A projection method for generating bounds to overlap

1976 J. Phys. A: Math. Gen. 9 211

(http://iopscience.iop.org/0305-4470/9/2/004)

J. Phys. A: Math. Gen., Vol. 9, No. 2, 1976. Printed in Great Britain. © 1976

A projection method for generating bounds to overlap

Anthea Grubb
Department of Mathematics, Brunel University, Kingston Lane, Uxbridge, Middlesex

Received 3 September 1975

Abstract. In this paper a simple method is presented which unifies and generalizes a variety of upper and lower bounds for overlaps in a Hilbert space context.

1. Introduction

Let H be the Hamiltonian of a quantum mechanical system, suppose that H has an increasing sequence of eigenvalues $\{E_K\}$ and that ψ_K is the eigenvector corresponding to E_K. If ϕ is an approximation to ψ_K, then assuming that $\|\psi_K\| = \|\phi\| = 1$, the overlap $S_K = \langle\phi, \psi_K\rangle$ is a measure of the accuracy of this approximation. A number of papers have been published in recent years obtaining variational upper and lower bounds to S_K. In particular, Weinhold (1970, 1973) has used his method of Gram determinants to generate his inequalities, and Hoffmann-Ostenhof and Hoffmann-Ostenhof (1975) have started from an operator inequality of Löwdin to get their results.

In this paper, a simple Hilbert space inequality is obtained showing that various bounds are just special cases of this inequality. It uses the standard result (for example, see Simmons 1963) that the best approximation to a vector by a linearly independent set of vectors which span a subspace W is the orthogonal projection of the vector onto W. The inequality follows by observing that the length of this projection cannot exceed the length of the vector itself. In this paper we shall consider a few of the inequalities for the ground state overlap $S_0 = S = \langle\phi, \psi_0\rangle$ where E_0 is assumed to be non-degenerate. The generalization to excited states is obvious.

2. The projection method

Let $\{y_1, \ldots, y_n\}$ be any set of vectors which span a subspace W, and let v be a given vector in the Hilbert space. The vector

$$y = \alpha_1 y_1 + \ldots + \alpha_n y_n \tag{2.1}$$

in W which is the best approximation to v in the sense that $\|v - y\|$ is a minimum is the orthogonal projection of v on W, and hence $v - y$ is orthogonal to W. This leads to the set of equations

$$\langle y_i, y\rangle = \langle y_i, v\rangle \qquad i = 1, \ldots, n. \tag{2.2}$$

Let G be the Gram matrix of $\{y_1, \ldots, y_n\}$ with elements

$$G_{ij} = \langle y_i, y_j\rangle \tag{2.3}$$

and let β be the column vector with elements

$$\beta_i = \langle y_i, v \rangle . \tag{2.4}$$

Then equations (2.1) and (2.2) can be written as

$$G\alpha = \beta. \tag{2.5}$$

The matrix G is Hermitian since

$$(G^\dagger)_{ij} = \overline{\langle y_j, y_i \rangle} = \langle y_i, y_j \rangle = G_{ij} \tag{2.6}$$

and non-singular provided that $\{y_1, \ldots, y_n\}$ are linearly independent so that the equation can be solved for α.

Since y is the projection of v on W, we have

$$\|y\|^2 \leqslant \|v\|^2 \tag{2.7a}$$

that is

$$\sum_{i,j}^{n} \bar{\alpha}_i \alpha_j \langle y_i, y_j \rangle \leqslant \|v\|^2. \tag{2.7b}$$

Substituting $\alpha = G^{-1}\beta$, we obtain

$$\beta^\dagger G^{-1} \beta \leqslant \|v\|^2 \tag{2.8}$$

which is the inequality which yields bounds for the overlap. If $\{y_1, \ldots, y_n\}$ form an orthogonal set then G is diagonal with

$$G_{ij} = \|y_i\|^2 \delta_{ij} \tag{2.9}$$

and in this case inequality (2.8) takes the simpler form

$$\sum_i \frac{|\beta_i|^2}{\|y_i\|^2} \leqslant \|v\|^2. \tag{2.10}$$

3. The bounds of Weinhold and Rayner

The vector

$$v = \phi - \langle \psi_0, \phi \rangle \psi_0 \tag{3.1}$$

is orthogonal to ψ_0, and

$$\|v\|^2 = 1 - |S|^2. \tag{3.2}$$

Let $\{\chi_1, \ldots, \chi_n\}$ be a linearly independent set of vectors, and let

$$y_i = (H - E_0)\chi_i. \tag{3.3}$$

Provided that ψ_0 does not lie in the subspace spanned by $\{\chi_1, \ldots, \chi_n\}$, the set $\{y_1, \ldots, y_n\}$ is linearly independent and we can approximate v by this set. We find that

$$\beta_i = \langle \chi_i, (H - E_0)\phi \rangle \tag{3.4a}$$

and

$$G_{ij} = \langle \chi_i, (H - E_0)^2 \chi_j \rangle \tag{3.4b}$$

which leads to the upper bound

$$|S|^2 \leqslant 1 - \beta^\dagger G^{-1} \beta. \quad (3.5)$$

If in addition $\{y_1, \ldots, y_n\}$ is an orthogonal set, we get Rayner's upper bound (see Weinhold 1970)

$$|S|^2 \leqslant 1 - \sum_{i=1}^{n} \frac{|\langle \phi, (H-E_0)\chi_i \rangle|^2}{\langle \chi_i, (H-E_0)^2 \chi_i \rangle}. \quad (3.6)$$

We can generate lower bounds for $|S|^2$ by approximating the vector

$$v = (H-E_1)^{1/2}(\phi - \langle \psi_0, \phi \rangle \psi_0). \quad (3.7)$$

Since $(H-E_1)$ is a positive operator on the subspace of vectors orthogonal to ψ_0, the square root and hence v are well defined. A simple calculation shows that

$$\|v\|^2 = \langle H \rangle - E_1 + (E_1 - E_0)|S|^2 \quad (3.8)$$

where

$$\langle H \rangle = \langle \phi, H\phi \rangle. \quad (3.9)$$

If we choose the vectors $\{\chi_1, \ldots, \chi_n\}$ so that

$$y_i = (H-E_0)(H-E_1)^{1/2} \chi_i \quad (3.10)$$

is a linearly independent set and use this set to approximate v, we find that

$$\beta_i = \langle \chi_i, (H-E_1)(H-E_0)\phi \rangle \quad (3.11a)$$

and

$$G_{ij} = \langle \chi_i, (H-E_1)(H-E_0)^2 \chi_j \rangle. \quad (3.11b)$$

With these values of β_i and G_{ij} we have the lower bound

$$|S|^2 \geqslant (E_1 - \langle H \rangle + \beta^\dagger G^{-1} \beta)(E_1 - E_0)^{-1} \quad (3.12)$$

and if $\{y_1, \ldots, y_n\}$ is an orthogonal set the result is Rayner's lower bound (see Weinhold 1970)

$$|S|^2 \geqslant \left(E_1 - \langle H \rangle + \sum_{i=1}^{n} \frac{|\langle \phi, (H-E_1)(H-E_0)\chi_i \rangle|^2}{\langle \chi_i, (H-E_1)(H-E_0)^2 \chi_i \rangle} \right) (E_1 - E_0)^{-1}. \quad (3.13)$$

4. The lower bounds of Hoffmann-Ostenhof and Hoffman-Ostenhof

The lower bounds of Hoffmann-Ostenhof and Hoffmann-Ostenhof (1975) are expressed in terms of the operator

$$M = m(H-E_0) + |\phi\rangle\langle\phi| \quad (4.1)$$

where m is a positive real parameter. If ϕ is a reasonable approximation to ψ_0, then $\langle \phi, \psi_0 \rangle \neq 0$ and for any vector x

$$\langle x, Mx \rangle = m \langle x, (H-E_0)x \rangle + |\langle \phi, x \rangle|^2 > 0 \quad (4.2)$$

showing that M is a strictly positive operator. Thus provided the real numbers p and q satisfy the conditions

$$p < \inf_x \frac{\langle x, M^2 x\rangle}{\|x\|^2} \tag{4.3}$$

and

$$q < \inf_x \frac{\langle x, Mx\rangle}{\|x\|^2} \tag{4.4}$$

the operators $M^2 - p$ and $M - q$ are both strictly positive and hence $(M^2-p)^{1/2}$ and $(M-q)^{1/2}$ are both defined.

Suppose that we initially approximate the vector

$$v = (M^2 - p)^{1/2}\psi_0 \tag{4.5}$$

by a linear combination of the set

$$y_i = M(M^2 - p)^{1/2}\chi_i. \tag{4.6}$$

This set will be linearly independent provided that $\{\chi_1, \ldots, \chi_n\}$ are linearly independent. Calculation shows that

$$\|v\|^2 = |S|^2 - p \tag{4.7}$$

$$G_{ij} = \langle \chi_i, (M^4 - pM^2)\chi_j\rangle \tag{4.8}$$

and

$$\beta_i = S\gamma_i \tag{4.9}$$

where

$$\gamma_i = \langle (M^2 - p)\chi_i, \phi\rangle \tag{4.10}$$

hence

$$|S|^2 \geqslant p(1 - \gamma^\dagger G^{-1}\gamma)^{-1}. \tag{4.11}$$

If $\{y_1, \ldots, y_n\}$ form an orthogonal set, then

$$|S|^2 \geqslant p\left(1 - \sum_{i=1}^{n} \frac{|\langle \phi, (M^2-p)\chi_i\rangle|^2}{\langle \chi_i, (M^4 - pM^2)\chi_i\rangle}\right)^{-1}. \tag{4.12}$$

The case $n = 1$ is the first inequality of Hoffmann-Ostenhof and Hoffmann-Ostenhof (1975).

Another bound is obtained by approximating

$$v = (M - q)^{1/2}\psi_0 \tag{4.13}$$

by the set

$$y_i = M(M - q)^{1/2}\chi_i \tag{4.14}$$

where $\{\chi_1, \ldots, \chi_n\}$ is again a linearly independent set. We find that

$$\|v\|^2 = |S|^2 - q \tag{4.15}$$

$$G_{ij} = \langle \chi_i, (M^3 - qM^2)\chi_j\rangle \tag{4.16}$$

and

$$\beta_i = S\delta_i \tag{4.17}$$

where

$$\delta_i = \langle \chi_i, (M-q)\phi \rangle \tag{4.18}$$

and these equations lead to

$$|S|^2 \geqslant q(1-\delta^\dagger G^{-1}\delta)^{-1}. \tag{4.19}$$

By imposing the requirement that $\{y_1, \ldots, y_n\}$ form an orthogonal set, the inequality (4.19) takes the simpler form

$$|S|^2 \geqslant q\left(1-\sum_{i=1}^{n} \frac{|\langle \phi, (M-q)\chi_i \rangle|^2}{\langle \chi_i, (M^3-qM^2)\chi_i \rangle}\right)^{-1}. \tag{4.20}$$

Again, if $n=1$, this is the second inequality of Hoffmann-Ostenhof and Hoffmann-Ostenhof (1975).

References

Hoffmann-Ostenhof T and Hoffmann-Ostenhof M 1975 *Chem. Phys. Lett.* **31** 277–80
Simmons G F 1963 *Introduction to Topology and Modern Analysis* (New York: McGraw-Hill)
Weinhold F 1970 *J. Math. Phys.* **11** 2127–38
—1973 *J. Chem. Phys.* **59** 355–62