A comparison of some lower bounds for eigenvalues of Schrodinger's equation

1983 J. Phys. A: Math. Gen. 16 3291

(http://iopscience.iop.org/0305-4470/16/14/021)

J. Phys. A: Math. Gen. **16** (1983) 3291–3311. Printed in Great Britain

A comparison of some lower bounds for eigenvalues of Schrödinger's equation

Peter G Hornby and Michael N Barber

Department of Applied Mathematics, University of New South Wales, PO Box 1, Kensington, NSW 2033, Australia

Received 24 January 1983

Abstract. A method proposed recently by Singh of finding lower bounds to the eigenenergies of Schrödinger's equation is reviewed and compared with the methods of Bazley and Fox and of Löwdin. Numerical results for an anharmonic oscillator suggest that the Singh method is less accurate than the others, but unlike the Löwdin single bracketing function technique does not suffer from any eigenvalue ordering problems.

1. Introduction

Upper bounds to the eigenvalues of the time independent Schrödinger equation

$$H\psi = E\psi, \tag{1.1}$$

where H is the Hamiltonian of the system, are readily obtained by the traditional Rayleigh–Ritz variational method. Lower bounds are less easily calculated but are necessary to locate the eigenvalues with certainty. This objective has led to the derivation of several methods to obtain lower bounds (see e.g. Temple 1928, Bazley and Fox 1961, Löwdin 1965a, b, Reid 1976, Hill 1980, Singh 1981 and other references cited in these papers). Some of these methods, notably those of Bazley and Fox (1961), Löwdin (1965a, b) and Hill (1980), are capable of producing lower bounds to both ground states and excited states of considerable accuracy. In a recent paper, Singh (1981) has proposed what appears to be a new method of constructing lower bounds that moreover yields a sequence of improving bounds. No numerical results are reported, however, so it is difficult to assess the practical significance of Singh's procedures. This is the main aim of this paper. In addition, we compare and contrast the *formal* basis of Singh's method with the Löwdin theory.

The paper is arranged as follows. In § 2 we review the formalism developed by Singh and compare it with that of Löwdin. Section 3 illustrates the procedure by explicitly evaluating lower bounds to a number of energy levels of an anharmonic oscillator. These bounds are compared both in accuracy and computational effort with other bounds in § 4 and the paper closes with an overall discussion and assessment in § 5.

2. Singh's procedure

2.1. Formal results

Consider a Hermitian operator H of the form

$$H = H_0 + V, \tag{2.1}$$

where H_0 and V are Hermitian with $V \geqslant 0$. The eigenvectors $|E_0^j\rangle$ and eigenvalues E_0^j of H_0 are assumed known and ordered such that $E_0^j < E_0^{j+1}$. We shall also require that

$$E_0^j < 0 \qquad \text{for } j < J_0. \tag{2.2}$$

This can, however, usually be arranged by a simple redefinition of the energy zero. Our aim will be to bound from below the eigenvalues E^j of H satisfying

$$H|E^j\rangle = E^j|E^j\rangle. \tag{2.3}$$

Define the orthogonal projections

$$P_0 = \sum_{i=0}^{J_0-1} |E_0^i\rangle\langle E_0^i|, \qquad O_0 = 1 - P_0, \tag{2.4a, b}$$

together with

$$\tilde{H} = \tilde{H}_0 + V = H_0 O_0 + V = H_0 - H_0 P_0 + V. \tag{2.5}$$

For notational ease, we shall denote the spectrum of an operator G by $\sigma(G)$, its image by Image(G) and the linear span of a set $\{|\phi^k\rangle\}_{k=1}^m$ by $\text{Span}_{k=1,m}\{|\phi^k\rangle\}$.

The Singh procedure depends upon using (2.5) to rewrite the eigenvalue equation (2.3) of H as

$$(\tilde{H} - E)|E\rangle = -H_0 P_0 |E\rangle. \tag{2.6}$$

This may be regarded as an implicit set of equations for E and $|E\rangle$ which can be 'solved' whenever $E \notin \sigma(\tilde{H})$. In this case, we may write

$$|E\rangle = (\tilde{H} - E)^{-1} P_0 (-H_0) P_0 |E\rangle. \tag{2.7}$$

Following Singh, we define

$$|\beta(E)\rangle = (-H_0)^{1/2} P_0 |E\rangle. \tag{2.8}$$

In view of (2.2), $-H_0$ is positive on Image(P_0) so that $|\beta(E)\rangle$ is non-zero whenever $\langle E_0^j|E\rangle \neq 0$ for some j in $0 \leqslant j \leqslant J_0 - 1$. The factor $(-H_0)^{1/2}$ appearing in (2.8) is not entirely necessary but it does simplify subsequent formulae.

Equation (2.7) still represents an implicit set of equations for E and $|E\rangle$, which can obviously be rewritten as

$$|E\rangle = (\tilde{H} - E)^{-1} P_0 (-H_0)^{1/2} |\beta(E)\rangle. \tag{2.9}$$

Multiplying on the left by $(-H_0)^{1/2} P_0$ yields the eigenvalue problem

$$|\beta(E)\rangle = \mathscr{B}(E)|\beta(E)\rangle, \tag{2.10}$$

where

$$\mathscr{B}(E) = (-H_0)^{1/2} P_0 (\tilde{H} - E)^{-1} P_0 (-H_0)^{1/2}. \tag{2.11}$$

Now we can identify $|\beta(E)\rangle$ as an eigenvector of $\mathscr{B}(E)$ in Image(P_0) with eigenvalue unity. Moreover, recalling (2.9), $|\beta(E)\rangle$ is rotated into $|E\rangle$, an eigenvector of H, by the operator $(\bar{H}-E)^{-1}P_0(-H_0)^{1/2}$.

It is natural now to examine the J_0 eigenvectors $|\beta^j(z)\rangle$ in Image(P_0) and their associated eigenvalues $\beta^j(z)$ of

$$\mathscr{B}(z)=(-H_0)^{1/2}P_0(\bar{H}-z)^{-1}P_0(-H_0)^{1/2}. \tag{2.12}$$

Let us confine our attention for a moment to one eigenvector $|\beta(z)\rangle$ of $\mathscr{B}(z)$ in Image(P_0). Since $\mathscr{B}(z)$ is continuous for $z\notin\sigma(\bar{H})$ it follows that $\beta(z)$ is also continuous on this set (Singh 1981). Hence, given that

$$\mathscr{B}(z)|\beta(z)\rangle=\beta(z)|\beta(z)\rangle \tag{2.13}$$

we define the 'trial wavefunction' associated with $\beta(z)$ by

$$|z\rangle=(\bar{H}-z)^{-1}P_0(-H_0)^{1/2}|\beta(z)\rangle. \tag{2.14}$$

The significance of those vectors $|\beta(z)\rangle$ for which $\beta(z)=1$ can be seen if we compute

$$\begin{aligned}(H-z)|z\rangle&=(\bar{H}-z+H_0P_0)(\bar{H}-z)^{-1}P_0(-H_0)^{1/2}|\beta(z)\rangle\\&=[1+H_0P_0(\bar{H}-z)^{-1}]P_0(-H_0)^{1/2}|\beta(z)\rangle\\&=(1-\beta(z))P_0(-H_0)^{1/2}|\beta(z)\rangle.\end{aligned} \tag{2.15}$$

Thus $\beta(z)=1$ implies $(H-z)|z\rangle=0$, which together with (2.7)–(2.11) establishes a correspondence between the eigenvalues E of H such that $\langle E_0^j|E\rangle\neq 0$ for some $j<J_0$ and the values of z for which $\mathscr{B}(z)$ has an eigenspace in Image(P_0) with eigenvalue one.

The preceding analysis is easily generalised to degenerate eigenvalues. Let $|E\rangle_\alpha$, $\alpha=1,2,\ldots,k$, be k linearly independent eigenvectors of H with eigenvalue E where $E\notin\sigma(\bar{H})$. Defining

$$|\beta(E)\rangle_\alpha=(-H_0)^{1/2}P_0|E\rangle_\alpha \tag{2.16}$$

we see that $|\beta(E)\rangle_\alpha$ satisfies (2.10) for all α with $\beta(E)=1$. Moreover, using (2.9) it is straightforward to show that the vectors $|\beta(E)\rangle_\alpha$ are linearly independent and hence the eigenspace of $\mathscr{B}(z;z=E)$ with eigenvalue unity is at least k-fold degenerate.

A similar investigation of (2.9) reveals that, for $E\notin\sigma(\bar{H})$, the existence of k linearly independent eigenvectors $|\beta(E)\rangle_\alpha$ of $\mathscr{B}(E)$ satisfying (2.10) implies the existence of k linearly independent eigenvectors $|E\rangle$ of H with eigenvalue E. Hence we conclude that for all $z\notin\sigma(\bar{H})$, $\mathscr{B}(z)$ has a k-fold degenerate eigenspace with eigenvalue unity if and only if z is a k-fold degenerate eigenvalue of H.

To illuminate this correspondence further, especially regarding the importance of $\sigma(\bar{H})$, suppose E is a k-fold degenerate eigenvalue of H. Let also J_0 be such that $k>J_0$. Hence the eigenspace of $\mathscr{B}(E)$ with eigenvalue unity is, at most, of dimension J_0, contradicting the proposed correspondence. Hence, we must have $E\in\sigma(\bar{H})$, which can be understood as follows.

In the k-fold degenerate eigenspace of H with eigenvalue E, there exist at least $k-J_0$ linearly independent vectors orthogonal to Image(P_0). Let $|E\rangle$ be such a vector; then

$$H|E\rangle=(\bar{H}+H_0P_0)|E\rangle=\bar{H}|E\rangle=E|E\rangle \tag{2.17}$$

so that $E\in\sigma(\bar{H})$. Now, although $E\in\sigma(\bar{H})$ implies that $(\bar{H}-E)^{-1}$ does not exist, we notice that the eigenspace of $\bar{H}$ with eigenvalue E is orthogonal to Image(P_0) and

hence $(\bar{H}-\bar{E})^{-1}$ is finite on Image(P_0), implying in turn that $\mathscr{B}(E)$ is bounded. Similarly, each of the other linearly independent eigenvectors of H with eigenvalue E *and* a non-zero overlap with Image(P_0) forces the existence of an independent vector $|\beta(E)=1\rangle$ in Image(P_0) satisfying (2.10). The same type of mechanism also takes care of disjoint sectors of H.

We turn now to the singularities of the functions $\beta(z)$. These can only occur when $z \in \sigma(\bar{H})$ and, as we have just seen, need not always occur then. What is necessary for some of the β's to become singular at $z=\bar{E}\in\sigma(\bar{H})$ is a non-zero overlap between Image(P_0) and the eigenspace of $\bar{H}$ with eigenvalue $\bar{E}$. Denote by $\bar{\mathscr{E}}$ the projection of this eigenspace of $\bar{H}$ onto Image(P_0). It is easy to see that, as $z\to\bar{E}$, a subspace of Image(P_0), equal in dimension to $\bar{\mathscr{E}}$, will become an eigenspace of $\mathscr{B}(z)$ associated with a diverging eigenvalue. The Hermiticity of $\mathscr{B}(z)$ as $z\to\bar{E}$ ensures that the rest of $\mathscr{B}$'s eigenvectors, being orthogonal to the diverging eigenvectors, are rotated orthogonal to these diverging eigenvectors quickly enough as $z\to\bar{E}$ to escape divergence themselves. This heuristic argument suggests that there is a one-to-one correspondence between the number of functions $\beta(z)$ that diverge as $z\to\bar{E}$ and the number of linearly independent vectors in $\bar{\mathscr{E}}$.

Whenever β is regular $\mathrm{d}\beta/\mathrm{d}z$ follows from the Feynman–Hellmann theorem (Singh 1981). Explicitly,

$$\begin{aligned}\mathrm{d}\beta/\mathrm{d}z &= \langle\beta(z)|\mathrm{d}\mathscr{B}/\mathrm{d}z\,|\beta(z)\rangle = \langle\beta(z)|(-H_0)^{1/2}P_0(\bar{H}-z)^{-2}P_0(-H_0)^{1/2}|\beta(z)\rangle \\ &= \langle z|z\rangle > 0 \qquad \text{for } |z\rangle \neq 0. \end{aligned} \tag{2.18}$$

On the other hand, if $|z\rangle=0$, then (2.14) implies that $(-H_0)^{1/2}P_0|\beta(z)\rangle=0$ which contradicts the fact that $|\beta(z)\rangle$ is a non-zero eigenvector. Thus $\mathrm{d}\beta/\mathrm{d}z>0$ whenever $\beta(z)$ is regular.

Equation (2.18) and the pole structure of the $\beta(z)$ functions lead us to conclude that if we trace the behaviour of one eigenvalue of $\mathscr{B}(z)$ as z varies over the reals we obtain a function $\beta(z)$ something like the function depicted by full curves in figure 1.

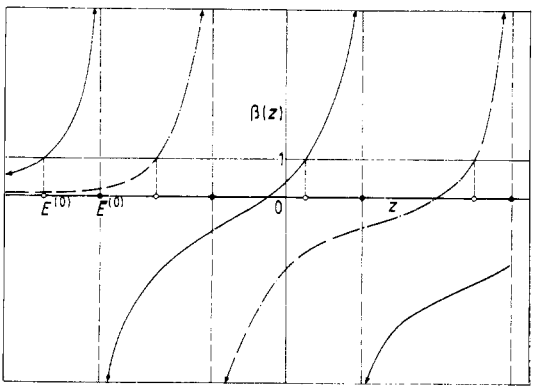

Figure 1. Typical behaviour of the β functions involved in a $J_0=2$ calculation on a Hamiltonian with non-degenerate eigenvalues and no disjoint sectors. The function depicted by full curves picks up $E^0, E^2, E^4\ldots$ and is singular for $\bar{E}^l$ such that $\langle\bar{E}^l|\beta(\bar{E}^l)\rangle\neq 0$, that is $\bar{E}^0, \bar{E}^2, \bar{E}^4\ldots$. The other eigenvalue of $\mathscr{B}(z)$ is depicted by broken lines and picks up $E^1, E^3, \ldots$.

The criterion, $\beta(z)=1$, for an eigenenergy of H is conveniently recast as a fixed point problem (Singh 1981). Let $\alpha(z)$ be defined by

$$\beta(z)=(E_u-\alpha(z))/(E_u-z), \qquad E_u \notin \sigma(H). \tag{2.19}$$

Some algebra establishes that

$$\mathscr{A}(z)|\beta(z)\rangle=\alpha(z)|\beta(z)\rangle, \tag{2.20}$$

where

$$\begin{aligned}\mathscr{A}(z)&=E_uP_0-(E_u-z)\mathscr{B}(z)=E_uP_0+(-H_0)^{1/2}P_0(\bar{H}-E_u-\bar{H}+z)(\bar{H}-z)^{-1}P_0(-H_0)^{1/2}\\&=(E_u+H_0)P_0+(-H_0)^{1/2}P_0(\bar{H}-E_u)(\bar{H}-z)^{-1}P_0(-H_0)^{1/2}.\end{aligned} \tag{2.21}$$

Setting $A=\bar{H}-E_u$ we obtain

$$\mathscr{A}(z)=(H_0+E_u)P_0+(-H_0)^{1/2}P_0A[A^2+(E_u-z)A]^{-1}AP_0(-H_0)^{1/2}, \tag{2.22}$$

which is the basis of Singh's exposition. There is a peculiarity at $z=E_u$ where

$$\mathscr{A}(E_u)=E_uP_0. \tag{2.23}$$

So Image(P_0) is an eigenspace of $\mathscr{A}(E_u)$ associated with the eigenvalue E_u. Apart from this exception at $z=E_u$ there is a one-to-one correspondence between the fixed points z^* of $\alpha(z)$ and the solutions to $\beta(z)=1$.

Since we are only interested in Image (P_0) when finding eigenvectors of $\mathscr{A}(z)$, we may regard $\mathscr{A}(z)$ as a finite-dimensional matrix. The operator $\mathscr{H}_0=(H_0+E_u)P_0$ is thus regarded as a $J_0\times J_0$ diagonal matrix with respect to the basis $\{|E_0^j\rangle\}_{j=0}^{J_0-1}$ and

$$\nu(z)=\mathscr{A}(z)-\mathscr{H}_0=(-H_0)^{1/2}P_0A[A^2+(E_u-z)A]^{-1}AP_0(-H_0)^{1/2} \tag{2.24}$$

as a $J_0\times J_0$ perturbing matrix.

2.2. *Lower bounds*

Figure 1 suggests that finding lower bounds to z^* such that $\beta(z^*)=1$ could be accomplished by finding a function $\beta_n(z)$ which upper bounds $\beta(z)$ on the entire real line. Singh's method is less ambitious in that it upper bounds the $\beta(z)$'s in the interval $(-\infty, E_u)$. This is done by finding $\alpha_n(z)$ which are lower bounds to $\alpha(z)$ on the interval $(-\infty, E_u)$. The $\beta_n(z)$ defined by

$$\beta_n(z)=(E_u-\alpha_n(z))/(E_u-z) \tag{2.25}$$

then have the required bounding property.

To lower bound $\alpha(z)$ it is necessary to find an operator lower bound to $\mathscr{A}(z)$, which can be achieved by modifying the $\nu(z)$ term (2.24). The success of the method requires the positivity of $\nu(z)$ which is in turn assured by the positive definiteness of A. This imposes the following conditions on E_u:

$$E_u<E_0^{J_0}, \qquad E_u<0. \tag{2.26}$$

It is also desirable to make E_u as large as possible so as to lower bound as many fixed points as possible. The condition

$$E_0^{J_0-1}<E_u \tag{2.27}$$

would seem reasonable.

Following Singh we approximate the operator

$$\bar{\nu}(z) = A[A^2 + (E_u - z)A]^{-1}A \tag{2.28}$$

from below by a sequence of operators $\{\bar{\nu}_n(z)\}$ obtained using Bubnov–Galerkin approximations (Mikhlin 1964, Singh 1977, 1981). Let $|\alpha\rangle \in \text{Image}(P_0)$ and define

$$|f\rangle = [A^2 + (E_u - z)A]^{-1}A|\alpha\rangle; \tag{2.29}$$

then $|f\rangle$ satisfies

$$[A^2 + (E_u - z)A]|f\rangle = A|\alpha\rangle. \tag{2.30}$$

Defining

$$L(z) = A^2 + (E_u - z)A \tag{2.31}$$

and

$$|\alpha'\rangle = A|\alpha\rangle \tag{2.32}$$

equation (2.30) becomes

$$L(z)|f\rangle = |\alpha'\rangle. \tag{2.33}$$

The Bubnov–Galerkin method approximates $|f\rangle$ with the finite sum

$$|f_n\rangle = \sum_{k=1}^{n} C_k|\phi^k\rangle, \tag{2.34}$$

where $\{|\phi^k\rangle\}_{k=1}^n$ is a set of vectors in the domain of A. The C_k are then determined by requiring that the 'error', $L|f_n\rangle - |\alpha'\rangle$, be orthogonal to $\text{Span}_{k=1,\ldots,n}\{|\phi^k\rangle\}$. This leads to the set of equations

$$\sum_{k=1}^{n} \langle\phi^l|A^2 + (E_u - z)A|\phi^k\rangle C_k = \langle\phi^l|A|\alpha\rangle. \tag{2.35}$$

In particular, $|f_n\rangle \in \text{Span}_{k=1,\ldots,n}\{|\phi^k\rangle\}$ so that

$$\langle f_n|\alpha'\rangle - \langle f_n|L|f_n\rangle = 0, \tag{2.36}$$

which on substituting (2.33), yields

$$\langle f_n|L(z)|f - f_n\rangle = 0. \tag{2.37}$$

Defining

$$\langle\alpha|\bar{\nu}_n(z)|\alpha\rangle = \langle\alpha'|f_n\rangle \tag{2.38}$$

it follows that

$$\langle\alpha|\bar{\nu}(z) - \bar{\nu}_n(z)|\alpha\rangle = \langle\alpha'|f - f_n\rangle = \langle f|L|f - f_n\rangle = \langle f|L|f - f_n\rangle - \langle f_n|L|f - f_n\rangle$$
$$= \langle f - f_n|L|f - f_n\rangle \geqslant 0. \tag{2.39}$$

The matrix elements of $\bar{\nu}_n(z)$ in $\text{Image}(P_0)$ are calculated by letting

$$|f_n^j\rangle = \sum_k C_k^j|\phi^k\rangle, \qquad j = 0, \ldots, J_0 - 1, \tag{2.40}$$

$$G_i^j = \langle\phi^i|A|E_0^j\rangle, \qquad j = 0, \ldots, J_0 - 1, \tag{2.41}$$

$$l_{ij}(z) = \langle\phi^i|L(z)|\phi^j\rangle, \tag{2.42}$$

and solving

$$\sum_{k=1}^{n} l_{ik}(z)C_k^j = G_i^j, \qquad i = 1, \ldots, n, \qquad j = 0, \ldots, J_0 - 1, \tag{2.43}$$

for the C_k^j. Thus

$$\bar{\nu}_n^{ij}(z) = \langle E_0^i | \bar{\nu}_n(z) | E_0^j \rangle = \sum_{k=1}^{n} C_k^i G_k^{j*}, \qquad i, j = 0, \ldots, J_0 - 1, \tag{2.44}$$

where * denotes complex conjugation. The required finite matrix approximation for $\mathscr{A}(z)$ is

$$\begin{aligned}\mathscr{A}_n^{ij}(z) &= \langle E_0^i | \mathscr{A}_n(z) | E_0^j \rangle = \langle E_0^i | \mathscr{H}_0 + (-H_0)^{1/2} \bar{\nu}_n(z)(-H_0)^{1/2} | E_0^j \rangle \\ &= (E_0^i + E_u)\delta_{ij} + \sum_{k,l=0}^{J_0-1} (-E_0^i)^{1/2} \delta_{ik} \bar{\nu}_n^{kl}(z) \delta_{lj} (-E_0^j)^{1/2}.\end{aligned} \tag{2.45}$$

The convergence of this method is assured by noticing that if the $|\phi_k\rangle$ are normalised according to

$$\langle \phi^i | A^2 | \phi^j \rangle = (\phi^i, \phi^j) = \delta_{ij} \tag{2.46}$$

then (2.35) becomes

$$\sum_{j=1}^{\infty} [\delta_{ij} + (E_u - z)(\phi^i, A^{-1}\phi^j)]C_j = (\phi^i, A^{-1}\alpha), \tag{2.47}$$

where the $\{|\phi^k\rangle\}_{k \in \mathbb{N}}$ span the closure with respect to $(\cdot, \cdot)$ of the domain of A. For $(E_u - z) > 0$, (2.47) has a unique solution. A sufficient additional condition for the convergence of $\langle \alpha | A | f_n \rangle$ to $\langle \alpha | A | f \rangle$ is the compactness of A^{-1} with respect to $(\cdot, \cdot)$ on the closure of the domain of A. In the case of the anharmonic oscillator A will be positive bounded below and unbounded above with discrete spectrum. This is sufficient to ensure convergence of the $\alpha_n(z)$'s to their respective $\alpha(z)$'s (Mikhlin 1964, Singh 1977).

Given the matrix $\mathscr{A}_n^{ij}(z)$ and the functions $\alpha_n(z)$ it remains to show that the functions $\beta_n(z)$ defined by (2.25) have the required properties of $\beta(z)$. This Singh does by constructing a matrix $\mathscr{B}_n^{ij}(z)$ from $\mathscr{A}_n^{ij}(z)$ with eigenvalues $\beta_n(z)$ and the same eigenvectors as $\mathscr{A}_n^{ij}(z)$. In addition the β_n and $\mathrm{d}\beta_n/\mathrm{d}z$ are all positive in $(-\infty, E_u)$, showing that the $\alpha_n(z)$ each have, at most, one fixed point in $(-\infty, E_u)$. These fixed points are moreover stable. The $\beta_n(z)$'s are, however, singular at $z = E_u$ unless $\alpha_n(z) \to E_u$ quickly enough as $z \to E_u$. Finally we note that the 'approximate trial wavefunction' associated with $\beta_n(z)$ is

$$|z\rangle_n = \sum_{j=0}^{J_0-1} (-E_0^j)^{1/2} \langle E_0^j | \beta_n(z) \rangle | f_n^j \rangle \tag{2.48}$$

and is an estimate of the eigenvector of H corresponding to the energy $E \simeq z_n^*$ such that $\beta_n(z_n^*) = 1$.

2.3. Comparison with Löwdin's bracketing function technique

The Singh procedure, described in §2.2, is, at least superficially, similar to the partitioning technique of Löwdin (1962, 1965a, b) based on a bracketing function. (For recent reviews see Reid 1976, Abdel-Raouf 1982.)

The Löwdin scheme centres on the reduced resolvent† (see e.g. Reid 1976)

$$T(z) = O(z - OHO)^{-1}O, \tag{2.49}$$

where $O = 1 - P$ with

$$P = \sum_{k=0}^{J_0-1} |\phi^k\rangle\langle\phi^k|. \tag{2.50}$$

The operator P is a projection operator into some appropriate 'reference space'. The 'bracketing functions' $f(z)$ are the eigenvalues of

$$F(z) = P[H + HT(z)H]P \tag{2.51}$$

associated with Image(P). These functions have fixed points for $z \in \sigma(H)$, singularities corresponding to the spectrum of OHO associated with Image(O), and negative gradients. Further, the singularities and fixed points of the $f(z)$ are subject to the same sort of overlap conditions with Image(P) as those found in the Singh scheme (Wilson 1967). Thus the functions $f(z)$ have similar properties to the functions α introduced in (2.25), though f and α are clearly not identical. For instance, the gradients of the α's are positive under the stated assumptions about the spectrum of H_0 and E_u.

Similarly, Singh's functions $\beta(z)$ are somewhat analogous to the eigenvalues $w(z)$ of

$$W(z) = P(z - H)^{-1}P \tag{2.52}$$

associated with Image(P), e.g. compare figure 1 with figure 1 of Löwdin (1965b). On the other hand, the β functions differ from the $w(z)$'s in three significant respects. Firstly, the gradients of the β functions are opposite in sign to those of the $w(z)$'s. Secondly, the fixed points of the bracketing functions $f(z)$ are determined by the singularities of the w's while it is the zeros of $\beta - 1$ that determine the fixed points of the α's. Thirdly, the singularities of the bracketing functions are determined by the zeros of the w's. In contrast, the singularities of the β's force the divergence of their respective α functions. These differences are due to the fact that the w's appear in the denominators of the f's (Löwdin 1965b). On the other hand, the β's are more directly related to the α's.

The factor of $(-H_0)^{1/2}$ appearing in (2.12) can now be seen as a non-orthogonal projection which lends uniformity to the important values of the β functions. This point will be discussed further in § 5.

While one can thus draw some analogies between the technical aspects of the formalisms of Löwdin and Singh, these should not be allowed to obscure what is a fundamental difference in philosophy. In the Löwdin scheme, the fixed point of the bracketing function is *not* explicitly calculated but rather the bracketing function (or in practice an approximation to it, see appendix 2) is evaluated at a *known* upper bound on the required eigenvalue. Singh's method, however, attempts to find the fixed point. In addition, the Löwdin method attempts to approximate all the fixed points of a single bracketing function ($J_0 = 1$) to obtain the energy lower bounds. The Singh method, on the other hand, bounds larger energy eigenvalues by enlarging the

† In earlier work (e.g. Löwdin 1962, 1965a, b) T is defined as $O(\alpha + O(E - H)O]^{-1}O$ where α is an arbitrary constant other than zero. We follow Reid (1976) and use the simpler form which is equivalent to the earlier. In order to agree with our earlier notation we have interchanged O and P compared with the notation of Reid and Löwdin.

reference space and finding the smallest fixed points of the set of α functions generated. This is also possible in the Löwdin scheme (Löwdin 1962, Wilson 1967, Abdel-Raouf 1981), but appears not to have been fully exploited. How the two methods fare in practice is a question we address in § 4.2.

3. The anharmonic oscillator

As a test of the practical significance of the theoretical developments of §§ 2.1 and 2.2, we consider an anharmonic oscillator with Hamiltonian

$$H = (-\mathrm{d}^2/\mathrm{d}x^2 + x^2) + \lambda x^4, \tag{3.1}$$

which we decompose as

$$H_0 = -\mathrm{d}^2/\mathrm{d}x^2 + x^2 \tag{3.2}$$

and

$$V = \lambda x^4. \tag{3.3}$$

The state spaces of both H_0 and H divide naturally into two sectors of well defined parity. We restrict attention to the even parity sector.

The eigenvalues and eigenvectors of H_0 are given by (see e.g. Schiff 1968)

$$H_0|E_0^i\rangle = E_0^i|E_0^i\rangle, \qquad E_0^i = 4i + 1, \tag{3.4}$$

and

$$\psi_i(x) = \langle x|E_0^i\rangle = c_i\, \mathrm{e}^{-x^2/2} H_{2i}(x), \qquad i = 0, 1, 2, \ldots, \tag{3.5}$$

where $H_j(x)$ are the Hermite polynomials and $c_i = 2^{-i}[(2i)!]^{-1/2}\pi^{-1/4}$. In this basis

$$\begin{aligned}\langle E_0^i|V|E_0^j\rangle/\lambda &= \tfrac{1}{4}[(2i+5)(2i+4)(2i+3)(2i+2)]^{1/2}\delta_{i+2,j} \\ &\quad + \tfrac{1}{2}(4i+3)[(2i+2)(2i+1)]^{1/2}\delta_{i+1,j} \\ &\quad + (8i^2+4i+1)\delta_{i,j} + \tfrac{1}{2}(4i+1)[2i(2i-1)]^{1/2}\delta_{i-1,j} \\ &\quad + \tfrac{1}{4}[2i(2i-1)(2i-2)(2i-3)]^{1/2}\delta_{i-2,j},\end{aligned} \tag{3.6}$$

so that H is a band matrix of lower band width two. This band character is essential to the finiteness of the Bubnov–Galerkin approximation if $\langle A\psi_i|A\psi_j\rangle$ is to be calculated in the $\{|E_0^i\rangle\}$ basis. Otherwise an assumption similar to a special choice (Bazley and Fox 1961) is needed to provide a finite basis in which $\langle A\phi^i|A\phi^j\rangle$ can be calculated.

To implement the bounding procedure described in § 2 we chose the first n (even parity) eigenstates of H_0 to span the manifold $\mathrm{Span}_{k=1,\ldots,n}\{|\phi^k\rangle\}$, the actual set $\{|\phi^k\rangle\}_{k=1}^n$ being calculated by a Gram–Schmidt orthogonalisation of $\{|E_0^j\rangle\}_{j=1}^n$ with respect to the bilinear form on A^2. This requires the calculation of inner products of the form $\langle A\psi_i|A\psi_j\rangle$ which in turn requires knowledge of $N = n + 2$ unperturbed eigenstates of H_0 and the matrix elements of A between them. This number, N, of basis elements necessary for the calculation will, in § 4, be an important criterion in comparing the Singh method with other bounds.

Next the number J_0 of basis elements projected out of H_0 is chosen together with an appropriate E_u satisfying (2.26). Since H_0 has a positive spectrum it is necessary

to shift its spectrum, along with E_u, in such a way that the J_0 low-lying eigenstates of H_0 acquire negative energies. A shift of $E_s \leqslant -E_u$ is sufficient for this purpose. Once the desired fixed point in $(-\infty, E_u)$ is found the energy background E_s is removed from the estimate to obtain the actual bounds to the energies of (3.1). The fixed points of the $\alpha_n(z)$ were found using a modified Müller method (Blatt 1975) to carry out an inverse interpolation on $\alpha_n(z)-z$. This usually required fewer evaluations of $\alpha_n(z)$ than a simple fixed point iteration.

Before comparing the Singh method results with those of alternative methods we consider two preliminary questions regarding the sensitivity of the fixed points z_n^* to choices of E_u and the negative bias E_s.

Figure 2 is a plot of the ground state energy estimates for $\lambda = 5$ against E_s with E_u chosen so that

$$E_u - E_s = 4.9. \tag{3.7}$$

Adjusting E_s in this way effectively slides the whole calculation along the real axis. We see that the estimates show, for modest N, small variation relative to the error in the estimate itself. The small N estimates can, however, be improved significantly by choosing E_s as large as permissible. The value of E_s cannot be made so large that the biased E_u becomes greater than zero, for then the positive definiteness of A may be lost. Consequently E_s is chosen to slide all subsequent calculations along the real axis in such a way that the biased value of E_u becomes zero. The operator A remains positive definite for $\lambda > 0$ in this case, since $V > 0$.

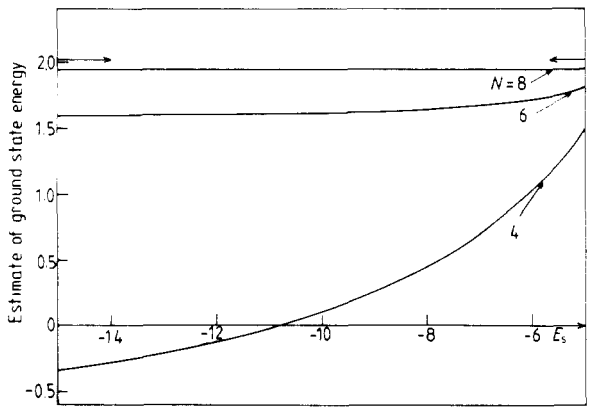

Figure 2. Variation of estimates of ground state energy for $\lambda = 5$ with the value of the energy bias E_s chosen for the calculation. The unbiased value of E_u is 4.9 and the arrows locate an accurate upper bound of 2.018 340 65 which is accurate to the indicated figure.

Figure 3 is a plot of the estimates obtained for the ground state of H for $\lambda = 5$ against the unbiased values of E_u chosen for the calculation, i.e. the values of E_u *before* the bias of $E_s = -E_u$ is applied to the system. Each time E_u crosses an unperturbed energy E_0^j, the associated eigenvector $|E_0^j\rangle$ is included in Image(P_0). Even so the estimates show continuously increasing behaviour until Image(P_0) becomes too big for the chosen $|\phi^k\rangle$ set to handle. This occurs when J_0 exceeds n. It should be noted that each time a new element is added to Image(P_0) the set of equations (2.43) must be re-solved for the extra vector $\boldsymbol{G}^{J_0}$. In addition the extra

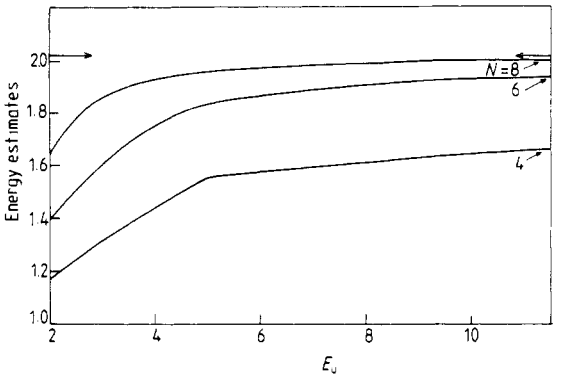

Figure 3. Variation of ground state energy estimates for $\lambda = 5$ with the unbiased value of E_u chosen for the calculation. E_s is chosen to slide the calculation along the real axis in such a way that the biased value of E_u becomes zero.

elements of $\bar{\nu}^{ij}(z)$ must be calculated according to (2.44). This has to be performed each time $\mathscr{A}_n^{ij}(z)$ is required for a particular z.

The estimates quoted in figure 4 and the tables of § 4 have been calculated using values of E_u which were as large as practical without making Image(P_0) unnecessarily large. For example, to find the ground state estimates in figure 4 the unbiased value of E_u was chosen as 4.9 so that the dimension of Image(P_0) was one. If, on the other hand, an eigenvalue E^j crossed the next unperturbed level E_0^{j+1} then the unperturbed state $|E_0^{j+1}\rangle$ is included in Image(P_0) and the unbiased value of E_u is set to a value slightly smaller than E_0^{j+2}. This occurs in table 2 for the $j = 2$ (quantum number = 4) eigenvalue estimates.

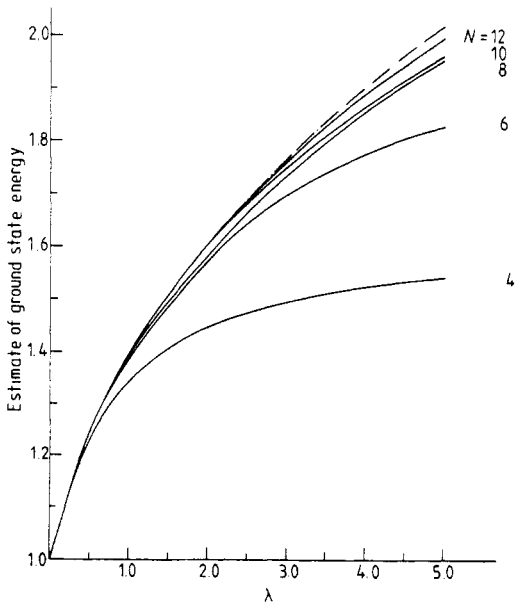

Figure 4. Dependence of the ground state energy estimates upon N and λ, compared with an accurate upper bound.

We now turn to the general behaviour of the estimates. The plot in figure 4 shows the λ dependence of the ground state estimates for several different values of N. These are compared with an accurate upper bound obtained by finding the smallest eigenvalue of the 60×60 matrix

$$\langle E_0^i | H | E_0^j \rangle, \qquad i, j = 0, 1, \ldots, 59. \tag{3.8}$$

This upper bound agrees with the results of Biswas *et al* (1973) to at least 14 figures.

The first thing to note is that, although the estimates worsen with increasing λ, they never actually diverge as would occur with a naive application of perturbation theory. The glitch around $N = 10$ appears to be a characteristic of the model rather than a characteristic peculiar to this approximation, a similar effect also being found in the sequence of Rayleigh–Ritz upper bounds obtained from matrices similar to (3.8).

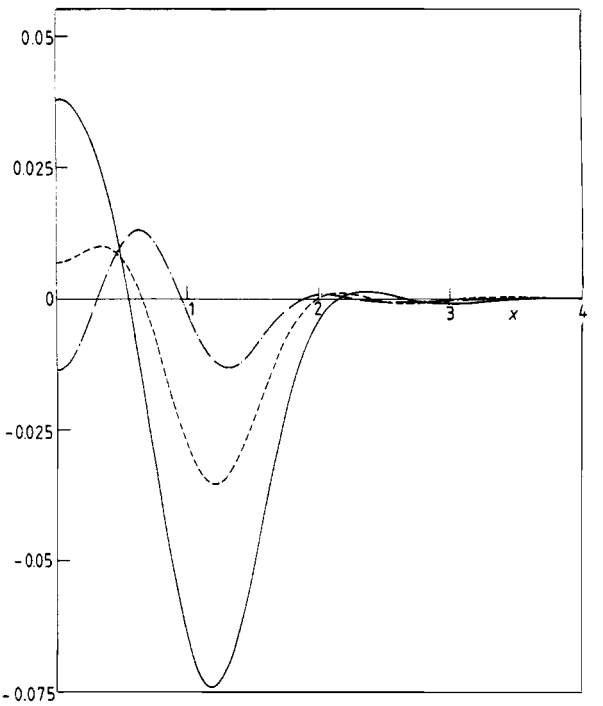

Figure 5. Error difference between the approximate Singh wavefunction and an accurate Rayleigh–Ritz wavefunction (see text) for ($n = 4$, $N = 6$) (——), ($n = 5$, $N = 7$) (- - - - -) and ($n = 6$, $N = 8$) (– · – ·) calculations.

In figure 5 the differences between the wavefunctions given by (2.48) and the 60-parameter ground state of (3.8) are plotted in configuration space. The comparison is for a few small values of N at $\lambda = 5$. The first plot shows how ψ_2 dominates the error in the four-parameter wavefunction generated by an $N = 6$ calculation. In fact ψ_2 dominates the wavefunction errors of the $N = 4, 5, 6$ calculations. The second of the plots shows significant contributions by ψ_3 to the errors of the $N = 7$ calculation whilst the last plot shows ψ_3 quite suddenly dominating the errors in the $N = 8$ (six-parameter wavefunction) calculation.

4. Numerical comparison with other methods

4.1. Bazley and Fox

Table 1 compares the results of the present method with those obtained by Bazley and Fox (1961) using the method of intermediate Hamiltonians in conjunction with special choice.

As explained in appendix 1, the Bazley–Fox method applied to the anharmonic oscillator requires the same input information as the Singh scheme, that is the first n columns of the matrix

$$\langle E_0^i | V | E_0^j \rangle.$$

Since this is a band matrix it is sufficient to consider an $N \times n$ matrix V_{Nn}. The $n \times n$ matrix formed by the first n rows of $V_{Nn}(V_{nn})$ must be inverted and the matrix product

$$\lambda (V_{Nn}(V_{nn})^{-1} V_{Nn}^{\dagger}) \tag{4.1}$$

calculated. Here † denotes the conjugate transpose. To the $N \times N$ matrix of (4.1) is added the diagonal matrix of H_0 to form an intermediate Hamiltonian which is a lower bound to (3.1). In the present case the non-diagonal part of this intermediate Hamiltonian is an $N \times N$ matrix which differs from the matrix

$$\langle E_0^i | H | E_0^j \rangle, \qquad i, j = 0, 1, \ldots, N-1, \tag{4.2}$$

by a 2×2 subarray of the matrices in the bottom right-hand corner of these arrays. The computational effort involved in finding this 2×2 matrix is comparable to the

Table 1. Comparison of Bazley and Fox and Singh lower bounds for the first three even parity levels of the anharmonic oscillator (3.1) using $N = 5$ unperturbed eigenstates.

Quantum number $= 2j$	Unbiased E_u	λ	Lower bounds		Upper bound (exact to no of figures quoted)
			Bazley and Fox	Singh	
0	4.99	0.2	1.118 255	1.116 848	1.118 292 6544
		0.4	1.204 738	1.200 829	1.204 810 3274
		0.6	1.275 773	1.269 843	1.275 983 5663
		0.8	1.336 760	1.328 079	1.337 545 2081
		1.0	1.390 301	1.377 591	1.392 351 642
	12.99	1.0	1.390 301	1.389 082	1.392 351 642
2	8.99	0.2	6.260 404	6.230 901	6.277 248 617
		0.4	6.979 830	6.886 007	7.072 598 726
		0.6	7.505 763	7.357 396	7.689 565 295
		0.8	7.942 661	7.732 751	8.205 677 392
		1.0	8.330 586	8.026 231	8.655 049 958
	12.99	1.0	8.330 586	8.327 684	8.655 049 958
4	12.99	0.2	12.225 85	12.075 909	12.440 601 8
		0.4	14.030 37	12.719 466	14.368 912 59
		0.6	14.906 30		15.823 505 46
		0.8	15.344 32	$\geq 12.99^{\dagger}$	17.022 827 08
		1.0	15.629 53		18.057 557 44

† The fixed point was too close to 12.99 to be clearly distinguishable. E_u cannot be increased significantly here without increasing N.

effort involved in solving the set of equations (2.43) and calculating $\mathscr{A}_n^{ij}(z)$ for $J_0 = 2$. Computer central memory requirements are of primary importance in this type of calculation and from this point of view it is the solution of these large sets of linear equations that causes the major problems in implementing all three methods compared in this paper.

If only a few low-lying eigenvalue estimates are required, the Bazley–Fox intermediate Hamiltonian eigenvalue problem could be handled efficiently with a Lanczos algorithm. This allows a rough comparison to be made between the calculations involved in finding a few eigenvalues of the Bazley–Fox intermediate Hamiltonian and the Gram–Schmidt orthogonalisation used to find $\{|\phi^k\rangle\}_{k=1}^n$.

It should be noted that the Gram–Schmidt process is not entirely necessary to the success of the Bubnov–Galerkin approximation (Mikhlin 1964), but it does reduce the time and storage required to set up the equations (2.43) for each new value of z. Since the equations (2.43) will generally be set up often in the search for z^* the present choice of $\{|\phi^k\rangle\}_{k=1}^n$ would seem reasonable, despite the initial overhead involved in the calculation of the $|\phi^k\rangle$ and the matrix elements of A between them.

The above discussion shows, at least on this model, that the computations involved in the Bazley–Fox method, implemented as described, are roughly equivalent to a subset of the computations involved in the Singh approach. This implies that the Bazley–Fox method is not only more accurate than the Singh scheme, as seen in table 1, but can also be arranged to require less computational effort.

4.2. Löwdin/Reid

Tables 2 and 3 are a comparison of the Singh scheme with an approximate reduced resolvent method of Löwdin as applied to the anharmonic oscillator by Reid (1965). The approximation itself is accomplished by applying the exact reduced resolvent formalism to an intermediate Hamiltonian obtained from (3.1) by doing an inner projection on V. This turns out to be the same intermediate Hamiltonian as that used by Bazley and Fox (see appendix 2). A consequence of this is the fact that the fixed points of the bracketing function obtained from the intermediate Hamiltonian are lower bounds to the fixed points of the bracketing function obtained from the Hamiltonian (3.1). The scheme then proceeds by presupposing an upper bound to the eigenvalue required and using this upper bound to construct a lower bound by evaluating the bracketing function of the intermediate Hamiltonian at the upper bound. The value of this approximate bracketing function at the upper bound is a lower bound to the fixed point of the approximate bracketing function and consequently a lower bound to the corresponding eigenvalue of (3.1). We note that this also means that the lower bound so obtained will always be worse than the lower bound obtained from the intermediate Hamiltonian.

The approach also harbours another problem: it is possible for the eigenvalue of the intermediate Hamiltonian to be such a bad approximation to the desired eigenvalue of the original Hamiltonian that the upper bound to the desired eigenvalue could lie very close to a singularity of the approximate bracketing function. Indeed, it is imaginable that the approximate bracketing function evaluation at the upper bound could pick up the wrong branch of the approximate bracketing function entirely (Löwdin 1965b). This is the cause of the peculiar Löwdin/Reid estimates for the $j = 2$ (quantum number = 4) estimates in the $N = 7$ calculations in table 2. No such eigenvalue ordering problem is encountered in the Singh scheme since the excited

Table 2. Comparison of Löwdin/Reid and Singh lower bounds for the first three even parity levels of the anharmonic oscillator (3.1) using $N = 7$ unperturbed eigenstates. These bounds should be compared with the exact upper bound results in table 1.

Quantum number $= 2j$	Unbiased E_u	λ	Lower bounds	
			Löwdin/Reid	Singh
0	4.99	0.2	1.118 292	1.118 284
		0.4	1.204 791	1.204 350
		0.6	1.275 909	1.273 727
		0.8	1.337 397	1.332 221
		1.0	1.392 131	1.383 341
	20.99	1.0	1.392 131	1.392 171
2	8.99	0.2	6.276 978	6.267 420
		0.4	7.070 812	7.045 157
		0.6	7.677 753	7.580 353
		0.8	8.163 597	7.906 090
		1.0	8.552 989	8.120 262
	20.99	1.0	8.552 989	8.595 571
4	12.99	0.2	12.370 33	12.141 767
	16.99	0.4	13.034 27	14.109 470
		0.6	8.879 68	15.323 326
	20.99	0.8	7.589 51	16.667 507
		1.0	3.354 64	17.615 728

Table 3. Comparison of Löwdin/Reid and Singh lower bounds for the anharmonic oscillator (3.1) using $N = 22$ unperturbed eigenstates. The bracketed figures are those that differ from a 20-dimensional Rayleigh–Ritz upper bound. This upper bound is presumably much more accurate than these lower bounds.

Quantum number $= 2j$	Unbiased E_u	λ	Lower bounds	
			Löwdin/Reid	Singh
0	4.99	0.25	1.141 901 839 539 14(98)	1.141 901 839 539 1(2)
		0.5	1.241 854 059 651 (94)	1.241 854 059 6(22)
		0.75	1.322 872 581 46(9 0)	1.322 872 58(1 29)
		1.0	1.392 351 641 (545)	1.392 351 6(39 1)
	40.99	1.0	1.392 351 641 (545)	1.392 351 641 (35)
2	8.99	0.25	6.500 905 725 74(33)	6.500 905 725 7(029)
		0.5	7.396 900 638 (80)	7.396 900 6(26 19)
		0.75	8.083 870 9(20 5)	8.083 870 (211 6)
		1.0	8.655 049 9(69)	8.655 04(6 132)
	40.99	1.0	8.655 049 9(69)	8.655 049 9(33)
4	16.99	0.25	12.991 025 904 8(83)	12.991 025 90(3 467)
		0.5	15.136 845 7(57)	15.136 845 (251)
		0.75	16.740 9(40 7)	16.740 (747 9)
	20.99	1.0	18.057 5(59)	18.057 (417)
	40.99	1.0	18.057 5(59)	18.057 5(44)

state estimates are obtained by enlarging the reference space Image(P_0) rather than attempting to pin down more fixed points of a single α function:

Computationally the Löwdin method has similar central memory requirements to the other two schemes. It involves the solution of an $n \times n$ set of equations comparable to the set of equations (2.43) with $J_0 = 1$. However, the value of the bracketing function is required at one point only so that such a set of equations need only be set up and solved once. Thus no special basis, like $\{|\phi^k\rangle\}_{k=1}^n$, is needed. This makes the Löwdin scheme faster than the Singh approach. There is, however, the problem of calculating an upper bound to the desired eigenvalue of (3.1). The tighter this upper bound, the better the Löwdin lower bound obtained. This is in direct contrast to the Singh scheme, wherein the worse the effective upper bound E_u, the better the lower bound. If we do something which is perhaps a little unfair to the Löwdin scheme, that is include an $N \times N$ eigenvalue calculation for the upper bound in the Löwdin scheme as part of the calculation, then the Löwdin scheme's computational difficulty approaches more closely that of the Singh approach for $J_0 = 1$.

From this it would appear that the Löwdin scheme is more efficient in terms of accuracy per unit calculation than the Singh method. On the other hand, the Singh approach suffers no eigenvalue ordering problems and yields, by itself, an approximate wavefunction. Further, by increasing E_u it is possible to make Singh's method more accurate than the Löwdin scheme though, we suspect, the Singh estimates will always be less accurate than the corresponding Bazley–Fox estimates, at least for the anharmonic oscillator.

5. Discussion and assessment

The formal manipulations leading to the eigenvalue equation (2.10) are common in the literature (see e.g. Reid 1976 and references therein). The Hamiltonian is decomposed as a sum

$$H = H_1 - H_2 \tag{5.1}$$

and the eigenvalue equation for H rearranged into a generalised eigenvalue problem:

$$H_2 G(E) H_2 |E\rangle = H_2 |E\rangle \tag{5.2}$$

where

$$G(z) = (H_1 - z)^{-1}. \tag{5.3}$$

This problem can in turn be viewed (see e.g. Abdel-Raouf 1982) as the extremisation of $\langle\phi|H_2 G H_2|\phi\rangle$ subject to $\langle\phi|H_2|\phi\rangle$ being constant. If this constraint is introduced into the objective function, along with a Lagrange multiplier $\beta(z)$, (5.2) is replaced by

$$H_2 G(z) H_2 |z\rangle = \beta(z) H_2 |z\rangle \tag{5.4}$$

for $z \neq E$. The eigenvalues of H correspond to $\beta(z = E) = 1$ and the formal similarity with Singh's formulation is evident.

Usually H_2 is taken as the potential term V and G the Green function of the diagonal, or kinetic, part H_0 (Hall *et al* 1969, 1970). Singh, however, has made a novel choice by using part of H_0 for H_2. Depending on the gradient of the $\beta(z)$, lower bounds can be obtained by bounding $\beta(z)$ from above or below. In Singh's case $d\beta/dz > 0$ so that an upper bound to $H_2 G H_2$ is required. If H_2 is definite the

generalised eigenvalues have a variational characterisation. In fact, the replacement $|\beta(z)\rangle = H_2^{1/2}|z\rangle$ can be used to construct the eigenvalue equation

$$H_2^{1/2}G(z)H_2^{1/2}|\beta(z)\rangle = \beta(z)|\beta(z)\rangle. \tag{5.5}$$

The problem is that $H_2^{1/2}G(z)H_2^{1/2}$ cannot, in general, be calculated, even on a subspace of the domain of $H_2^{1/2}$. This removes the possibility of using a direct Rayleigh–Ritz variational upper bound to $\beta(z)$. The difficulty can be attributed to the use of the 'complicated part' of the Hamiltonian H in the Green function $G(z)$. Singh, nonetheless, ingeniously constructs upper bounds to the $\beta(z)$ using the method reproduced in § 2.2. The significance of this bounding procedure is that it directly addresses the problem of approximating a seemingly intractable operator $H_2^{1/2}GH_2^{1/2}$. Furthermore the approximation is controlled. Clearly the utility of similar techniques in approximating $G(z)$ arising out of other choices for H_2 and H_1 could well bear investigation.

We have mentioned that the Löwdin approximation quoted in this article is obtained by applying the exact reduced resolvent formalism to the intermediate Hamiltonians of Bazley and Fox. In Löwdin's own opinion this is not a good approach (Löwdin 1965b, Reid 1976) but is excusable since the actual fixed point of the bracketing function is not sought.

This naturally, however, raises the question as to whether the Singh approximation scheme (§ 2.2) is equivalent to applying the *exact* formalism (§ 2.1) to an intermediate Hamiltonian. To explore this question set

$$H_2^{1/2} = (-H_0)^{1/2}P_0, \tag{5.6}$$

$$G(z) = (\bar{H} - z)^{-1}, \tag{5.7}$$

so that

$$\mathscr{B}(z) = H_2^{1/2}G(z)H_2^{1/2}. \tag{5.8}$$

The operator $\bar{\nu}_n$ approximating $\bar{\nu}$ in (2.28) can be rewritten

$$\bar{\nu}_n = \sum_{k,l=1}^{n} (\bar{H} - E_u)|\phi^k\rangle[\langle\phi^k|(\bar{H} - E_u)(\bar{H} - z)|\phi^l\rangle]^{-1}\langle\phi^l|(\bar{H} - E_u) \tag{5.9}$$

by using (2.32), (2.34), (2.35) and (2.38).

To clarify (5.9) let

$$|\chi^k\rangle = (\bar{H} - E_u)^{1/2}(\bar{H} - z)^{1/2}|\phi^k\rangle \tag{5.10}$$

and define

$$\bar{P}_n = \sum_{k,l=1}^{n} |\chi^k\rangle[\langle\chi^k|\chi^l\rangle]^{-1}\langle\chi^l| \tag{5.11}$$

so that $\bar{P}_n^2 = \bar{P}_n$. Equation (5.9) then becomes

$$\bar{\nu}_n = \bar{\nu}^{1/2}\bar{P}_n\bar{\nu}^{1/2}, \tag{5.12}$$

clearly showing the projective nature of the Bubnov–Galerkin method as well as the lower bounding property of

$$\mathscr{A}_n = \mathscr{H}_0 + H_2^{1/2}\bar{\nu}_n H_2^{1/2}. \tag{5.13}$$

To continue the investigation we turn to the Aronszajn form of (5.12). Let

$$|\eta^k\rangle = (\bar{H} - z)|\phi^k\rangle \tag{5.14}$$

and define the Aronszajn projections

$$P_n^{\nu^+} = \sum_{k,l=1}^{n} |\eta^k\rangle[\langle\eta^k|\bar{\nu}|\eta^l\rangle]^{-1}\langle\eta^l|\bar{\nu}, \tag{5.15a}$$

$$P_n^{\nu} = \sum_{k,l=1} \bar{\nu}|\eta^k\rangle[\langle\eta^k|\bar{\nu}|\eta^l\rangle]^{-1}\langle\eta^l|, \tag{5.15b}$$

together with

$$O_n^{\nu} = 1 - P_n^{\nu}, \qquad O_n^{\nu^+} = 1 - P_n^{\nu^+}. \tag{5.15c}$$

Hence

$$\bar{\nu}_n = \bar{\nu}P_n^{\nu^+} = P_n^{\nu}\bar{\nu} = P_n^{\nu}\bar{\nu}P_n^{\nu^+}. \tag{5.16}$$

We can recover $\mathscr{B}_n(z)$ from $\mathscr{A}_n(z)$ by reversing the derivation of $\mathscr{A}(z)$ from $\mathscr{B}(z)$ in § 2.1. Now

$$\mathscr{A}_n(z) = E_u P_0 + H_2^{1/2}(-1 + \bar{\nu}_n)H_2^{1/2} \tag{5.17}$$

so that a factor of $z - E_u$ needs to be pulled out of

$$-1 + \bar{\nu}_n = (z - E_u)G_n(z) \tag{5.18}$$

which also serves to define $G_n(z)$. Expanding (5.18) yields

$$\begin{aligned}(z - E_u)G_n(z) &= -1 + (\bar{H} - E_u)(\bar{H} - z)^{-1}P_n^{\nu^+} \\ &= (\bar{H} - z)^{-1}[(z - \bar{H})(P_n^{\nu^+} + O_n^{\nu^+}) + (\bar{H} - E_u)P_n^{\nu^+}] \\ &= (z - E_u)[(\bar{H} - z)^{-1}P_n^{\nu^+} + (E_u - z)^{-1}O_n^{\nu^+}],\end{aligned} \tag{5.19}$$

so that

$$G_n(z) = (\bar{H} - z)^{-1}\left(P_n^{\nu^+} + \frac{\bar{H} - z}{(E_u - z)}O_n^{\nu^+}\right)$$

$$= (\bar{H} - z)^{-1}\left(1 + \frac{\bar{H} - E_u}{(E_u - z)}O_n^{\nu^+}\right) \tag{5.20a}$$

$$= (\bar{H} - z)^{-1} + (E_u - z)^{-1}O_n^{\nu}\bar{\nu}O_n^{\nu^+}. \tag{5.20b}$$

Now $E_u - z > 0$ and for arbitrary $|\xi\rangle$ we have

$$(E_u - z)^{-1}\langle\xi|O_n^{\nu}\bar{\nu}O_n^{\nu^+}|\xi\rangle = \langle\eta|\bar{\nu}|\eta\rangle(E_u - z)^{-1} \geqslant 0, \tag{5.21}$$

so $G_n(z) \geqslant G(z)$ for $z < E_u$ as expected.

Manipulations such as these using projections like P_n^{ν} and $\bar{P}_n$ can be found in Gay (1964), Bazley and Fox (1966) or Wilson (1965). These all yield expressions like (5.20) for $G(z)$. However, we have *not* been able to reproduce a simple intermediate Hamiltonian from the various expressions for $G_n(z)$ we have tried.

On the other hand, (5.20*b*) does make a start toward an error analysis for the method and confirms the observed E_u dependence of the estimates.

In addition, this equation also confirms the convergence property of the estimates as well as the singularity structure of the eigenvalues $\beta_n(z)$ of

$$\mathcal{B}_n(z) = H_2^{1/2} G_n(z) H_2^{1/2} \tag{5.22}$$

discovered by Singh.

To finish we display the form that the original Schrödinger equation takes after the Singh approximation has been applied. Let z_n^* be such that $\beta_n(z_n^*) = 1$ and define $|z\rangle_n$ according to (2.48), that is

$$|z\rangle_n = G_n(z) H_2^{1/2} |\beta_n(z)\rangle. \tag{5.23}$$

Substituting (5.20*a*) yields, after some algebra,

$$\begin{aligned}(H - z)|z\rangle_n &= (\bar{H} - H_2 - z)(\bar{H} - z)^{-1}[1 + (\bar{H} - E_u)(E_u - z)^{-1} O_n^{\nu^+}] H_2^{1/2} |\beta_n(z)\rangle \\ &= [1 - \beta_n(z) + (E_u - z)^{-1}(\bar{H} - E_u) O_n^{\nu^+}] H_2^{1/2} |\beta_n(z)\rangle,\end{aligned} \tag{5.24}$$

so that finally

$$\begin{aligned}(H - z_n^*)|z_n^*\rangle_n &= (z_n^* - E_u)^{-1}(\bar{H} - E_u) O_n^{\nu^+} P_0 H_0^{1/2} P_0 |\beta_n(z_n^*)\rangle \\ &= (z_n^* - E_u)^{-1}(\bar{H} - E_u) O^{\nu^+} P_0 H_0 P_0 |z_n^*\rangle_n.\end{aligned} \tag{5.25}$$

In part this final section has attempted to relate the Singh approximation to other developments. It appears that the Singh scheme is a distinct approximation with some similarities to other methods. It is similar in assumptions to the Bazley–Fox method in that it requires special choice to make the Bubnov–Galerkin approximation tractable, and it has a superficial similarity to the Löwdin method (see in particular the discussion in Wilson (1967)).

In addition the β functions are, in the $J_0 \to \infty$ limit, directly related to a renormalisation of configuration space (Hall *et al* 1969, 1970). In this context the β functions are regarded as variational parameters which are used to minimise the Rayleigh–Ritz quotient of some trial wavefunction which undergoes a renormalisation of configuration space as β varies. In the present case the Rayleigh–Ritz quotient obtained from (2.14) is

$$\langle z|H|z\rangle / \langle z|z\rangle = z + \beta(z)(1 - \beta(z))/\beta'(z) \tag{5.26}$$

where (2.13), (2.14), (2.15) and (2.18) have been used.

Numerically the results presented in §§ 3 and 4 show that the procedure avoids the eigenvalue ordering problem that can arise in the Löwdin technique used by Reid. However, the results are less accurate and involve a greater degree of computational effort than the classical Bazley and Fox approximations.

Acknowledgment

One of us (PGH) is grateful to the Australian Department of Education for the award of a Commonwealth Postgraduate Scholarship.

Appendix 1. Bazley/Fox bound

Let $\{|p_i\rangle\}_{i=0}^{n-1}$ define a submanifold of the Hilbert space of the system in question and define

$$P_n^{v^+} = \sum_{i,j=0}^{n-1} |p_i\rangle(\langle p_i|V|p_j\rangle)^{-1}\langle p_j|V \tag{A1.1}$$

together with

$$P_n^{v} = \sum_{i,j=0}^{n-1} V|p_i\rangle(\langle p_i|V|p_i\rangle)^{-1}\langle p_j|; \tag{A1.2}$$

then

$$H_N = H_0 + VP_n^{v^+} \tag{A1.3}$$

will be an operator lower bound to H whenever $V \geqslant 0$. Special choice then searches for $|p_i\rangle$ such that

$$V|p_i\rangle = \sum_{l=0}^{N-1} \beta_{il}|E_0^l\rangle, \qquad N \text{ finite.} \tag{A1.4}$$

The band structure of the potential term displayed in (3.6) allows the choice

$$|p_i\rangle = |E_0^i\rangle. \tag{A1.5}$$

This means $N = n+2$ in this case so that β_{il} is essentially the $n \times N$ matrix V_{nN}^{+} mentioned in equation (4.1). Thus

$$\langle E_0^i|VP_n^{v^+}|E_0^j\rangle = \begin{cases} \sum_{k,l=0}^{n-1} \beta_{ki}^{*}[\beta_{kl}]^{-1}\beta_{lj}, & i,j < N, \\ 0 & \text{otherwise,} \end{cases} \tag{A1.6}$$

from which the discussion in § 4.1 follows.

Appendix 2. Approximate resolvent approximation for bracketing function

As defined in (2.51) the bracketing function is not a practical representation. The particular approximation used by Reid (1965) to obtain a tractable problem involves setting $J_0 = 1$ and replacing H in (2.49) and (2.51) with H_N defined by (A1.1), (A1.3) and (A1.5). This defines a tractable approximate bracketing function $f_N(z)$ which is exact on H_N. The actual expressions for $f_N(z)$ are

$$f_N(z) = \langle \phi^0|H_0(1+T_0H_0)|\phi^0\rangle + \boldsymbol{a}^{+}D^{-1}\boldsymbol{a} \tag{A2.1}$$

where

$$T_0 = O(z - OH_0O)^{-1}O, \qquad |\phi^0\rangle = |E_0^0\rangle, \tag{A2.2}$$

and O are defined by (2.50).

The D and $\boldsymbol{a}$ are given in terms of the $\{|p_i\rangle\}_{i=0}^{n-1}$ appearing in equations (A1.1) and (A1.2). They are

$$a_j = \langle p_j|V(1+T_0H_0)|\phi^0\rangle \tag{A2.3}$$

and

$$D_{ij} = \langle p_i | V(1 - T_0 V) | p_j \rangle. \tag{A2.4}$$

Once again the special choice (A1.5) can be made so that the basic input information must be sufficient to calculate

$$\langle E_0^i | VT_0 V | E_0^j \rangle = \langle V\psi_i | T_0 | V\psi_j \rangle, \qquad i, j = 0, \ldots, n-1. \tag{A2.5}$$

This again requires knowledge of $N = n + 2$ eigenstates of H_0 and the matrix elements in the top left-hand $N \times n$ subarray of the matrix of V.

References

Abdel-Raouf M A 1981 *J. Math. Phys.* **22** 2849
—— 1982 *Phys. Rep.* **84** 165
Bazley N W and Fox D W 1961 *Phys. Rev.* **124** 483
—— 1966 *Phys. Rev.* **148** 90
Biswas S N, Datta K, Saxena R P, Srivastava P K and Varma V S 1973 *J. Math. Phys.* **14** 1190
Blatt J M 1975 *Aust. Comp. J.* **7** 51
Gay J G 1964 *Phys. Rev.* **135** A 1220
Hall G G, Hyslop J and Rees D 1969 *Int. J. Quant. Chem.* **3** 195
—— 1970 *Int J. Quant. Chem.* **4** 5
Hill R N 1980 *J. Math. Phys.* **21** 2182
Löwdin P-O 1962 *J. Math. Phys.* **3** 969
—— 1965a *Phys. Rev.* **139** A 357
—— 1965b *J. Chem. Phys.* **43** suppl S175
Mikhlin S G 1964 *Variational Methods in Mathematical Physics* (Oxford: Pergamon)
Reid C E 1965 *J. Chem. Phys.* **43** suppl S186
—— 1976 *Quantum Science Methods and Structure—A tribute to Per-Olov Lowdin* ed L Colois, O Goscinski, J Linderberg and Y Ohrn (New York: Plenum) pp 315–47
Schiff L I 1968 *Quantum Mechanics* (New York: McGraw-Hill)
Singh S R 1977 *J. Math. Phys.* **18** 1466
—— 1981 *J. Math. Phys.* **22** 893
Temple G 1928 *Proc. R. Soc.* A **119** 276
Wilson E B Jr 1965 *J. Chem. Phys.* **43** suppl S172
Wilson T M 1967 *J. Chem. Phys.* **47** 3912