A variational-iterative technique applied to quantum mechanical problems

B L Burrows and P W Core

Department of Mathematics, North Staffordshire Polytechnic, Beaconside, Stafford, UK

Abstract. A technique for solving nonlinear operator equations iteratively is applied to quantum mechanical calculations. It is shown that alternative methods of applying the technique lead to another way of approaching variational calculations and generalisations of perturbation theory respectively.

1. Introduction

On the subject of quantum mechanical calculations two methods are predominant in the literature, these being the variational method and perturbation theory. In this paper we consider two related methods of obtaining approximate solutions of an equation of the form

$$T\phi = f(\phi) \tag{1.1}$$

where T is a self-adjoint operator and $f(\phi)$ may be nonlinear. One of these methods uses the theory of solving nonlinear equations and can be shown to be equivalent to the usual variational methods. The second method can be described as a generalisation of perturbation theory and includes second-order perturbation theory as a special case. In all cases we suppose that we can divide our quantum mechanical Hamiltonian H into $H_0 + V$ where

$$H_0\phi_0 = E_0\phi_0 \tag{1.2}$$

is a solved problem and that ϕ_0 is an acceptable initial approximation to the wavefunction. The use of the subscript '0' does not imply that these methods are restricted to the ground state. The methods are illustrated with example calculations for

$$H = -\tfrac{1}{2}\,d^2/dx^2 + \tfrac{1}{2}x^2 + Ax^4 \tag{1.3}$$

so that we can take

$$H_0 = -\tfrac{1}{2}\,d^2/dx^2 + \tfrac{1}{2}x^2 \tag{1.4}$$

the harmonic oscillator Hamiltonian.

2. Basic theory

The problems considered can be written in the form

$$T\psi = f(\psi) \tag{2.1}$$

where $\psi \in D(T)$ the domain of a self-adjoint operator T on a Hilbert space and T^{-1} exists. In general $f(\psi)$ may be nonlinear and (2.1) may have several solutions. The theory involved has been described in the references. The method of solution employed by Burrows and Perks (1981a, b) was to consider the related sequence of equations

$$T\Psi_{n+1} = f(\psi_n) \tag{2.2}$$

where ψ_{n+1} is a variational approximation to Ψ_{n+1}. To be more precise in this paper we will assume that ψ_n belongs to a set of parameterised functions so that

$$\psi_n \in S = \{\theta \in D(T): \theta(x) = \Phi(x, a_1, a_2 \dots a_N), a_i \in \mathbb{R}\}. \tag{2.3}$$

Let

$$G(\hat{\psi}_{n+1}) = \langle \hat{\psi}_{n+1} | T\hat{\psi}_{n+1} \rangle - 2\langle \hat{\psi}_{n+1} | f(\psi_n) \rangle \tag{2.4}$$

then $\delta G = 0$ implies that for fixed ψ_n and arbitrary $\delta \hat{\psi}_{n+1}$

$$\langle \delta \hat{\psi}_{n+1} | T\hat{\psi}_{n+1} - f(\psi_n) \rangle = 0. \tag{2.5}$$

This is satisfied when $\hat{\psi}_{n+1} = \Psi_{n+1}$. ($T^{-1}$ exists so that for fixed ψ_n this solution is unique.) A variational solution ψ_{n+1} for Ψ_{n+1} is found by solving

$$\langle \delta \hat{\psi}_{n+1} | T\hat{\psi}_{n+1} - f(\psi_n) \rangle = 0 \tag{2.6}$$

for fixed ψ_n and arbitrary $\delta \hat{\psi}_{n+1} \in S$. Successive iterations can be obtained in this way and if the procedure converges so that $\lim_{n \to \infty} \psi_n = \Psi$ and if $\psi \in S$ then $\Psi = \psi$ where ψ is some solution of (2.1). If $\psi \notin S$ then we can enlarge S and repeat the procedure. An alternative method, rather than consider the successive approximations, is to solve the nonlinear equation

$$\langle \delta \hat{\psi} | T\hat{\psi} - f(\hat{\psi}) \rangle = 0 \qquad \hat{\psi}, \delta \hat{\psi} \in S \tag{2.7}$$

to obtain an approximation for ψ in S. Standard methods can be used to solve the nonlinear equations. However, both methods rely on a good initial approximation ψ_0. In this paper the set S will consist of functions of the form

$$\Phi(x, \boldsymbol{a}) = e_0(x) + \sum_{j=1}^{N} a_j e_j(x)$$

where

$$\langle e_i | Te_j \rangle = \alpha_i \delta_{ij} \qquad (\alpha_i > 0). \tag{2.8}$$

In this notation (2.7) becomes

$$\langle e_i | T\hat{\psi} - f(\hat{\psi}) \rangle = 0 \qquad i = 1, \dots, N \tag{2.9}$$

or more conveniently

$$a_i = (1/\alpha_i)\langle e_i | f(\hat{\psi}) - Te_0 \rangle \tag{2.10}$$

i.e.

$$\boldsymbol{a} = F_i(\boldsymbol{a}). \tag{2.11}$$

The standard techniques (see Ortega and Rheinboldt 1970) for speeding up the convergence of the iterative procedure corresponding to (2.11) are of the form

$$\boldsymbol{a}^{n+1} = \boldsymbol{a}^n - \beta(\boldsymbol{a}^n - F(\boldsymbol{a}^n)) \tag{2.12}$$

where β is a linear operator. If $\beta = I$ (the identity operator) then this corresponds to the direct iterative procedure.

Choosing

$$\beta = I - F^1(\boldsymbol{a}^n) \tag{2.13}$$

gives Newton's method, but a simpler choice is

$$\beta = I - F^1(\boldsymbol{a}^0). \tag{2.14}$$

This is identical with Newton's method for the first iteration but saves on the computation time involved in the recalculation $F^1(\boldsymbol{a}^n)$ at each stage for Newton's method. In the examples considered in this paper $F(\boldsymbol{a})$ was simple enough for Newton's method to be applied throughout.

3. Variational approximations

Consider a quantum mechanical system described by the Hamiltonian H so that

$$H\psi = (H_0 + V)\psi = E\psi. \tag{3.1}$$

We may rewrite this in the form

$$H_0\psi = (\langle\psi|H\psi\rangle/\langle\psi|\psi\rangle)\psi - V\psi = f(\psi). \tag{3.2}$$

This is now of the form of (2.1) and the theory described in § 2 can be applied.

We suppose that

$$H_0\phi_k = E_k\phi_k \tag{3.3}$$

and that $\phi_k \approx \psi$ so that we may take $\psi_0 = \phi_k$. The set of functions used will be, corresponding to (2.8),

$$\phi_k + \sum_{\substack{j=0 \\ j \neq k}}^{m} a_j\phi_j. \tag{3.4}$$

Here we suppose

$$H_0\phi_j = E_j\phi_j \tag{3.5}$$

$$\langle\phi_j|\phi_i\rangle = \delta_{ji} \tag{3.6}$$

and that

$$E_j \leqslant E_{j+1}. \tag{3.7}$$

The solution of (2.7) in this case can be shown to be equivalent to the variation of

$$\langle\hat{\psi}|H\hat{\psi}\rangle/\langle\hat{\psi}|\hat{\psi}\rangle \tag{3.8}$$

with $\hat{\psi} \in S$. Thus the method provides an alternative way of solving this problem. The method can also be applied using a simpler function $f(\psi)$. Consider the equation

$$(H_0 - E_k)\psi = (\hat{E} - V)\psi = f(\psi). \tag{3.9}$$

For consistency we require

$$\hat{E} = \langle\phi_k|V\psi\rangle/\langle\phi_k|\psi\rangle. \tag{3.10}$$

We can now use the procedures described in § 2 with the set of functions defined by (3.4). For (3.9) to be exactly in the form of (2.1) we need to relax the condition that T^{-1} exists to T^{-1} exists on the space orthogonal to ϕ_k. With $\hat{\psi}$ defined by (3.4), (2.7) becomes

$$\langle \phi_j | (H_0 - E_k)\hat{\psi} - (\hat{E} - V)\hat{\psi} \rangle = 0$$

where

$$\hat{E} = \langle \phi_k | V\hat{\psi} \rangle. \tag{3.11}$$

In this case $f(\psi)$ is a much simpler function and again the solutions obtained are equivalent to the variation of (3.8). The major advantage in doing the Rayleigh–Ritz calculation this way is that roundoff errors only occur at the final iteration whereas the usual eigenvalue techniques can be subject to a cumulation of errors. It is also easily applied to an excited state provided a good initial approximation is available.

4. Generalisation of perturbation theory

Now consider the equation

$$(H_0 - E_0)\Psi_{n+1} = (E - V)\psi_n \tag{4.1}$$

where

$$E = \langle \phi_0 | V\psi_n \rangle \tag{4.2}$$

and

$$(H_0 - E_0)\phi_0 = 0. \tag{4.3}$$

This is of the form (2.2) and we will seek to use the method of successive iterations solving (2.6) to obtain the sequence $\{\psi_n\}$. The set S will consist of functions of the form of (3.4) with $k = 0$. The sequence $\{\psi_n\}$ can be constructed in many ways. One method is to take $\psi_0 = \phi_0$ and

$$\psi_n = \phi_0 + \sum_{j=1}^{n} a_{n_j}\phi_j \qquad (n \geq 1). \tag{4.4}$$

As in the previous section we are assuming the relationships (3.5), (3.6) and (3.7). The solution of (2.6) leads to

$$a_{n+1,j} = \langle \phi_j | (E - V)\psi_n \rangle / (E_j - E_0), \qquad j = 1 \ldots (n+1)$$

$$= \left[a_{nj}(1 - \delta_{jn+1})E - \sum_{k=1}^{n} a_{nk}\langle \phi_j | V\phi_k \rangle \right] \Big/ (E_j - E_0) \tag{4.5}$$

where

$$E = \langle \phi_0 | V\phi_0 \rangle + \sum_{k=1}^{n} a_{nk}\langle \phi_0 | V\phi_k \rangle. \tag{4.6}$$

The coefficients calculated using (4.5) differ from those obtained using conventional variational perturbation theory in that $E^1 = \langle \phi_0 | V\phi_0 \rangle$ is replaced by E defined in (4.6). Consequently the method differs from conventional perturbation theory when $n \geq 2$ in the sequence. We note that no iterations are necessary for the calculation of the

a_{nj} and the work involved is similar to that of perturbation theory involving sums of the matrix elements $\langle \phi_j | V\phi_k \rangle$. At the nth stage of the calculation, when ψ_{n+1} is found, the coefficients of ϕ_j ($j \leqslant n$) are re-estimated. Thus the differences $a_{nj} - a_{n+1j}$ ($j \leqslant n$) provide a practical measure of convergence. This procedure can also be applied to excited states. A problem with such an application is that it is necessary to define an ordering of the ϕ_j. For the ground state we have used $\phi_i > \phi_j$ if $E_i > E_j$ and an arbitrary ordering for $E_i = E_j$. A possible ordering for the excited state ϕ_k is

$$\phi_i > \phi_j \qquad \text{if} \qquad |E_i - E_k| > |E_j - E_k| \tag{4.7}$$

with again arbitrary ordering for equality.

An alternative choice of the sequence $\{\psi_n\}$ is $\psi_0 = \phi_0$,

$$\psi_n = \phi_0 + \sum_{j=1}^{m} a_{nj}\phi_j \qquad (n \geqslant 1) \tag{4.8}$$

for some fixed m. This leads to

$$a_{n+1,j} = \langle \phi_j | (E - V)\psi_n \rangle / (E_j - E_0), \qquad j = 1 \ldots m \tag{4.9}$$

where

$$E = \langle \phi_0 | V\phi_0 \rangle + \sum_{k=1}^{m} \langle \phi_0 | V\phi_k \rangle a_{nk}. \tag{4.10}$$

It is easy to verify that at the first stage of the calculation

$$E = \langle \phi_0 | V\phi_0 \rangle = E^1 \tag{4.11}$$

the first-order energy in (truncated) perturbation theory and that the first iteration produces the first-order wavefunction so that at the second stage $E = E^1 + E^2$ the sum of the first- and second-order perturbation energies. This pattern does not continue but proceeding iteratively we obtain successive approximations to the energy as in higher-order perturbation theory. This form of sequence $\{\psi_n\}$ can be applied directly for excited states with ϕ_k replacing ϕ_0 and (4.8) becoming

$$\psi_n = \phi_k + \sum_{\substack{j=0 \\ j \neq k}}^{m} a_{nj}\phi_j. \tag{4.12}$$

5. Scaling for generalised perturbation theory

As in conventional perturbation theory the convergence of the method can be improved by a scaling of H_0 which is equivalent to a redivision of H into H_0 and V. We have

$$H = H_0 + V = \mu H_0 + [V + (1 - \mu)H_0] \tag{5.1}$$

$$= \hat{H}_0 + \hat{V} \qquad \text{(say)} \tag{5.2}$$

for any real μ. We now wish to choose μ so that $\hat{V}$ is as small as possible. The eigenfunctions of $\hat{H}_0$ are ϕ_k, the eigenfunction of H_0, but the eigenvalues are now μE_k. One choice of μ to produce $\hat{V}$ small in some sense is that μ such that

$$S = \langle \phi_0 | (V + (1 - \mu)H_0)^2 \phi_0 \rangle \tag{5.3}$$

Table 1. Ground state.

V	x^4	$\frac{1}{4}x^4$	$\frac{1}{10}x^4$	$\frac{7}{100}x^4$	$\frac{1}{20}x^4$	$\frac{1}{100}x^4$
First variational solution (3.2)	0.805 870 (5)	0.621 292 (4)	0.559 166 (4)	0.543 793 (4)	0.532 643 (3)	0.507 256 (3)
Second variational solution (3.9)	0.805 870 (7)	0.621 292 (6)	0.559 166 (5)	0.543 793 (4)	0.532 643 (3)	0.507 256 (3)
Standard perturbation theory	1.25	0.6875	0.575	0.5525	0.5375	0.5075
	−1.375	0.523 44	0.565 03	0.546 81	0.533 55	0.507 26
	No convergence	No convergence	No convergence	No convergence	0.5326	0.507 256
Standard PT adjusted by μ	1.25	0.6875	0.575	0.5525	0.5375	0.5075
	0.2000	0.568 18	0.552 17	0.540 86	0.5314	0.507 24
	No convergence	No convergence	No convergence	0.54	0.533	0.507 256
Approximation based on (4.1) and (4.8)	1.25	0.6875	0.575	0.5525	0.5375	0.5075
	−1.375	0.523 44	0.565 03	0.546 81	0.533 55	0.507 26
	No convergence	No convergence	No convergence	No convergence	0.5326	0.507 256
Approximation based on (4.1) and (4.8) with μ	1.25	0.6875	0.575	0.5525	0.5375	0.5075
	0.2000	0.568 12	0.552 17	0.540 81	0.5314	0.507 24
	No convergence	No convergence	No convergence	0.544	0.532 64	0.507 256
Approximation based on (4.1) and (4.4)	No convergence	No convergence	0.56	0.54	0.533	0.507 25
Approximation based on (4.1) and (4.4) with μ	No convergence	0.6	0.56	0.544	0.533	0.507 26

Table 2. First excited state.

V	x^4	$\frac{1}{4}x^4$	$\frac{1}{10}x^4$	$\frac{7}{100}x^4$	$\frac{1}{20}x^4$	$\frac{1}{100}x^4$
First variational solution (3.2)	2.843 872 (5)	2.027 398 (5)	1.770 324 (4)	1.703 398 (4)	1.653 703 (4)	1.535 651 (3)
Second variational solution (3.9)	2.843 872 (9)	2.027 398 (7)	1.770 324 (6)	1.703 398 (5)	1.653 703 (4)	1.535 651 (3)
Standard perturbation theory	5.25 −15.375 No convergence	2.4375 1.1484 No convergence	1.875 1.668 75 No convergence	1.7625 1.6614 1.7	1.6875 1.6359 1.7	1.5375 1.5354 1.535 651
Standard PT adjusted by μ	5.25 −6.4286 No convergence	2.4375 1.6442 No convergence	1.875 1.7100 2	1.7625 1.6765 1.7	1.6875 1.6417 1.654	1.5375 1.5355 1.535 651
Approximation based on (4.1) and (4.8)	5.25 −15.375 No convergence	2.4375 1.1484 No convergence	1.875 1.668 75 No convergence	1.7625 1.6614 No convergence	1.6875 1.6359 1.654	1.5375 1.5354 1.535 651
Approximation based on (4.1) and (4.8) with μ	5.25 −6.4286 No convergence	2.4375 1.6442 No convergence	1.875 1.7100 No convergence	1.7625 1.6755 1.705	1.6875 1.6417 1.653 71	1.5375 1.5355 1.535 651
Approximation based on (4.1) and (4.4)	No convergence	No convergence	2	1.7	1.7	1.536
Approximation based on (4.1) and (4.4) with μ	No convergence	No convergence	1.8	1.7	1.65	1.536

is a minimum.

$$S = \langle \phi_0 | V^2 \phi_0 \rangle + 2(1-\mu)E_0 E^1 + (1-\mu)^2 (E_0)^2 \qquad (5.4)$$

$$\partial S/\partial \mu = 0 \Rightarrow \mu = (E_0 + E^1)/E_0 \qquad (5.5)$$

where $E^1 = \langle \phi_0 | V \phi_0 \rangle$. The new first-order energy is $\langle \phi_0 | \hat{V} \phi_0 \rangle = 0$. An alternative choice of μ is to minimise

$$\sum_{k=0}^{m} \langle \phi_k | (V + (1-\mu)H_0)^2 \phi_k \rangle \qquad (5.6)$$

which leads to

$$\mu = 1 + \left(\sum_{k=0}^{m} E_k \langle \phi_k | V \phi_k \rangle \right) \Big/ \sum_{k=0}^{m} E_k^2. \qquad (5.7)$$

The former criteria will be more useful when ϕ_0 is the dominant eigenfunction in the expansion for the exact solution ψ whereas the latter is more appropriate when there are several ϕ_n giving a significant contribution to ψ. The theory of scaling has been discussed by Musher and Schulman (1968) and Amos (1970) and these authors obtain the value of μ given by (5.5). We do not know if our natural generalisation of this, (5.7), has been used explicitly.

6. Example calculations

In tables 1 and 2 we present the results of calculations with H and H_0 given by (1.3) and (1.4) respectively. The values listed are approximations to the eigenvalues of H. In table 1 the ground state is considered whereas in table 2 the results refer to the first excited state.

The first two rows give the results from the iterative schemes (3.2) and (3.9) which are equivalent to the minimisation of (3.8). The numbers in the brackets are the number of iterations needed so that the results are consistent to six decimal places. It can be seen that a maximum of seven iterations are needed and consequently this is a fairly trivial numerical problem and provides a useful alternative to the direct variation of (3.8). In fact increasing the number of iterations to a maximum of nine ensures the results consistent to ten decimal places.

Only seven eigenstates of H_0 are used in the trial functions in this model calculation but these provide fairly accurate results. Exact results for x^4 and $\frac{1}{10}x^4$ have been given by Hioe *et al* (1978) and these are 0.803 771 and 0.559 146. Our results may also be compared with the calculations done by Turbiner (1981) who obtains 0.804 468 and 0.561 658 respectively. The perturbation and generalised perturbation calculations which are given in tables 1 and 2 also used only seven eigenstates so that the variational calculations provide the 'exact' values within this model for the perturbation results.

The third and fourth rows give the estimates provided by standard perturbation theory and the equivalent calculations when H_0 is scaled by μ obtained from (5.5). The first value is $E_0 + E^1$, the second is $E_0 + E^1 + E^2$ and the final value is the results obtained by tenth-order perturbation theory when this gives answers consistent with those obtained from the variational methods. In other cases we do not have convergence for the tenth-order perturbation theory. The standard perturbation theory used is truncated perturbation theory so that only the first seven eigenstates of H_0 are used.

The use of μ causes convergence for $\frac{7}{100}x^4$ and, as expected, is unnecessary for the smaller perturbations such as $\frac{1}{100}x^4$.

In rows 5 and 6 the results of the generalisation of standard perturbation given in (4.1) and (4.8) are given with and without the adjustment involving μ. The theory predicts $E_0 + E^1$ and $E_0 + E^1 + E^2$ are the first two energies from the iteration and the final figure is that obtained from the tenth iteration. The results are approximately the same as in standard perturbation theory.

The results given in the final two rows are based on the generalisation of perturbation theory presented in (4.1) and (4.2) with and without the use of μ. Since seven eigenstates were used (4.5) was used only 6 times to produce the final results. It can be seen that the convergence of the method is superior to the previous perturbation results. Only the final energy (given by (4.6)) is listed and the places given are the number which agree with the variational calculations. These results encourage us to believe that this generalisation provides a useful alternative to standard perturbation theory.

To summarise, the abstract theory described by Amos (1970) and Burrows and Perks (1981a) has been applied in two ways to quantum mechanical calculations. The first variant gives alternative iterative methods of using the conventional variational theory. The advantages of the method are twofold. Firstly roundoff errors are confined to the final iteration and secondly the calculations can easily be applied to excited states provided that good initial approximations are available. The second variant provides various generalisations of perturbation theory depending on how the sequence is chosen. One of the choices considered (equation (4.4)) provides a method for which there is some evidence of superior convergence properties. It is a simple method which involves the calculation of sums similar to conventional perturbation theory so that the work involved is comparable. The usual scaling methods can be employed and lead to improved results.

References

Amos A T 1970 *J. Chem. Phys.* **52** 603
Burrows B L and Core P W 1984 *J. Math. Anal. Appl.* to be published
Burrows B L and Perks A J 1981a *J. Math. Anal. Appl.* **83** 77–86
—— 1981b *J. Phys. A: Math. Gen.* **14** 797–808
Hioe F T, McMillen D and Montroll E W 1978 *Phys. Rep.* **43C** 307–35
Musher J I and Schulman J 1968 *J. Chem. Phys.* **49** 4845
Turbiner A V 1981 *J. Phys. A: Math. Gen.* **14** 1641–9
Ortega J M and Rheinboldt W F 1970 *Iterative solution of Non-linear Equations in General Variables* (London: Academic)