(z)$ is the resolvent operator (Goldberger and Watson 1964, Messiah 1962) and the contour in equation (48) lies above the real axis and all the singularities of the integrand. Explicitly

$$G(z) = (z - H)^{-1} \tag{49}$$

and the resolvent satisfies the equation

$$(z - H)G(z) = 1. \tag{50}$$

Operating on equation (50) with $|\psi(0)\rangle$ we obtain

$$(z - H)\{G(z)|\psi(0)\rangle\} = |\psi(0)\rangle \tag{51}$$

which is of the same form as equation (1) with

$$A \equiv z - H, \qquad |x\rangle \equiv G(z)|\psi(0)\rangle, \qquad |b\rangle \equiv |\psi(0)\rangle. \tag{52}$$

Thus, if one wants the probability amplitude for a particular state $|1\rangle$ say, this is found by inverting $\langle 1|G(z)|\psi(0)\rangle \equiv \langle 1|x\rangle \equiv x_1$ according to equation (48) and x_1 may be found from equation (13) under the replacements (52), or from equation (29) or (34) for $n = 1$ under the replacements

$$a_{ij} = (z - E_i^0)\delta_{ij} - V_{ij} \equiv A_{ii}^0\delta_{ij} + A_{ij}^1, \qquad b_j \equiv \langle 1|\psi(0)\rangle. \tag{53}$$

Explicitly,

$$\langle 1|G(z)|\psi(0)\rangle = \frac{b_1}{\mathscr{D}_1} + \sum_{j=2}^{N}{}^{*}\frac{V_{1j}b_j}{\mathscr{D}_{1j}} + \sum_{j,k=2}^{N}{}^{*}\frac{V_{1j}V_{jk}b_k}{\mathscr{D}_{ijk}} + \ldots \tag{54}$$

$$\mathscr{D}_1 = z - E_1^0 - V_{11} - \sum_j{}^{*}\frac{V_{1j}V_{j1}}{\mathscr{D}_j^{(1)}} - \sum_{j,k}{}^{*}\frac{V_{1j}V_{jk}V_{ki}}{\mathscr{D}_{jk}^{(1)}} + \ldots \tag{55}$$

which agrees with results obtained previously (Swain 1975). We have assumed of course that the Hamiltonian is decomposed as

$$H = H^0 + V \tag{56}$$

and that the states $|i\rangle$ are eigenvectors of H^0 as in equation (37).

The time-dependent probability amplitude is found from equation (48) and the transition probability is obtained by taking the modulus squared.

In the case of n-fold degeneracy it is again necessary to treat the n degenerate states on the same footing. With the replacements (53) equation (29) becomes

$$\sum_{\beta=1}^{n}\left((z - E_\alpha^0)\delta_{\alpha\beta} - V_{\alpha\beta} - \sum_j{}^{*}\frac{V_{\alpha j}V_{j\beta}}{\mathscr{D}_j^{(1-n)}} - \sum_{j,k}{}^{*}\frac{V_{\alpha j}V_{jk}V_{k\beta}}{\mathscr{D}_{jk}^{(1\ldots n)}} + \ldots\right)x_\beta$$

$$= b_\alpha + \sum_j{}^{*}\frac{V_{\alpha j}b_j}{\mathscr{D}_j^{(1\ldots n)}} + \sum_{j,k}{}^{*}\frac{V_{\alpha j}V_{jk}b_k}{\mathscr{D}_{jk}^{(1\ldots n)}} + \ldots, \qquad \alpha = 1, 2, \ldots, n. \tag{57}$$

Thus for the case where $|1\rangle$ and $|2\rangle$ are degenerate we obtain the equations

$$\begin{pmatrix} z-E_1^0+R_{11}(z) & R_{12}(z) \\ R_{21}(z) & z-E_2^0+R_{22}(z) \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1-\sum_{i=1}^{N} R_{1i}b_i/\mathscr{D}_i^{(12)} \\ b_2-\sum_{i=1}^{N} R_{2i}b_i/\mathscr{D}_i^{(12)} \end{pmatrix} \tag{58}$$

which give the probability amplitudes for the two degenerate states. The amplitudes for the other states, x_i $(i \neq 1, 2)$, are found by substituting for x_1 and x_2 in equation (27).

The extension to several degenerate states proceeds along exactly similar lines.

Finally, we note that we may recover the standard treatment by regarding the operator equation (50) as being of the form (1) with the correspondences

$$A^0 \equiv z-H^0, \qquad A^1 \equiv -V, \qquad |x\rangle \equiv G(z), \qquad |b\rangle \equiv 1 \equiv \sum_i |i\rangle\langle i| \tag{59}$$

where the states $|i\rangle$ form a complete set. We now take $|x\rangle$ and $|b\rangle$ as operators, not vectors, but this causes no change in the argument up to equation (10), which in the present case becomes

$$(z-H^0+R)PG = S \tag{60}$$

where

$$R = -PVP - PV(z-H^0-QV)^{-1}QVP \tag{61}$$

$$S = P + PV(z-H^0-QV)^{-1}Q. \tag{62}$$

These operators may be expanded in power series in V according to equations (18) and (19):

$$R = -P\sum_{k=0}^{\infty}\left(V\frac{Q}{z-H^0}\right)^k VP \tag{63}$$

$$S = P\sum_{k=0}^{\infty}\left(V\frac{Q}{z-H^0}\right)^k. \tag{64}$$

Equation (61) is just the negative of the usual definition of the level shift operator (e.g. Cohen-Tannoudji 1968) and the standard treatment may be developed as in that reference.

Acknowledgment

One of us (RIJ) thanks the Department of Education for Northern Ireland for financial support.

References

Cohen-Tannoudji C 1968 *Cargèse Lectures in Physics* vol 2, ed M Levy
Feenberg E 1948a *Phys. Rev.* **74** 206–8
—— 1948b *Phys. Rev.* **74** 664–9
Feshbach M 1948 *Phys. Rev.* **74** 1548–9
Goldberger M L and Watson K M 1964 *Collision Theory* (New York: Wiley)
Löwdin P-O 1962 *J. Math. Phys.* **3** 969–82
Messiah A 1962 *Quantum Mechanics* (Amsterdam: North-Holland)
Mower L 1980 *Phys. Rev.* A **22** 882–97
Richards P E 1948 *Phys. Rev.* **74** 835–6
Swain S 1975 *J. Phys. A: Math. Gen.* **8** 1277–97
—— 1976 *J. Phys. A: Math. Gen.* **9** 1811–21
—— 1977 *J. Phys. A: Math. Gen.* **10** 155–65