THE GENERATING EQUATIONS OF RAYLEIGH–SCHRÖDINGER PERTURBATION THEORY

Carlos E. SOLIVEREZ ☆ and Eduardo GAGLIANO ★
Centro Atómico Bariloche ‡ and Instituto Balseiro ‡, 8400 Bariloche, Argentina

Received 19 February 1982; in final form 20 May 1982

A set of non-linear equations is given which solves the stationary Schrödinger equation in terms of a known subproblem. An iterative solution of the equations yields the degenerate version of Rayleigh–Schrödinger perturbation theory, but other approximation schemes, as well as a purely numerical solution, are possible. As an illustration the Stark effect of the planar rigid rotator is discussed.

The quantum-mechanical characterization of a closed system described by a hamiltonian H, whether it be a single atom or an aggregate of many, requires solving the stationary Schrödinger equation in Hilbert space. Only in a few cases do closed solutions exist, and one is forced in general to make simplifying assumptions. The most fruitful one has been to restrict oneself to a manifold of Hilbert space spanned by some reduced set $\{\varphi_j\}$ of basis functions, a good choice being one such that a reasonable approximation to the actual wavefunction is given by a linear combination of a few φ_j. Such is the case of the atomic or molecular shell model where one takes as basis functions some suitable atomic orbitals or adequately symmetrized linear combination of them. The basis set is easily characterized when there is some soluble subproblem h which sufficiently resembles H, in which case one takes $\{\varphi_j\}$ to be some selected set of eigenfunctions of h. As in most cases the φ_j belong to the discrete spectrum, when the remainder $H - h$ is small one may use Rayleigh–Schrödinger perturbation theory in order to find an approximate solution [1]. In what follows it will be shown that a related approach may be used to solve the Schrödinger equation when it is not required that $H - h$ be small, but instead that the eigenfunctions satisfy a certain overlap condition. Perturbation theory is obtained when one solves by iteration the non-linear equations which determine the solution, this being the reason why we have called them the generating equations of Rayleigh–Schrödinger perturbation theory.

The generating equations may be obtained by specializing results from the theory of effective hamiltonians [2–7]. As this theory is usually formulated in the language of abstract vector spaces, projection operators, biorthogonal basis vectors, and the like – a language which is not necessary for solving the simpler problem discussed below – we will instead make an independent discussion based solely on standard linear algebra.

Let us consider the Schrödinger equation

$$H\Phi_{ja} = E_{ja}\Phi_{ja} \tag{1}$$

in the finite dimensional space spanned by some set $\{\varphi_{ja}\}$ of eigenfunctions of h,

$$h\varphi_{ja} = e_j\varphi_{ja}, \quad a = 1, 2, \ldots, d_j, \tag{2}$$

where the second subindex distinguishes the different eigenfunctions belonging to the d_j-degenerate eigenvalue e_j. The first subindices have been chosen in such a way that

$$e_j \neq e_k \quad \text{if} \quad j \neq k, \tag{3}$$

$$E_{ja} = e_j + \Delta_{ja}, \quad \lim_{V \to 0} \Delta_{ja} = 0, \tag{4}$$

☆ Member of the Carrera del Investigador Cientifico, Consejo Nacional de Investigaciones Cientificas y Técnicas.
★ Instituto Balseiro student.
‡ Comisión Nacional de Energia Atómica.
‡ Comisión Nacional de Energia Atómica and Universidad Nacional de Cuyo.

0 009-2614/82/0000--0000/$ 02.75

and no restrictions are imposed on the degeneracies of E_{ja}. The remainder V is defined by

$$V = H - h, \tag{5}$$

and all operators are taken to be hermitian as corresponds to a closed system. Therefore all wavefunctions may be orthonormalized

$$\langle \varphi_{ja} | \varphi_{kb} \rangle = \delta_{jk}\delta_{ab} = \langle \Phi_{ja} | \Phi_{kb} \rangle. \tag{6}$$

In the chosen space one may always write

$$\Phi_{ja} = \sum_{k,b} \varphi_{kb}\, U_{kb,ja}, \tag{7}$$

where from (6) it follows that the matrix $\mathbf{U}$ of elements $U_{kb,ja}$ is unitary. Using eqs. (4) and (5), eq. (1) may be written

$$(h - e_j)\,\Phi_{ja} = (\Delta_{ja} - V)\,\Phi_{ja}. \tag{8}$$

Replacing Φ_{ja} from eq. (7), using eq. (2), taking scalar product with φ_{lc}, and making some changes of indices, eq. (8) becomes

$$(e_k - e_j)\,U_{kb,ja} = U_{kb,ja}\Delta_{ja} - \sum_{l,c} V_{kb,lc}\, U_{lc,ja}, \tag{9}$$

where $V_{kb,lc}$ is the matrix element $\langle \varphi_{kb} | V | \varphi_{lc} \rangle$. From eq. (3), the first member of eq. (9) vanishes only when $j = k$, in which case one obtains

$$U_{jb,ja}\Delta_{ja} = \sum_{l,c} V_{jb,lc} U_{lc,ja}. \tag{10}$$

It is now convenient to define the $d_j \times d_k$ rectangular matrix $\mathbf{A}_{jk}$, and the $d_j \times d_j$ square matrix $\mathbf{\Delta}_j$ such that its ab matrix elements are given by

$$(\mathbf{A}_{jk})_{ab} = A_{ja,kb}\,, \quad (\mathbf{\Delta}_j)_{ab} = \Delta_{ja}\delta_{ab}\,. \tag{11}$$

Eqs. (9) and (10) may then be more concisely written in terms of matrix products as

$$(e_k - e_j)\mathbf{U}_{kj} = \mathbf{U}_{kj}\mathbf{\Delta}_j - \sum_l \mathbf{V}_{kl}\mathbf{U}_{lj}\,, \tag{12}$$

$$\mathbf{U}_{jj}\,\mathbf{\Delta}_j = \sum_l \mathbf{V}_{jl}\mathbf{U}_{lj}\,. \tag{13}$$

When keeping j fixed and making k run over all possible values, one obtains a set of coupled non-linear equations for the matrices $\mathbf{U}_{1j}, \mathbf{U}_{2j}, \ldots, \mathbf{U}_{jj}, \ldots$, which when solved gives all the coefficients in eq. (7) as well as the energy shifts Δ_{jn}. Indeed, when $\mathbf{U}_{jj}$ is non-singular it follows that

$$\mathbf{\Delta}_j = \mathbf{U}_{jj}^{-1} \sum_l \mathbf{V}_{jl}\mathbf{U}_{lj}\,, \tag{14}$$

which may be used in eq. (12) to decouple the two sets of equations. Upon multiplication of the resulting expression by the right with $\mathbf{U}_{jj}^{-1}$ one gets

$$\mathbf{R}_{kj} = (e_k - e_j)^{-1} \sum_l (\mathbf{R}_{kj}\mathbf{V}_{jl}\mathbf{R}_{lj} - \mathbf{V}_{kl}\mathbf{R}_{lj})\,, \quad j \neq k\,, \tag{15}$$

where we have defined

$$\mathbf{R}_{kj} = \mathbf{U}_{kj}\mathbf{U}_{jj}^{-1}\,. \tag{16}$$

The non-linear equations (15) completely determine $\mathbf{R}_{kj}$ and are the generating equations mentioned at the beginning. Being non-linear, these equations do have in general more than one solution in which case the "continuity" condition eq. (4), or

$$\lim_{V\to 0} \mathbf{R}_{kj} = \mathbf{1}_j \delta_{kj}\,, \tag{17}$$

where $\mathbf{1}_j$ is the unit $d_j \times d_j$ square matrix, unambiguously identifies the proper solution. Eq. (15) was first given in operator form by Bloch [2] though in a somewhat different context. In the present approach it was first discussed by one of the authors [3] in the particular case where

$$\mathbf{U}_{jj} = \mathbf{1}_j\,. \tag{18}$$

This condition is in general not fulfilled, and it is therefore necessary to discuss how to determine $\mathbf{U}_{jj}$. From eq. (14) it immediately follows

$$\mathbf{\Delta}_j = \mathbf{U}_{jj}^{-1}\mathbf{W}_{jj}\mathbf{U}_{jj}\,, \tag{19}$$

where we have defined

$$\mathbf{W}_{jj} = \sum_k \mathbf{V}_{jk}\mathbf{R}_{kj}\,. \tag{20}$$

As long as all $\mathbf{R}_{kj}$ with fixed j are known, $\mathbf{W}_{jj}$ is also known, and $\mathbf{U}_{jj}$, being the matrix which diagonalizes $\mathbf{W}_{jj}$, is obtained together with the Δ_{ja} from the eigenvectors and eigenvalues of $\mathbf{W}_{jj}$. It should be pointed out that as $\mathbf{W}_{jj}$ is in general non-hermitian, its eigenvectors are not orthogonal, and therefore $\mathbf{U}_{jj}$ is not unitary. As eq. (19) remains invariant when $\mathbf{U}_{jj}$ is multiplied by an arbitrary complex number, this indetermination must be removed. This may be done by

recalling the orthonormality condition eq. (6) which in terms of $\mathbf{U}_{jj}$ and $\mathbf{R}_{kj}$ reads

$$\langle \Phi_{ja} | \Phi_{jb} \rangle = \sum_k (\mathbf{U}_{jj}^\dagger \mathbf{R}_{kj}^\dagger \mathbf{R}_{kj} \mathbf{U}_{jj})_{ab} = \delta_{ab} , \tag{21}$$

that is

$$\mathbf{U}_{jj} \mathbf{U}_{jj}^\dagger = \left[\sum_k \mathbf{R}_{kj}^\dagger \mathbf{R}_{kj} \right]^{-1} , \tag{22}$$

which together with eq. (19) completely determines $\mathbf{U}_{jj}$ apart from irrelevant phase factors.

An essential hypothesis was the non-singular character of $\mathbf{U}_{jj}$, which corresponds to saying that the functions

$$\Phi'_{ja} = \sum_b \varphi_{jb} U_{jb,ja} , \quad a = 1, 2, ..., d_j , \tag{23}$$

are linearly independent. As Φ'_{ja} is the projection of Φ_{ja} over the subspace spanned by the eigenfunctions of h with eigenvalue e_j, the assumption imposes a restriction on how different H and h may be. It is known that in some cases, notably in superconductivity, the condition is not fulfilled, but this is the exception rather than the rule. No prescription is known at present concerning the ways of partitioning H, eq. (5), that fulfill the linear independence condition, so one is forced here to take a pragmatic approach and assume the condition to be true unless the final results prove the contrary.

It is interesting to point out that one can give more convenient properties to the matrix $\mathbf{W}_{jj}$ by making a change of basis to some linear combination over the functions φ_{ja} with j fixed. When discussing the properties of such a choice of new basis functions one is in the realm of the theory of effective hamiltonians.

Only in a few cases are explicit solutions to eqs. (15) known [4], and although the equations may always be solved numerically, one would most often resort to some perturbative scheme. When solving the system by the iteration method one obtains degenerate Rayleigh–Schrödinger perturbation theory [1], which we now discuss.

Both h and V have characteristic scale factors S_h and S_V whose quotient gives the perturbation parameter

$$\lambda = S_V / S_h . \tag{24}$$

It is therefore convenient to write

$$e_j = S_h \epsilon_j , \quad \mathbf{V}_{lm} = S_V \mathbf{v}_{lm} , \quad \mathbf{W}_{jj} = S_h \mathbf{w}_{jj} , \tag{25}$$

$$\mathbf{R}_{kj} = \delta_{kj} \mathbf{1}_j + \lambda \alpha_{kj} \left[\mathbf{R}_{kj} \sum_l \mathbf{v}_{jl} \mathbf{R}_{lj} - \sum_l \mathbf{v}_{kl} \mathbf{R}_{lj} \right] , \tag{26}$$

where

$$\alpha_{kj} = 0 , \quad \text{if } k = j ,$$

$$= (\epsilon_k - \epsilon_j)^{-1} , \quad \text{if } k \neq j . \tag{27}$$

Defining the pth order approximation $\mathbf{R}_{lj}^{(p)}$ by

$$\mathbf{R}_{lj} = \sum_{p=0}^{\infty} \lambda^p \mathbf{R}_{lj}^{(p)} , \tag{28}$$

from eq. (26) it is easily proved that $\mathbf{R}_{kj}^{(p)}$ satisfies the recurrence relations

$$\mathbf{R}_{kj}^{(0)} = \delta_{kj} \mathbf{1}_j ,$$

$$\mathbf{R}_{kj}^{(p)} = \alpha_{kj} \left[- \sum_l \mathbf{v}_{kl} \mathbf{R}_{lj}^{(p-1)} + \sum_{q=1}^{p-1} \mathbf{R}_{kj}^{(q)} \sum_l \mathbf{v}_{jl} \mathbf{R}_{lj}^{(p-1-q)} \right] , \quad p \geqslant 1 . \tag{29}$$

In the same fashion from eq. (20) it is obtained that

$$\mathbf{w}_{jj} = \sum_{p=1}^{\infty} \lambda^p \mathbf{w}_{jj}^{(p)} , \quad \mathbf{w}_{jj}^{(p)} = \sum_k \mathbf{v}_{jk} \mathbf{R}_{kj}^{(p-1)} . \tag{30}$$

Up to second-order terms in $\mathbf{R}_{kj}$ one gets

$$\mathbf{w}_{jj}^{(1)} = \mathbf{v}_{jj} , \quad \mathbf{R}_{kj}^{(1)} = -\alpha_{kj} \mathbf{v}_{kj} ,$$

$$\mathbf{w}_{jj}^{(2)} = - \sum_k \mathbf{v}_{jk} \alpha_{kj} \mathbf{v}_{kj} ,$$

$$\mathbf{R}_{kj}^{(2)} = \alpha_{kj} \sum_l \mathbf{v}_{kl} \alpha_{lj} \mathbf{v}_{lj} - \alpha_{kj}^2 \mathbf{v}_{kj} \mathbf{v}_{jj} ,$$

$$\mathbf{w}_{jj}^{(3)} = \sum_{k,l} \mathbf{v}_{jk} \alpha_{kj} \mathbf{v}_{kl} \alpha_{lj} \mathbf{v}_{lj} - \sum_k \mathbf{v}_{jk} \alpha_{kj}^2 \mathbf{v}_{kj} \mathbf{v}_{jj} , \tag{31}$$

where it is seen that the first non-hermitian terms in $\mathbf{W}_{jj}$ appear in third order of perturbation theory.

One may also solve the generating equations using modified iteration methods [8] or other algorithms as the one used at the end.

In order to illustrate the formalism we will discuss its application to the case of the Stark effect of a rigid planar rotator. The system is particularly interesting because: (a) In the absence of applied electric field $\mathbf{\varepsilon}$

all states except the ground one are doubly degenerate; (b) the degeneracy of the mth excited state is removed only in the $2m$th order of perturbation theory; (c) the exact solutions are known.

Let I be the rotator's moment of inertia, $\boldsymbol{L}$ its orbital angular momentum, and $\boldsymbol{p}$ its dipole moment which lies on the plane normal to $\boldsymbol{L}$. As only the component of $\boldsymbol{\epsilon}$ normal to $\boldsymbol{L}$ contributes to the energy, one may take without loss of generality the x axis along $\boldsymbol{\epsilon}$ and the z axis along $\boldsymbol{L}$. Then the total hamiltonian may be written

$$H = L_z^2/2I - \boldsymbol{p}\cdot\boldsymbol{\epsilon} = -(\hbar^2/2I)\,\mathrm{d}^2/\mathrm{d}\phi^2 - p\epsilon\cos\phi, \quad (32)$$

where ϕ is the azimuthal angle. The Schrödinger equation is here exactly soluble [9,10] as eq. (1) leads to

$$(\mathrm{d}^2/\mathrm{d}u^2 + a - 2q\cos 2u)f(u) = 0\,, \quad (33)$$

$$a = 8IE/\hbar^2\,, \quad q = -4Ip\epsilon/\hbar^2\,, \quad \Phi(\phi) = f(\phi/2)\,, \quad (34)$$

which is the canonical form of Mathieu's equation [11]. We now take

$$h = -S_h(L_z/\hbar)^2\,, \quad V = S_V 2\cos\phi\,,$$

$$S_h = \hbar^2/2I\,, \quad S_V = -p\epsilon/2\,. \quad (35)$$

From eqs. (2), (6), (35), (24) and (25) it is found that

$$\epsilon_0 = 0\,, \quad \epsilon_j = j^2\,, \quad \varphi_0 = (2\pi)^{-1/2}\,,$$

$$\varphi_{j+} = \pi^{-1/2}\cos j\phi\,, \quad \varphi_{j-} = \pi^{-1/2}\sin j\phi\,,$$

$$\lambda = -Ip\epsilon/\hbar^2 = q/4\,, \quad j = 1, 2, 3, \ldots\,, \quad (36)$$

$$\mathbf{v}_{00} = 0\,, \quad \mathbf{v}_{0l} = (2^{1/2}\ 0)\,\delta_{l1}\,, \quad \mathbf{v}_{l0} = \mathbf{v}_{0l}^{\dagger}\,,$$

$$\mathbf{v}_{lm} = (\delta_{lm-1} + \delta_{lm+1})\mathbf{1}\,, \quad l, m = 1, 2, \ldots\,, \quad (37)$$

where $(2^{1/2}\ 0)$ is a row matrix and $\mathbf{1}$ is the unit 2×2 matrix. The degeneracy indices $a, b = +, -$ correspond to the sign of the eigenvalues of the reflection operator σ_y such that

$$\sigma_y g(\phi) = g(-\phi), \quad \sigma_y\varphi_0 = \varphi_0\,, \quad \sigma_y\varphi_{j\pm} = \pm\varphi_{j\pm}\,. \quad (38)$$

As σ_y commutes with H the total wavefunction Φ may be taken to be a simultaneous eigenfunction of both operators. One may therefore set equal to zero all non-diagonal elements in $\mathbf{R}_{kj}$, $\mathbf{U}_{kj}$ and $\mathbf{W}_{jj}$.

Using eqs. (31) we obtain

$$\Delta_0/S_h = w_{00} = -2\lambda^2 + \tfrac{7}{2}\lambda^4 + \ldots\,,$$

$$\Delta_{1+}/S_h = w_{1+,1+} = \tfrac{5}{3}\lambda^2 - (763/216)\lambda^4 + \ldots\,,$$

$$\Delta_{1-}/S_h = w_{1-,1-} = -\tfrac{1}{3}\lambda^2 + (5/216)\lambda^4 + \ldots\,,$$

$$\Delta_{2+}/S_h = w_{2+,2+} = (2/15)\lambda^2 + (433/13500)\lambda^4 + \ldots\,,$$

$$\Delta_{2-}/S_h = w_{2-,2-} = (2/15)\lambda^2 - (317/13500)\lambda^4 + \ldots\,,$$

$$\Delta_{j\pm}/S_h = w_{j\pm,j\pm} = [2/(4j^2-1)]\lambda^2 + [(20j^2+7)/2(4j^2-1)^3(j^2-1)]\lambda^4 + \ldots\,,$$

$$j > 2\,. \quad (39)$$

Because the perturbation v connects only adjacent levels, all odd-order terms in λ ought to vanish. The results eqs. (39) coincide with those obtained by McLachlan [11] in a completely different fashion. It may be seen that the lowest order in which the degeneracy of the Φ_{j+} and Φ_{j-} states is removed is that in which the matrix product $\mathbf{v}_{10}\mathbf{v}_{01}$ appears for the first time in the expansion (30). This happens in the $2j$th order of perturbation theory, and the resulting splittings turns out to be

$$(\Delta_{1+} - \Delta_{1-})/S_h = 2\lambda^2\,,$$

$$(\Delta_{j+} - \Delta_{j-})/S_h = (2/j^2)(\alpha_{jj-1}\alpha_{jj-2}\ldots\alpha_{j2}\alpha_{j1})^2\lambda^{2j}\,,$$

$$j \geqslant 1\,, \quad (40)$$

an expression which has not been given before and seems difficult to obtain in any other way. One may also write without difficulty the coefficients $R_{k\pm,j\pm}$ and from them $U_{k\pm,j\pm}$. For this purpose one may either normalize exactly the $R_{k\pm,j\pm}$ coefficients by using eq. (6) or, being more consistent with the perturbation scheme, normalize them to the desired order by an expansion of eqs. (22) and (16) in a manner similar to that of eqs. (28)–(30).

Because the coefficients $R_{k\pm,j\pm}$ decrease very rapidly when k increases, a simple approximation is to solve eqs. (15) assuming that they vanish from a certain value of k onwards. As an illustration we will approximate the ground-state energy and coefficients taking $R_{k+,0} = 0$ for $k \geqslant 2$. Remembering that by definition $R_{00} = 1$, and making use of eqs. (15), (20), (22) and (16), it is obtained

$$R_{1+,0} = [1 \pm (1+8\lambda^2)^{1/2}]/2^{3/2}\lambda ,$$

$$\Delta_0/S_h = 2^{1/2}\lambda R_{1+,0} ,$$

$$U_{00} = (1+R_{1+,0}^2)^{-1/2} , \quad U_{1+,0} = R_{1+,0}U_{00} , \quad (41)$$

where according to the limiting condition (17) one should take in $R_{1+,0}$ the lower sign in the square root. For $\lambda = 2$ ($q = 8$), a value which is well outside the range of perturbation theory, one obtains

$$\Delta_0/S_h = -2.4 , \quad U_{00} = 0.77 , \quad U_{1+,0} = -0.64 , \quad (42)$$

to be compared with the rounded exact values [11, pp. 31,32]

$$\Delta_0/S_h = -2.7 , \quad U_{00} = 0.71 , \quad U_{1+,0} = -0.67 , \quad (43)$$

a remarkable agreement for such a crude approximation. The precision increases very rapidly if one allows higher coefficients to be non-vanishing. Notice that from eq. (39) one would obtain $\Delta_0/S_h = 48$.

In conclusion, it has been shown that Rayleigh–Schrödinger perturbation theory is just a particular way (iteration method) of solving the generating equations, but by no means the only one.

References

[1] A. Dalgarno, in: Quantum theory, Vol. 1, ed. D.R. Bates (Academic Press, New York, 1961) pp. 171–209.
[2] C. Bloch, Nucl. Phys. 6 (1958) 329.
[3] C.E. Soliverez, Am. J. Phys. 35 (1967) 624.
[4] C.E. Soliverez, J. Phys. C2 (1969) 2161.
[5] F. Jorgensen, Mol. Phys. 29 (1975) 1137.
[6] C.E. Soliverez, Phys. Letters 79A (1980) 410.
[7] C.E. Soliverez, Phys. Rev. A24 (1981) 4.
[8] E. Durand, Solutions numériques des équations algébriques, Vol. 2 (Masson, Paris, 1961) ch. 5.
[9] M.P. Silverman, Phys. Rev. A24 (1981) 339.
[10] M.P. Silverman, Am. J. Phys. 49 (1981) 871.
[11] N.W. McLachlan, Theory and application of Mathieu functions (Clarendon Press, Oxford, 1951).