Variational bounds on overlaps (in Hilbert space)

1975 J. Phys. A: Math. Gen. 8 654

(http://iopscience.iop.org/0305-4470/8/5/003)

J. Phys. A: Math. Gen., Vol. 8, No. 5, 1975. Printed in Great Britain.

Variational bounds on overlaps

M F Barnsley and P D Robinson
School of Mathematics, Bradford University, Bradford BD7 1DP, UK

Received 24 October 1974

Abstract. A general formalism is presented which furnishes upper and lower variational bounds on squared overlaps $|\langle\theta, \chi\rangle|^2$, where χ and θ are vectors in a Hilbert space. The bounding functionals involve a self-adjoint operator A and a second operator B, such that $A\theta = 0$ and $B\chi = b\chi$. Most of the non-variational bounds derived by other workers on quantum-mechanical overlaps emerge from the formalism, together with optimal variational improvements in every case. In addition many new bounds are obtained.

1. Introduction

In this paper we present a general theory of variational bounds on squared overlaps $|\langle\theta, \chi\rangle|^2$, where χ and θ are vectors in a Hilbert space. The method relies on the availability of a self-adjoint operator A such that $A\theta = 0$ and an operator B such that $B\chi = b\chi$, where b is a nonzero (possibly complex) number. In particular the theory applies directly to quantum mechanics with θ an atomic or molecular eigenfunction and χ some given function which is usually an approximation to θ. The squared overlap is an obvious measure of the accuracy of approximate eigenfunctions and is also used a great deal in estimating quantum-mechanical properties (for example, Weinhold 1970).

In § 2 we present a general theory in terms of the operators A and B. Ultimately the bounds are written in the form

$$f(A, B) - \hat{Y}(A, B; \Phi) \geqslant |\langle\theta, \chi\rangle|^2 \geqslant g(A, B) + \hat{Z}(A, B; \Phi),$$

where $f(A, B)$ and $g(A, B)$ are non-variational bounds, with variational improvements supplied by $\hat{Y}(A, B; \Phi)$ and $\hat{Z}(A, B; \Phi)$, Φ being the trial vector.

In §§ 3 and 4 we particularize to quantum mechanics and consider a variety of choices for A and B. In this way many of the non-variational bounds derived by other workers emerge from our theory as examples of f and g, and we are led immediately to corresponding variational improvements. We furthermore obtain a variety of new bounds.

In § 5 we briefly discuss the feasibility of evaluating the variational functionals and also consider a class of bounds due to Weinhold which at first sight appear to lie outside the reach of the present theory.

2. General theory

Let A be a self-adjoint operator in a Hilbert space $\mathscr{H}$ with inner product $\langle , \rangle$. Let the eigenvalue spectrum of A be partly or wholly discrete and include zero as an isolated, non-degenerate member. Suppose that the eigenvector θ specified by

$$A\theta = 0, \qquad \theta \in \mathscr{H}, \qquad \langle\theta, \theta\rangle = 1 \tag{2.1}$$

is unknown and that we seek bounds on $|\langle\theta, \chi\rangle|^2$ for an arbitrary, known vector $\chi \in \mathscr{H}$.

If l is the eigenvalue (positive or negative) of A which is closest to the zero one and if h is the subspace of $\mathscr{H}$ which is orthogonal to θ, then

$$\langle\xi, A^2\xi\rangle \geqslant l^2\langle\xi, \xi\rangle \qquad \text{for all } \xi \in h. \tag{2.2}$$

Further we suppose that

$$u^2\langle\xi, \xi\rangle \geqslant \langle\xi, A^2\xi\rangle, \tag{2.3}$$

although in many applications the number u^2 will be infinite. These inequalities are the basis of complementary (upper and lower) variational bounds on $|\langle\theta, \chi\rangle|^2$.

To derive these bounds directly we consider the variation of the two functionals

$$F(\Phi) = \left\langle (A\Phi-\chi), \left(1-\frac{A^2}{u^2}\right)(A\Phi-\chi)\right\rangle, \tag{2.4}$$

$$\Phi \in \mathscr{H},\ \chi \in \mathscr{H},$$

$$G(\Phi) = \left\langle (A\Phi-\chi), \left(1-\frac{A^2}{l^2}\right)(A\Phi-\chi)\right\rangle, \tag{2.5}$$

around their common stationary value achieved when Φ is the solution vector ϕ of the equation in h:

$$A\phi = Q\chi = \chi - \theta\langle\theta, \chi\rangle, \tag{2.6}$$

Q being the projection operator

$$Q = 1 - |\theta\rangle\langle\theta|. \tag{2.7}$$

With the help of (2.1) and (2.6) we find that

$$F(\Phi) = F(\phi) + \left\langle A(\Phi-\phi), \left(1-\frac{A^2}{u^2}\right)A(\Phi-\phi)\right\rangle \tag{2.8}$$

and

$$G(\Phi) = G(\phi) - \left\langle A(\Phi-\phi), \left(\frac{A^2}{l^2}-1\right)A(\Phi-\phi)\right\rangle \tag{2.9}$$

where

$$F(\phi) = |\langle\theta, \chi\rangle|^2 = G(\phi). \tag{2.10}$$

Accordingly from (2.2) and (2.3) it follows that since $A(\Phi-\phi) \in h$, then

$$F(\Phi) \geqslant |\langle\theta, \chi\rangle|^2 \geqslant G(\Phi). \tag{2.11}$$

These complementary bounds can also be derived in what is perhaps a more logical manner by taking as a starting point dual variational principles associated with the linear equation

$$A^2\phi = AQ\chi, \tag{2.12}$$

with the operator A^2 decomposed in two different ways (cf Arthurs 1970, Barnsley and Robinson 1974).

It is desirable to make a further generalization in (2.11). Let us replace the vector χ by $B\chi$ where B is a linear operator in $\mathscr{H}$ with eigenvector θ satisfying

$$B\theta = b\theta \qquad b \neq 0 \tag{2.13}$$

and

$$|\langle\theta, B\chi\rangle|^2 = |b|^2|\langle\theta, \chi\rangle|^2 \qquad \chi \in \mathscr{H}. \tag{2.14}$$

Condition (2.14) is met if B is self-adjoint and b is real, and also less trivially if B is not self-adjoint but satisfies

$$\langle\theta, B\xi\rangle = 0 \qquad \text{for all } \xi \in h, \tag{2.15}$$

as can be seen if we decompose χ in the form

$$\chi = \theta\langle\theta, \chi\rangle + \xi. \tag{2.16}$$

The complementary variational bounds now become

$$F(A, B; \Phi) \geqslant |\langle\theta, \chi\rangle|^2 \geqslant G(A, B; \Phi) \tag{2.17}$$

with

$$F(A, B; \Phi) = \frac{1}{|b|^2}\left\langle(A\Phi - B\chi), \left(1 - \frac{A^2}{u^2}\right)(A\Phi - B\chi)\right\rangle \tag{2.18}$$

and

$$G(A, B; \Phi) = \frac{1}{|b|^2}\left\langle(A\Phi - B\chi), \left(1 - \frac{A^2}{l^2}\right)(A\Phi - B\chi)\right\rangle. \tag{2.19}$$

It is evident that whole families of bounds can now be generated by appropriate choices of the operators A and B. Certain choices lead to variational bounds which are already known. Furthermore, most of the *non-variational* bounds cited in the literature can be obtained by taking Φ as the zero vector. Calling such bounds f and g we have

$$f(A, B) = F(A, B; 0) > |\langle\theta, \chi\rangle|^2 > G(A, B; 0) = g(A, B) \tag{2.20}$$

where

$$f(A, B) = \left\langle\frac{B}{b}\chi, \left(1 - \frac{A^2}{u^2}\right)\frac{B}{b}\chi\right\rangle \tag{2.21}$$

and

$$g(A, B) = \left\langle\frac{B}{b}\chi, \left(1 - \frac{A^2}{l^2}\right)\frac{B}{b}\chi\right\rangle. \tag{2.22}$$

These bounds do not in general equal $|\langle\theta, \chi\rangle|^2$. However, the functionals $F(A, B; \Phi)$ and $G(A, B; \Phi)$ can always attain their common stationary value $|\langle\theta, \chi\rangle|^2$ and hence can furnish variational improvements to the cruder bounds $f(A, B)$ and $g(A, B)$. We rewrite (2.17) as

$$\begin{aligned} F(A, B; \Phi) = f(A, B) - Y(A, B; \Phi) \geqslant |\langle\theta, \chi\rangle|^2 &\geqslant g(A, B) + Z(A, B; \Phi) \\ &= G(A, B; \Phi) \end{aligned} \tag{2.23}$$

where

$$Y(A,B;\Phi) = \frac{1}{|b|^2}\left[-\left\langle A\Phi,\left(1-\frac{A^2}{u^2}\right)A\Phi\right\rangle + \left\langle B\chi,\left(1-\frac{A^2}{u^2}\right)A\Phi\right\rangle + \left\langle A\Phi,\left(1-\frac{A^2}{u^2}\right)B\chi\right\rangle\right] \tag{2.24}$$

and

$$Z(A,B;\Phi) = \frac{1}{|b|^2}\left[-\left\langle A\Phi,\left(\frac{A^2}{l^2}-1\right)A\Phi\right\rangle + \left\langle B\chi,\left(\frac{A^2}{l^2}-1\right)A\Phi\right\rangle + \left\langle A\Phi,\left(\frac{A^2}{l^2}-1\right)B\chi\right\rangle\right]. \tag{2.25}$$

The optimized versions of $Y(A, B; c\Phi)$ and $Z(A, B; c\Phi)$ with respect to a parameter c are

$$\hat{Y}(A,B;\Phi) = \left|\left\langle A\Phi,\left(1-\frac{A^2}{u^2}\right)\frac{B}{b}\chi\right\rangle\right|^2 \left\langle A\Phi,\left(1-\frac{A^2}{u^2}\right)A\Phi\right\rangle^{-1} \tag{2.26}$$

and

$$\hat{Z}(A,B;\Phi) = \left|\left\langle A\Phi,\left(\frac{A^2}{l^2}-1\right)\frac{B}{b}\chi\right\rangle\right|^2 \left\langle A\Phi,\left(\frac{A^2}{l^2}-1\right)A\Phi\right\rangle^{-1}. \tag{2.27}$$

The latter functionals are always non-negative and hence lead to improvements on the bounds $f(A, B)$ and $g(A, B)$. We denote such optimized bounds by

$$\hat{F}(A,B;\Phi) = f(A,B) - \hat{Y}(A,B;\Phi) \tag{2.28}$$

and

$$\hat{G}(A,B;\Phi) = g(A,B) + \hat{Z}(A,B;\Phi). \tag{2.29}$$

More generally the functionals Y and Z can be optimized within a linear basis set of trial vectors, yielding standard generalizations of (2.26) and (2.27).

In the event that zero is a degenerate eigenvalue of A, so that the subspace of $\mathscr{H}$ corresponding to this zero eigenvalue is spanned by t orthonormal vectors $\theta^1, \theta^2, \ldots \theta^t$, then the foregoing theory supplies bounds on the quantity

$$\sum_{s=1}^{t} |\langle \theta^s, \chi\rangle|^2. \tag{2.30}$$

3. Overlap with ground state

3.1. Preliminaries

Suppose that a quantum-mechanical system is described by a Hamiltonian H with normalized eigenvectors $\{\theta_i\}$ and eigenvalues $\{E_i\}$ so that

$$H\theta_i = E_i\theta_i \qquad i = 0, 1, 2 \ldots. \tag{3.1}$$

Let us assume that the eigenvalues are non-degenerate and satisfy

$$E_0 < E_1 < E_2 < \ldots \tag{3.2}$$

(there is little loss of generality here in view of the remark preceding (2.30)).

In this section we set $\theta = \theta_0$ and discuss bounds on the square-overlap $|\langle\theta_0, \chi\rangle|^2$ of an arbitrary vector χ with the ground state θ_0.

The question arises as to how we should select the operators A and B, for which there are very many possibilities. A natural starting point is suggested by the sequence of lower bounds

$$|\langle\theta_0, \chi\rangle|^2 > \left\langle \chi, \left(\frac{E_1 - H}{E_1 - E_0}\right)\left(\frac{H-\alpha}{E_0-\alpha}\right)^{-n} \chi \right\rangle = g_n(\alpha) \qquad (\alpha < E_0) \tag{3.3}$$

of the type presented by Rayner (1962b). These bounds approach the exact square-overlap both as $n \to \infty$ and as $\alpha \to E_0$.

The Rayner-type bounds $g_n(\alpha)$ and generalizations thereof arise from equation (2.22) on choosing

$$A = (H - E_0)^{1/2} \qquad B = (H-\alpha)^{-n/2} \qquad n = \ldots -1, 0, 1, 2 \ldots \tag{3.4}$$

so that

$$l^2 = E_1 - E_0 \qquad u^2 = \infty \qquad |b|^2 = (E_0 - \alpha)^{-n}. \tag{3.5}$$

We note that there is nothing to be gained by generalizing the choice for A to

$$A = (H - E_0)^m \qquad m > 0, \tag{3.6}$$

with

$$l^2 = (E_1 - E_0)^{2m} \qquad u^2 = \infty. \tag{3.7}$$

To see this, consider the fundamental bounds (2.17) in the resulting form

$$\langle\Theta, \Theta\rangle \geqslant |\langle\theta_0, \chi\rangle|^2 \geqslant \left\langle \Theta, \left[1 - \left(\frac{H - E_0}{E_1 - E_0}\right)^{2m}\right]\Theta \right\rangle \tag{3.8}$$

with

$$\Theta = (A\Phi - B\chi)/b. \tag{3.9}$$

We may assume that $(A\Phi) \in h$ rather than $\Phi \in \mathscr{H}$ is the arbitrary vector in (3.9). It follows that, for a given choice of B, any upper bound derived from (3.8) does not depend on the value of m and lower bounds are improved by taking m as small as possible. In practice this indicates $m = \frac{1}{2}$.

3.2. *Lower bounds*

Substitution of (3.4) and (3.5) into (2.29) leads to optimized lower bounds of the form

$$|\langle\theta_0, \chi\rangle|^2 \geqslant g_n(\alpha) + Z_n(\alpha; \Phi) = G_n(\alpha; \Phi) \qquad n = \ldots -1, 0, +1 \ldots \tag{3.10}$$

where $g_n(\alpha)$ is the lower bound in (3.3) and

$$\hat{Z}_n(\alpha; \Phi) = \frac{(E_0 - \alpha)^n}{(E_1 - E_0)} \frac{|\langle (H-E_0)^{1/2}\Phi, (H-E_1)(H-\alpha)^{-n/2}\chi\rangle|^2}{\langle\Phi, (H-E_0)(H-E_1)\Phi\rangle}. \tag{3.11}$$

Optimal variational improvements to Rayner-type lower bounds are supplied by these functionals $\hat{Z}_n(\alpha; \Phi)$.

When $n \geqslant 0$ it can be convenient to make the substitution

$$\Phi = (H - E_0)^{1/2}(H-\alpha)^{n/2}\Psi \qquad \Psi \in \mathscr{H}, \tag{3.12}$$

so that

$$\hat{Z}_n(\alpha;\Phi) = \hat{Z}'_n(\alpha;\Psi) \qquad \text{(say)}$$
$$= \frac{(E_0-\alpha)^n}{(E_1-E_0)} \frac{|\langle\Psi,(H-E_0)(H-E_1)\chi\rangle|^2}{\langle\Psi,(H-E_0)^2(H-E_1)(H-\alpha)^n\Psi\rangle}. \tag{3.13}$$

A related substitution is available for negative n.

When $n = 0$, explicit reference to the parameter α disappears. The resulting lower bound

$$|\langle\theta_0,\chi\rangle|^2 \geqslant g_0(\alpha)+\hat{Z}'_0(\alpha;\Psi) \tag{3.14}$$

has previously been given by Wang (1969). This variational principle improves $g_0(\alpha)$, which is the simple Eckart bound (Eckart 1930). Further, as a special case it yields the moment-theoretical lower bounds derived by Gordon (1968) (see Wang 1969).

The choice $n = 1$ is of especial interest. It has been noted by Weinhold and Wang (1971) that the Braun–Rebane (1969) variational lower bound for $|\langle\theta_0,\chi\rangle|^2$ which is

$$R(\alpha;\Omega) = \frac{E_0-\alpha}{E_1-E_0}\left(\frac{(E_1-\alpha)|\langle\Omega,\chi\rangle|^2}{\langle\Omega,(H-\alpha)\Omega\rangle}-\langle\Omega,\Omega\rangle\right), \qquad \Omega\in\mathscr{H}, \tag{3.15}$$

is in fact stationary about $g_1(\alpha)$ so that

$$|\langle\theta_0,\chi\rangle|^2 > g_1(\alpha) \geqslant R(\alpha;\Omega). \tag{3.16}$$

The Braun–Rebane bound is obtained by exploiting the Schwarz inequality

$$\langle\chi,(H-\alpha)^{-1}\chi\rangle \geqslant |\langle\Omega,\chi\rangle|^2/\langle\Omega,(H-\alpha)\Omega\rangle, \tag{3.17}$$

and is not a member of the general family (2.19). However it follows from (3.16) that the optimal variational correction $\hat{Z}'_1(\alpha;\Psi)$ to $g_1(\alpha)$ is also the requisite correction to the Braun–Rebane bound; thus

$$|\langle\theta_0,\chi\rangle|^2 \geqslant R(\alpha;\Omega)+\hat{Z}'_1(\alpha;\Psi), \qquad \Omega,\Psi\in\mathscr{H}. \tag{3.18}$$

In this section we have assumed so far that the parameter α is real and less than E_0, in line with (3.3). However, results are still available for any value of α (including complex values) which differs from all E_i. The choice $(H-\alpha)^{-n/2}$ for B is still self-adjoint if α is real and n is even. For complex α or odd n the choice $(H-\alpha)^{-n/2}$ satisfies the weaker conditions (2.13) and (2.15). In these latter cases the expression in (3.3) for $g_n(\alpha)$ must be modified to read

$$g_n(\alpha) = \left\langle\chi, \left(\frac{E_1-H}{E_1-E_0}\right)\left|\frac{H-\alpha}{E_0-\alpha}\right|^{-n}\chi\right\rangle, \tag{3.19}$$

and appropriate modifications involving $|H-\alpha|$ and $|E_0-\alpha|$ made to (3.11)–(3.13). The Braun–Rebane result is no longer relevant.

3.3. Upper bounds

Bounds complementary to those in § 3.2 are obtained by substituting (3.4) and (3.5) into the general optimum upper bound (2.28). We obtain

$$|\langle\theta_0,\chi\rangle|^2 \leqslant f_n(\alpha)-\hat{Y}_n(\alpha;\Phi) \qquad n = \ldots -1,0,1\ldots \tag{3.20}$$

where

$$f_n(\alpha) = (E_0-\alpha)^n\langle\chi, (H-\alpha)^{-n}\chi\rangle \tag{3.21}$$

and

$$\hat{Y}(\alpha;\Phi) = (E_0-\alpha)^n|\langle(H-E_0)^{1/2}\Phi, (H-\alpha)^{-n/2}\chi\rangle|^2/\langle\Phi, (H-E_0)\Phi\rangle. \tag{3.22}$$

Here the sequence of upper bounds $\{f_n(\alpha)\}$ is the natural complement to the sequence of lower bounds $\{g_n(\alpha)\}$. The bound $f_n(\alpha)$ approaches the exact result both when $n \to \infty$ and when $\alpha \to E_0$. Variational improvements are supplied by the functionals $\hat{Y}_n(\alpha;\Phi)$. We note that the restriction $\alpha < E_0$ can be relaxed in much the same way as in § 3.2. For simplicity we consider situations where $(H-\alpha)^{-n/2}$ is self-adjoint.

When $n \geq 0$ the variational improvements can again be made more practicable with the aid of (3.12). We have

$$\hat{Y}_n(\alpha;\Phi) = \hat{Y}'_n(\alpha;\Psi) \qquad \text{(say)}$$
$$= (E_0-\alpha)^n|\langle\Psi, (H-E_0)\chi\rangle|^2/\langle\Psi, (H-E_0)^2(H-\alpha)^n\Psi\rangle, \qquad \Psi\in\mathscr{H}. \tag{3.23}$$

Alternative substitutions are available when $n < 0$.

In the case $n = 0$, expression (3.23) leads to the result

$$|\langle\theta_0, \chi\rangle|^2 \leq \langle\chi, \chi\rangle - |\langle\Psi, (H-E_0)\chi\rangle|^2/\langle\Psi, (H-E_0)^2\Psi\rangle. \tag{3.24}$$

This variational upper bound seems first to have been discovered by Rayner (1962a) and subsequently by other workers (Weinhold and Wang 1971). The moment-theoretical upper bounds of Gordon (1968) are a special case of (3.24) (Wang 1969).

4. Overlap with excited states

4.1. *The choice $A = H-E_k$*

We now set θ equal to θ_k, an excited state eigenvector, and discuss bounds on the square-overlap $|\langle\theta_k, \chi\rangle|^2$ of an arbitrary vector χ with θ_k.

It is not possible to pick $A = (H-E_k)^{1/2}$ as we did with $k = 0$, for in general this operator is not self-adjoint. However, guided by arguments similar to those in § 3.1, the next best thing is to choose

$$A = (H-E_k), \tag{4.1}$$

corresponding to which

$$u^2 = \infty \qquad l^2 = (E_j-E_k)^2 = \min[(E_{k+1}-E_k)^2, (E_{k-1}-E_k)^2]. \tag{4.2}$$

If we take as before

$$B = (H-\alpha)^{-n/2} \qquad n = \ldots -1, 0, 1 \ldots, \tag{4.3}$$

and suppose that α and n are such that B is self-adjoint, then we are led to families of variational upper and lower bounds on $|\langle\theta_n, \chi\rangle|^2$ which are rather similar to those obtained for the ground state. In obvious notation they are

$$\hat{F}^k_n(\Phi) = f^k_n(\alpha) - \hat{Y}^k_n(\alpha;\Phi) \geq |\langle\theta_k, \chi\rangle|^2 \geq g^k_n(\alpha) + \hat{Z}^k_n(\alpha;\Phi) = \hat{G}^k_n(\Phi) \tag{4.4}$$

where the non-variational parts are

$$f^k_n = (E_k-\alpha)^n\langle\chi, (H-\alpha)^{-n}\chi\rangle, \tag{4.5}$$

$$g_n^k = \frac{(E_k-\alpha)^n}{(E_j-E_k)^2}\langle\chi, (H-\alpha)^{-n}[(E_j-E_k)^2-(H-E_k)^2]\chi\rangle \tag{4.6}$$

and the optimal variational improvements are

$$\hat{Y}_n^k(\alpha;\Phi) = (E_k-\alpha)^n|\langle\Phi, (H-E_k)(H-\alpha)^{-n/2}\chi\rangle|^2/\langle\Phi, (H-E_k)^2\Phi\rangle, \tag{4.7}$$

$$\hat{Z}_n^k(\alpha;\Phi) = \frac{(E_k-\alpha)^n}{(E_j-E_k)^2}\frac{|\langle\Phi, (H-E_k)[(H-E_k)^2-(E_j-E_k)^2](H-\alpha)^{-n/2}\chi\rangle|^2}{\langle\Phi, (H-E_k)^2[(H-E_k)^2-(E_j-E_k)^2]\Phi\rangle}. \tag{4.8}$$

In the case $n = 0$ we have the familiar upper bound

$$|\langle\theta_k, \chi\rangle|^2 \leqslant \langle\chi, \chi\rangle - |\langle\Phi, (H-E_k)\chi\rangle|^2/\langle\Phi, (H-E_k)^2\Phi\rangle \tag{4.9}$$

(Rayner 1962a, Weinhold and Wang 1971), which is actually the same as (3.24) when $k = 0$. The corresponding lower bound is however quite distinct from (3.14) when $k = 0$. With the exception of (4.9) the families of variational bounds (4.4) appear to be new. Inconvenient negative or fractional powers of $(H-\alpha)$ can be removed from (4.7) and (4.8) by appropriate choice of the trial vectors Φ.

4.2. The bounds of Weinhold and Wang and variational improvements.

An alternative choice for A is suggested by the bounds of Weinhold and Wang (1971) which may be expressed in the form

$$\begin{aligned} f_\pm &= \left(\frac{E_k-\beta_\pm}{E_{k\pm1}-E_k}\right)\langle\chi, (E_{k\pm1}-H)(H-\beta_\pm)^{-1}\chi\rangle \geqslant |\langle\theta_k, \chi\rangle|^2 \\ &\geqslant \left(\frac{E_k-\beta_\pm}{E_{k\mp1}-E_k}\right)\langle\chi, (E_{k\mp1}-H)(H-\beta_\pm)^{-1}\chi\rangle = g_\pm, \end{aligned} \tag{4.10}$$

where

$$E_{k-1} < \beta_- < E_k < \beta_+ < E_{k+1}. \tag{4.11}$$

There are separate bounds here according as to whether β_+ or β_- is employed. As they stand these bounds are not suitable for calculations. However, with the aid of an operator inequality of Rosenberg *et al* (1960) it is possible to derive practical bounds on $\langle\chi, (H-\beta_\pm)^{-1}\chi\rangle$ which can be used in (4.10) (Epstein 1968, Weinhold and Wang 1971).

In the general theory of § 2 if we set

$$A_\pm = [(H-E_k)(H-\beta_\pm)^{-1}]^{1/2}, \qquad B = 1, \tag{4.12}$$

then we are led to these bounds of Weinhold and Wang, together with variational improvements for them. First we note that $A_\pm$ is self-adjoint since $(H-E_k)(H-\beta_\pm)^{-1}$ is a non-negative operator. Further we find that $A_\pm^2$ is bounded below and above respectively by

$$l_\pm^2 = \left(\frac{E_{k\mp1}-E_k}{E_{k\mp1}-\beta_\pm}\right), \qquad u_\pm^2 = \left(\frac{E_{k\pm1}-E_k}{E_{k\pm1}-\beta_\pm}\right), \tag{4.13}$$

the upper or lower signs being chosen consistently throughout. On making the substitutions (4.12) and (4.13) in (2.28) and (2.29) we obtain

$$\hat{F}_\pm(\Phi) = f_\pm - \hat{Y}_\pm(\Phi) \geqslant |\langle\theta_k, \chi\rangle|^2 \geqslant g_\pm + \hat{Z}_\pm(\Phi) = \hat{G}_\pm(\Phi) \tag{4.14}$$

where

$$\hat{Y}_{\pm}(\Phi) = \left(\frac{E_k-\beta_\pm}{E_k-E_{k\pm 1}}\right)\frac{|\langle\Phi,(H-E_{k\pm 1})(H-E_k)\chi\rangle|^2}{\langle\Phi,(H-E_{k\pm 1})(H-\beta_\pm)(H-E_k)^2\Phi\rangle} \tag{4.15}$$

and

$$\hat{Z}_{\pm}(\Phi) = \left(\frac{E_k-\beta_\pm}{E_{k\mp 1}-E_k}\right)\frac{|\langle\Phi,(H-E_{k\mp 1})(H-E_k)\chi\rangle|^2}{\langle\Phi,(H-E_{k\mp 1})(H-\beta_\pm)(H-E_k)^2\Phi\rangle}. \tag{4.16}$$

The functionals $\hat{Y}_{\pm}(\Phi)$ and $\hat{Z}_{\pm}(\Phi)$ give optimal variational improvements to the bounds of Weinhold and Wang.

5. Discussion

5.1. Feasibility of evaluating the variational functionals

The variational corrections, for example (3.13) and (3.23), are found in general to involve matrix elements of H^3 or even higher powers of H. Furthermore, it appears necessary to know certain energies, such as E_0 and E_1, exactly. Here we observe that neither requirement implies insuperable difficulties.

Regarding the evaluation of the matrix elements of higher powers of H, we observe the 'freedom' of the variational corrections. They need not involve quantities like $\langle\chi, H^n\chi\rangle$, which may well diverge (for a given n) depending on how accurately χ approximates an eigenfunction of H (for example, Gordon 1968). Instead, any trial vector Ψ such that say $\langle\Psi, H^n\chi\rangle$ and $\langle\Psi, H^n\Psi\rangle$ exist will lead to valid corrections. Such trial vectors can always be found. An interesting discussion of how to choose them is given by Weinhold (1973) who has established the feasibility of calculating $\hat{Z}'_0(\Psi)$ in (3.14) for helium.

Regarding the energy levels such as E_0 and E_1 which occur in (3.13), we note first of all that these numbers are usually known fairly accurately from experiment. The requisite energy levels need not be known *exactly* for the following reason. Once a bound has been obtained under the assumption that the energies are exact, the latter constraint can then be weakened and the bound correspondingly modified. As a simple example consider the variational upper bound (3.24) wherein we put $\Psi = \chi$:

$$|\langle\theta_0, \chi\rangle|^2 \leqslant \langle\chi,\chi\rangle - |\langle\chi,(H-E_0)\chi\rangle|^2/\langle\chi,(H-E_0)^2\chi\rangle. \tag{5.1}$$

If we assume $\langle\chi,\chi\rangle = 1$ so that

$$|\langle\theta_0, \chi\rangle|^2 \leqslant 1, \tag{5.2}$$

then we can rewrite (5.1) as

$$|\langle\theta_0, \chi\rangle|^2 \leqslant \frac{\langle\chi, H^2\chi\rangle - \langle\chi, H\chi\rangle^2}{\langle\chi, H^2\chi\rangle - 2E_0\langle\chi, H\chi\rangle + E_0^2}. \tag{5.3}$$

We now relax the constraint that E_0 is exact by assuming we know only that

$$E_0^{\mathrm{L}} < E_0 < E_0^{\mathrm{U}}. \tag{5.4}$$

The bound (5.3) is correspondingly modified to read

$$|\langle\theta_0, \chi\rangle|^2 \leqslant \frac{\langle\chi, H^2\chi\rangle - \langle\chi, H\chi\rangle^2}{\langle\chi, H^2\chi\rangle - 2E_0^{\mathrm{U}}\langle\chi, H\chi\rangle + (E_0^{\mathrm{L}})^2}. \tag{5.5}$$

Hence, providing the bounds (5.4) are tight enough to ensure that

$$\langle \chi, H\chi \rangle^2 > 2E_0^{\mathrm{U}}\langle \chi, H\chi \rangle - (E_0^{\mathrm{L}})^2, \tag{5.6}$$

the bound (5.5) is an improvement on (5.2). Thus we do not need to know the energies *exactly* in order to obtain meaningful bounds.

Finally we remark on a pleasing feature of having bounds in the form 'non-variational part' plus 'variational improvement'. The evaluation of a variational improvement is entirely independent of the means by which the non-variational part is calculated. For example, a lower bound for $g_n(\alpha)$ in (3.10) might be obtained by a subsidiary variational method similar to the Braun–Rebane bound in the case $n = 1$; *any* variational correction $\hat{Z}_n'(\alpha; \Phi)$ may be added to *any* such lower bound to $g_n(\alpha)$.

5.2. *Other types of bound?*

In §§ 3 and 4 we made several particular types of choice for the operators A and B. Many other choices are possible in theory; we could consider for example

$$\tilde{A}_{\pm} = [(H - E_k)(H - \beta_{\pm})]^{1/2} \tag{5.7}$$

in place of $A_{\pm}$ in § 4.2, or take different B's. However, we have restricted attention to those A's and B's which lead to results related in one way or another to the bounds of other workers. It was felt that such variational bounds would be the most practical ones.

One type of result apparently beyond the confines of our theory is exemplified by the variation lower bound on $|\langle \theta_0, \chi \rangle|^2$ derived by Weinhold (1967) with the aid of a Gram determinant. This bound is distinct from the lower bounds which we give in that it involves an auxiliary unknown overlap $|\langle \theta_0, \xi \rangle|^2$ with ξ a trial vector. If this quantity is eliminated with the aid of an Eckart bound, Weinhold's variational principle becomes equivalent to the Braun–Rebane one. More generally one is always faced with the need to eliminate $|\langle \theta_0, \xi \rangle|^2$, and it is interesting to speculate what occurs if, as an alternative to using the Eckart bound, one of the other variational bounds of § 3.2 is used. Surely nothing will be gained.

References

Arthurs A M 1970 *Complementary Variational Principles* (Oxford: Clarendon Press)
Barnsley M F and Robinson P D 1944 *J. Inst. Math. Applic.* **14** 229–49
Braun P A and Rebane T K 1969 *Int. J. Quant. Chem.* 3 1058–61
Eckart C 1930 *Phys. Rev.* **36** 878–92
Epstein S T 1968 *J. Chem. Phys.* **49** 1411–2
Gordon R G 1968 *J. Chem. Phys.* **48** 4984–8
Rayner Margaret E 1962a *Q. J. Math.* **13** 61–74
—— 1962b *Q. J. Math.* **13** 137–50
Rosenberg L, Spruch L and O'Malley T F 1960 *Phys. Rev.* **118** 184–92
Wang P S C 1969 *Int. J. Quant. Chem.* 3 57–66
Weinhold F 1967 *J. Chem. Phys.* **46** 2448–9
—— 1970 *J. Math. Phys.* **11** 2127–38
—— 1973 *J. Chem. Phys.* **59** 355–62
Weinhold F and Wang P S C 1971 *Int. J. Quant. Chem.* **5** 215–20