Linear inequalities and overlap bounds: a novel use of an operator inequality

1976 J. Phys. A: Math. Gen. 9 1605

(http://iopscience.iop.org/0305-4470/9/10/012)

J. Phys. A: Math. Gen., Vol. 9, No. 10, 1976. Printed in Great Britain. © 1976

Linear inequalities and overlap bounds: a novel use of an operator inequality

M Cohen and J G Leopold
Department of Physical Chemistry, Hebrew University, Jerusalem, Israel

Received 11 February 1976, in final form 9 June 1976

Abstract. Bounds to overlap integrals are obtained by an operator inequality technique. The formulation involves a set of linear inequalities with several overlap integrals as unknowns. The 'solution' of this set of inequalities leads to bound expressions which complement those of Weinberger and of Cohen and Feldmann.

1. Introduction

Operator inequalities have been used extensively to derive bounds to quantum mechanical properties. In many cases, they provide a more convenient framework than do other procedures, such as determinantal inequalities (for a review, see Weinhold 1972). The starting point is as follows. Given two operators, A and B, the inequality

$$A \geqslant B \tag{1}$$

means that, for any function ϕ in the common domain of the operators A and B,

$$\langle\phi|A|\phi\rangle \geqslant \langle\phi|B|\phi\rangle. \tag{2}$$

For such a pair of operators, it has been shown (Löwdin 1965) that the ordered eigenvalues of $A(a_0, a_1, a_2, \ldots)$ and the ordered eigenvalues of $B(b_0, b_1, b_2, \ldots)$ satisfy:

$$a_0 \geqslant b_0, \qquad a_1 \geqslant b_1, \qquad a_2 \geqslant b_2, \qquad \ldots. \tag{3}$$

If we rewrite inequality (1) in the form:

$$A - B \geqslant 0 \tag{4}$$

it is clear that the composite operator $A - B$ can have no negative eigenvalues. The best known application of this result is the variational principle, which may be written in operator form:

$$H - E_0 U \geqslant 0. \tag{5}$$

Here, H is the system Hamiltonian, U the identity operator and E_0 is the exact ground state energy. For operators with indefinite sign one can use a projection technique (Löwdin 1971) whereby the indefinite operator is projected onto a suitably chosen subspace in which the projected operator has definite sign. For example, if E_1 denotes the energy of the first excited state, one may project the operator $H - E_1 U$ onto a subspace orthogonal to ψ_0, the exact ground state eigenfunction, to obtain the non-negative projected operator (Weinhold 1970)

$$Q_0(H-E_1U)Q_0 \geqslant 0. \tag{6}$$

The projection operator Q_0 has the usual properties:

$$Q_0^2 = Q_0, \qquad Q_0^\dagger = Q_0 \tag{7a}$$

and its explicit form is clearly

$$Q_0 = 1 - |\psi_0\rangle\langle\psi_0|. \tag{7b}$$

For any trial function f_0, inequality (6) implies that

$$a_{00}^2 = |\langle f_0|\psi_0\rangle|^2 \geqslant \frac{E_1 - I_0}{E_1 - E_0} \tag{8a}$$

where

$$I_0 = \langle f_0|H|f_0\rangle. \tag{8b}$$

(8*a*) contains Eckart's (1930) lower bound to the ground state overlap integral.

2. The basic operator inequality

A more general projection technique leads to the following inequality for the Hamiltonian (Löwdin 1965):

$$(H-\alpha_n U)^{-1} \geqslant |f\rangle\langle f|H-\alpha_n U|f\rangle^{-1}\langle f|. \tag{9}$$

Inequality (9) is valid for any α_n for a given n-dimensional orthonormal basis set f chosen such that the matrix $\langle f|H|f\rangle$ and the operator H both have the same number of eigenvalues below the parameter α_n. Thus, we assume that

$$I_{n-1} \leqslant \alpha_n \leqslant E_n \tag{10}$$

where

$$\langle f_i|H|f_j\rangle = I_i\delta_{ij} \qquad i, j = 0, \ldots, n-1 \tag{11}$$

while, by the excited state variation theorem (Hylleraas and Undheim 1930, MacDonald 1933) we know that:

$$E_i \leqslant I_i \qquad i = 0, \ldots, n-1. \tag{12}$$

If we employ the first k functions of the n-dimensional basis, inequality (9) remains valid for *each* α_k chosen so as to satisfy

$$I_{k-1} \leqslant \alpha_k \leqslant E_k. \tag{13}$$

Thus, we have a *set* of operator inequalities

$$(H-\alpha_k U)^{-1} \geqslant |f\rangle\langle f|H-\alpha_k U|f\rangle^{-1}\langle f| \qquad k = 0, 1, \ldots, n \tag{14}$$

which form the basis of our theory.

3. Overlap bounds

The operator inequalities (14) are valid for any function in the common domain of the operators. For any eigenfunction ψ_β of H we have:

$$\langle\psi_\beta|(H-\alpha_k U)^{-1}|\psi_\beta\rangle \geq \langle\psi_\beta|f\rangle\langle f|H-\alpha_k U|f\rangle^{-1}\langle f|\psi_\beta\rangle. \tag{15}$$

We now expand each function f_i in terms of the complete set of eigenfunctions of H:

$$f_i=\sum_{j=0}^{\infty} a_{ij}\psi_j \qquad i=0,1,\ldots,n-1 \tag{16}$$

and obtain from (15) and (16):

$$\frac{1}{E_\beta-\alpha_k} \geq \sum_{i=0}^{m-1} \frac{a_{i\beta}^2}{I_i-\alpha_k} \qquad m\geq k. \tag{17}$$

Consider now the set of $m+1$ linear equations:

$$\frac{1}{E_\beta-\alpha_k} = \sum_{i=0}^{m-1} \frac{b_{i\beta}^2}{I_i-\alpha_k} \qquad k=0,1,\ldots,m \tag{18}$$

in m unknowns $b_{i\beta}^2$. The solution of any selected set of m linear equations chosen from these $m+1$ equations (we have $m+1$ sets in all) is given explicitly by:

$$b_{i\beta}^2(s)=\prod_{\substack{j=0\\ j\neq i}}^{m-1} \frac{I_j-E_\beta}{I_j-I_i} \prod_{\substack{k=0\\ k\neq s}}^{m} \frac{I_i-\alpha_k}{E_\beta-\alpha_k}, \qquad s=0,1,\ldots,m. \tag{19}$$

Combining (17) and (18) we obtain

$$\sum_{i=0}^{m-1} \frac{a_{i\beta}^2}{I_i-\alpha_k} \leq \sum_{i=0}^{m-1} \frac{b_{i\beta}^2(s)}{I_i-\alpha_k} = \frac{1}{E_\beta-\alpha_k} \qquad k=0,1,\ldots,m;\, k\neq s \tag{20}$$

comprising $m+1$ sets of linear inequalities. The solutions $b_{i\beta}^2$ of the linear equations (18) are given explicitly by (19), while the overlap integrals

$$a_{i\beta}^2=|\langle f_i|\psi_\beta\rangle|^2 \tag{21}$$

remain to be determined. To this end, we define the quantities $x_i(s)$ as:

$$x_i(s)=a_{i\beta}^2-b_{i\beta}^2(s) \tag{22}$$

and rewrite (20):

$$\sum_{i=0}^{m-1} \frac{x_i(s)}{I_i-\alpha_k} \leq 0 \qquad k=0,1,\ldots,m;\, k\neq s. \tag{23}$$

In defining $x_i(s)$ we may drop the index β since our entire procedure is valid for *any* β. However, β will specify a particular solution of equations (19).

In the following section we describe a general procedure for establishing the *signs* of *some* of the $x_i(s)$. This allows us to determine when a particular $b_{i\beta}^2(s)$ constitutes a bound to an individual $a_{i\beta}^2$. For, whenever $x_i(s)$ is positive, *then for any* β:

$$a_{i\beta}^2 \geq b_{i\beta}^2(s) \qquad \text{(lower bound)} \tag{24}$$

while for negative $x_i(s)$,

$$a_{i\beta}^2 \leq b_{i\beta}^2(s) \qquad \text{(upper bound).} \tag{25}$$

4. Solution for a set of linear inequalities

The set of inequalities (23) is similar to the classical linear programming problem (see, for example, Gass 1975). Our treatment differs from the conventional solution of the linear programming problem in that we are satisfied if we can determine bounds to *some* of the variables ($a_{i\beta}^2$). Consider a general set of linear inequalities:

$$\mathbf{Ax} \leqslant \mathbf{0} \tag{26}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{27}$$

The set of linear inequalities (26) is equivalent to a set of linear equations:

$$\mathbf{Ax} = -\mathbf{d} \tag{28}$$

where

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \tag{29}$$

and *every* d_i is non-negative, but is otherwise unknown. We now assume the existence of a transformation $\mathbf{T}_1$ which transforms $\mathbf{A}$ into nearly triangular form

$$\mathbf{T}_1\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & & a_{1n} \\ & a'_{22} & a'_{23} & \cdots & & a'_{2n} \\ & & a'_{33} & \cdots & & a'_{3n} \\ & 0 & & \ddots & & \vdots \\ & & & & a'_{n-1,n-1} & a'_{n-1,n} \\ & & & & a'_{n,n-1} & a'_{nn} \end{bmatrix} \tag{30}$$

and the vector $\mathbf{d}$ into a new vector $\mathbf{c}$,

$$\mathbf{T}_1\mathbf{d} = \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n-1} \\ c_n \end{bmatrix} \tag{31}$$

and assume that

$$c_{n-1} \geqslant 0, \qquad c_n \geqslant 0. \tag{32}$$

We now consider the pair of transformed equations:

$$\begin{bmatrix} a'_{n-1,n-1} & a'_{n-1,n} \\ a'_{n,n-1} & a'_{nn} \end{bmatrix} \begin{bmatrix} x_{n-1} \\ x_n \end{bmatrix} = -\begin{bmatrix} c_{n-1} \\ c_n \end{bmatrix}. \tag{33}$$

A transformation $\mathbf{T}$ which reduces the 2×2 matrix in (33) to triangular form is easily seen to be:

$$\mathbf{T}=\begin{bmatrix} 1 & 0 \\ -a'_{n,n-1}/a'_{n-1,n-1} & 1 \end{bmatrix}. \tag{34}$$

Then:

$$\mathbf{T}\begin{bmatrix} c_{n-1} \\ c_n \end{bmatrix}=\begin{bmatrix} c_{n-1} \\ c_n-(c_{n-1}a'_{n,n-1}/a'_{n-1,n-1}) \end{bmatrix} \tag{35}$$

where we require

$$c_n-c_{n-1}\frac{a'_{n,n-1}}{a'_{n-1,n-1}}\geqslant 0. \tag{36}$$

This implies that we must have

$$\frac{a'_{n,n-1}}{a'_{n-1,n-1}}\leqslant 0 \tag{37}$$

(since the magnitudes of c_n, c_{n-1} are unknown) and without loss of generality, we choose

$$a'_{n,n-1}\leqslant 0, \qquad a'_{n-1,n-1}\geqslant 0. \tag{38}$$

The triangular matrix remaining after the application of this transformation is:

$$\mathbf{T}\begin{bmatrix} a'_{n-1,n-1} & a'_{n-1,n} \\ a'_{n,n-1} & a'_{nn} \end{bmatrix}=\begin{bmatrix} a'_{n-1,n-1} & a'_{n-1,n} \\ 0 & a''_{nn} \end{bmatrix} \tag{39}$$

where

$$a''_{nn}=\left(1-\frac{a'_{n-1,n}a'_{n,n-1}}{a'_{n-1,n-1}a'_{nn}}\right)a'_{nn}=(1-\rho)a'_{nn}. \tag{40}$$

The *sign* of x_n is now determined entirely by the *sign* of a''_{nn}, and is always opposite to it. Thus, once we have determined the sign of a''_{nn}, we have a 'solution' for x_n. From (40) it is clear that a''_{nn} and a'_{nn} have the same sign whenever $\rho<1$, and have opposite signs when $\rho>1$. In view of our choice of signs in (38), ρ is negative (and thus less than unity) whenever $a'_{n-1,n}$ and a'_{nn} have the same sign. However, when $a'_{n-1,n}$ and a'_{nn} have different signs, ρ is positive and we must examine its magnitude. We summarize the possible cases in table 1. (Here and in the following, the signs of the elements will be denoted + and −.)

Table 1. Elements to determine the signs of x_n and x_{n-1}.

Case	$a'_{n-1,n}$	a'_{nn}	ρ	a''_{nn}	x_n	x_{n-1}
1	+	+	−	+	−	
2	−	−	−	−	+	
3	+	−	<1	−	+	−
4	+	−	>1	+	−	
5	−	+	<1	+	−	−
6	−	+	>1	−	+	

Cases 1 and 2 require no knowledge of the magnitudes of the individual elements, but still yield the sign of x_n alone. In all the remaining cases, we require the magnitudes of all the elements of the 2×2 matrix in order to determine ρ. However, if $\rho<1$ we not only determine the sign of x_n as before, but also find that the sign of x_{n-1} is always negative.

It is clear that to proceed further (i.e. to determine the signs of more than two of the x_i) we shall require more detailed knowledge of the elements of the matrix $\mathbf{A}$. Under certain restrictive conditions, it may be possible to deal with a 3×3 matrix equation in place of (33) and to determine the signs of three x_i's. In the present work, we do not seek a complete solution of inequalities (23) but find that the results summarized in table 1 enable us to obtain useful bounds for many cases of interest.

5. Solution of the set of inequalities (23)

Following the notation of the preceeding section, we rewrite our set of inequalities (23) in matrix form:

$$\mathbf{A}^s\mathbf{x}(s)=-\mathbf{d}^s \tag{41}$$

where now:

$$\mathbf{A}^s=\begin{bmatrix} \dfrac{1}{I_0-\alpha_0} & \dfrac{1}{I_1-\alpha_0} & \cdots & \dfrac{1}{I_{m-1}-\alpha_0} \\ \dfrac{1}{I_0-\alpha_1} & \dfrac{1}{I_1-\alpha_1} & \cdots & \dfrac{1}{I_{m-1}-\alpha_1} \\ \vdots & & & \vdots \\ \dfrac{1}{I_0-\alpha_m} & \dfrac{1}{I_1-\alpha_m} & \cdots & \dfrac{1}{I_{m-1}-\alpha_m} \end{bmatrix}^s ; \qquad \mathbf{d}^s=\begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_m \end{bmatrix}^s \tag{42}$$

and

$$\mathbf{x}(s)=\begin{bmatrix} x_0(s) \\ x_1(s) \\ \vdots \\ x_{m-1}(s) \end{bmatrix}. \tag{43}$$

The rows of $\mathbf{A}$ are ordered according to increasing α_k, and the columns according to increasing I_i so that if $j>i$ and $l>k$,

$$(I_j-I_i)>0, \qquad (\alpha_l-\alpha_k)>0. \tag{44}$$

This will be referred to as 'normal' ordering. The superscripts s in (42) indicate that the row containing α_s is deleted from $\mathbf{A}$ and the element d_s from $\mathbf{d}$. The index s in (43) serves to denote this particular solution.

Now consider:

$$\begin{bmatrix} \dfrac{1}{I_i-\alpha_k} & \dfrac{1}{I_j-\alpha_k} \\ \dfrac{1}{I_i-\alpha_l} & \dfrac{1}{I_j-\alpha_l} \end{bmatrix}\begin{bmatrix} x_i \\ x_j \end{bmatrix}=-\begin{bmatrix} d_k \\ d_l \end{bmatrix}. \tag{45}$$

We assume (cf (38) above) that

$$\frac{1}{I_i-\alpha_k}>0, \qquad \frac{1}{I_i-\alpha_l}<0 \tag{46}$$

and obtain

$$1-\rho=-\frac{(\alpha_l-\alpha_k)(I_j-I_i)}{(I_i-\alpha_l)(I_j-\alpha_k)}. \tag{47}$$

In cases of normal ordering and in view of the assumption of (46) it follows that

$$\frac{1}{I_j-\alpha_k}>0 \tag{48}$$

so that $(1-\rho)$ is necessarily positive. Thus the sign of x_j depends only on the sign of $(I_j-\alpha_l)$ (case 1 or 3 in table 1).

We have seen that assumption (46) is crucial to our development. However, it may happen (we give some examples in the next section) that $(I_i-\alpha_k)$ and $(I_i-\alpha_l)$ are of the same sign. In that case, by interchanging the columns of the matrix in (45) (and, of course, the rows of the solution vector) we obtain

$$\begin{bmatrix} \dfrac{1}{I_j-\alpha_k} & \dfrac{1}{I_i-\alpha_k} \\ \dfrac{1}{I_j-\alpha_l} & \dfrac{1}{I_i-\alpha_l} \end{bmatrix} \begin{bmatrix} x_j \\ x_i \end{bmatrix} = -\begin{bmatrix} d_k \\ d_l \end{bmatrix} \tag{49}$$

and we may now proceed provided that:

$$\frac{1}{I_j-\alpha_k}>0, \qquad \frac{1}{I_j-\alpha_l}<0. \tag{50}$$

Normal ordering together with the assumption of (50) now assures that

$$\frac{1}{I_i-\alpha_l}<0 \tag{51}$$

and here, we obtain

$$1-\rho=\frac{(\alpha_l-\alpha_k)(I_j-I_i)}{(I_j-\alpha_l)(I_i-\alpha_k)} \tag{52}$$

whose sign depends only on the sign of $(I_i-\alpha_k)$. Thus, the sign of x_i is determined in this case (case 2 or 4 in table 1).

We have thus seen that each of the cases 1–4 may arise and a solution *for at least one* of x_i and x_j is obtained in each case. Cases 5 and 6 do not arise in our specific problem.

Before proceeding, we sketch the reduction leading to the 2×2 matrix problem just solved. We consider the 3×3 problem

$$\begin{bmatrix} \dfrac{1}{I_p-\alpha_q} & \dfrac{1}{I_i-\alpha_q} & \dfrac{1}{I_j-\alpha_q} \\ \dfrac{1}{I_p-\alpha_k} & \dfrac{1}{I_i-\alpha_k} & \dfrac{1}{I_j-\alpha_k} \\ \dfrac{1}{I_p-\alpha_l} & \dfrac{1}{I_i-\alpha_l} & \dfrac{1}{I_j-\alpha_l} \end{bmatrix} \begin{bmatrix} x_p \\ x_i \\ x_j \end{bmatrix} = -\begin{bmatrix} d_q \\ d_k \\ d_l \end{bmatrix}. \tag{53}$$

The transformation

$$\mathbf{T}=\begin{bmatrix} 1 & 0 & 0 \\ \dfrac{I_p-\alpha_q}{\alpha_k-\alpha_q} & -\dfrac{(I_p-\alpha_k)(I_i-\alpha_q)}{(\alpha_k-\alpha_q)(I_i-I_p)} & 0 \\ \dfrac{I_p-\alpha_q}{\alpha_l-\alpha_q} & 0 & -\dfrac{(I_p-\alpha_l)(I_j-\alpha_q)}{(\alpha_l-\alpha_q)(I_j-I_p)} \end{bmatrix} \tag{54}$$

contains only positive elements (in the case of normal ordering) and reduces the matrix of (53) to the form

$$\begin{bmatrix} \dfrac{1}{I_p-\alpha_q} & \dfrac{1}{I_i-\alpha_q} & \dfrac{1}{I_j-\alpha_q} \\ 0 & \dfrac{1}{I_i-\alpha_k} & \dfrac{1}{I_j-\alpha_k} \\ 0 & \dfrac{1}{I_i-\alpha_l} & \dfrac{1}{I_j-\alpha_l} \end{bmatrix}. \tag{55}$$

This procedure can be generalized, and forms the justification of our treatment of the 2×2 matrix problem only.

6. An example

To demonstrate the application of our procedure, we give an example of a 4×3 matrix $\mathbf{A}$. It will be sufficient to give only the *signs* of the elements in the final triangular matrix but the rows and columns are indexed for clarity. We have here

$$\begin{array}{c} \\ \mathbf{A}= \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} + & + & + \\ - & + & + \\ - & - & + \\ - & - & - \end{bmatrix} \end{array} \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \tag{56}$$

yielding four 3×3 matrices $\mathbf{A}^0, \mathbf{A}^1, \mathbf{A}^2$ and $\mathbf{A}^3$. The transformations described in § 5 now lead to the following forms:

$$\mathbf{A}^0 \rightarrow \begin{array}{c} \begin{array}{ccc} 1 & 2 & 0 \end{array} \\ \begin{bmatrix} + & + & - \\ 0 & + & - \\ 0 & 0 & - \end{bmatrix} \end{array} \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \qquad x_0(0)\geqslant 0 \text{ (case 2)} \tag{57}$$

$$\mathbf{A}^1 \rightarrow \begin{array}{c} \begin{array}{ccc} 1 & 2 & 0 \end{array} \\ \begin{bmatrix} + & + & + \\ 0 & + & + \\ 0 & 0 & + \end{bmatrix} \end{array} \begin{array}{c} \\ 0 \\ 2 \\ 3 \end{array} \qquad x_0(1)\leqslant 0 \text{ (case 1)} \tag{58a}$$

$$\mathbf{A}^1 \to \begin{array}{c} \begin{array}{ccc} 0 & 2 & 1 \end{array} \\ \left[\begin{array}{ccc} + & + & + \\ 0 & + & - \\ 0 & 0 & - \end{array}\right] \end{array} \begin{array}{c} \\ 0 \\ 2 \\ 3 \end{array} \qquad x_1(1) \geq 0 \text{ (case 2)} \tag{58b}$$

$$\mathbf{A}^2 \to \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \left[\begin{array}{ccc} + & + & + \\ 0 & + & + \\ 0 & 0 & - \end{array}\right] \end{array} \begin{array}{c} \\ 0 \\ 1 \\ 3 \end{array} \qquad \left.\begin{array}{l} x_1(2) \leq 0 \\ x_2(2) \geq 0 \end{array}\right\} \text{(case 3)} \tag{59a}$$

$$\mathbf{A}^2 \to \begin{array}{c} \begin{array}{ccc} 0 & 2 & 1 \end{array} \\ \left[\begin{array}{ccc} + & + & + \\ 0 & + & + \\ 0 & 0 & + \end{array}\right] \end{array} \begin{array}{c} \\ 0 \\ 1 \\ 3 \end{array} \qquad x_1(2) \leq 0 \text{ (case 4)} \tag{59b}$$

$$\mathbf{A}^3 \to \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \left[\begin{array}{ccc} + & + & + \\ 0 & + & + \\ 0 & 0 & + \end{array}\right] \end{array} \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \qquad x_2(3) \leq 0 \text{ (case 1).} \tag{60}$$

These results are evidently quite general and enable us to conclude that regardless of the dimension of the original **A** (which reflects the size of the basis set employed in (17)) we shall always have:

$$x_{s-1}(s) \leq 0, \qquad x_s(s) \geq 0 \qquad \text{(all } s\text{).} \tag{61}$$

7. Results

From (61), we have quite generally:

$$a^2_{s-1,\beta} \leq b^2_{s-1,\beta}(s) \qquad \text{all } \beta; s = 0, \ldots, m \tag{62}$$

$$a^2_{s\beta} \geq b^2_{s\beta}(s) \qquad \text{all } \beta; s = 0, \ldots, m \tag{63}$$

and it remains to examine the properties of the bounds $b^2_{s-1,\beta}(s)$ and $b^2_{s\beta}(s)$. The general expression of these is given by equation (19).

7.1. Upper bounds

We have explicitly

$$b^2_{s-1,\beta}(s) = \frac{I_{s-1} - \alpha_0}{E_\beta - \alpha_0} \prod_{\substack{k=0 \\ k \neq s-1}}^{m-1} \frac{(I_k - E_\beta)(I_{s-1} - \alpha_{k+1})}{(I_k - I_{s-1})(E_\beta - \alpha_{k+1})} \tag{64}$$

and note that quite generally

$$\frac{(I_k - E_\beta)(I_{s-1} - \alpha_{k+1})}{(I_k - I_{s-1})(E_\beta - \alpha_{k+1})} = 1 - \frac{(\alpha_{k+1} - I_k)(I_{s-1} - E_\beta)}{(I_{s-1} - I_k)(\alpha_{k+1} - E_\beta)}. \tag{65}$$

Now from (13), we see that no E_β or I_{s-1} can lie in the interval (I_k, α_{k+1}) and $(\alpha_{k+1} - I_k) \geqslant 0$. It thus follows from (65) that when $I_{s-1} > E_\beta$ each term in the product (64) exceeds unity *unless* I_k and α_{k+1} lie in the interval (E_β, I_{s-1}). Such terms will be set equal to unity by choosing

$$\alpha_{k+1} \to I_k. \tag{66}$$

Also, when $\beta < s - 1$ the factor involving α_0 is greater than unity, and may be set equal to unity by choosing

$$\alpha_0 \to -\infty. \tag{67}$$

The optimal choice of all the remaining α's is seen to be given by

$$\alpha_k \to E_k \tag{68}$$

and we obtain the upper bound

$$a^2_{s-1,\beta} \leqslant \prod_{k=\beta}^{s-2} \frac{(I_k - E_\beta)(I_{s-1} - E_{k+1})}{(I_k - I_{s-1})(E_\beta - E_{k+1})} \qquad (\beta < s-1). \tag{69}$$

When $\beta = s - 1$ every term of (64) exceeds unity and the bound is trivial. When $I_{s-1} < E_\beta$, each term in product (64) exceeds unity when I_k and α_{k+1} lie in the interval (E_β, I_{s-1}). Thus, we obtain in this case

$$a^2_{s-1,\beta} \leqslant \frac{I_{s-1} - E_0}{E_\beta - E_0} \prod_{\substack{k=0 \\ k \neq s-1,\ldots,\beta-1}}^{m-1} \frac{(I_k - E_\beta)(I_{s-1} - E_{k+1})}{(I_k - I_{s-1})(E_\beta - E_{k+1})} \qquad (\beta > s-1). \tag{70}$$

7.2. *Lower bounds*

Here, it is convenient to rewrite (19) in the form:

$$b^2_{s\beta}(s) = \frac{I_m - I_s}{I_m - E_\beta} \prod_{\substack{k=0 \\ k \neq s}}^{m} \frac{(I_k - E_\beta)(\alpha_k - I_s)}{(\alpha_k - E_\beta)(I_k - I_s)} \tag{71}$$

where $(I_m - I_s)$ is always positive and none of the I_s can occur in any interval (α_k, I_k) (cf (13) above). Thus, each term of the product in (71) is positive *except* when $E_\beta = E_k$. This single negative term clearly cannot occur when $E_\beta > I_m$ but in that case, the factor $(I_m - E_\beta)$ makes the bound trivial. On the other hand, when $I_m > E_\beta$, the bound is again trivial *except* when $E_\beta = E_s$. In this case, we obtain the optimal lower bound ($\alpha_k \to E_k$ as before)

$$a^2_{ss} \geqslant \frac{I_m - I_s}{I_m - E_s} \prod_{\substack{k=0 \\ k \neq s}}^{m} \frac{(I_k - E_s)(E_k - I_s)}{(E_k - E_s)(I_k - I_s)}. \tag{72}$$

8. Discussion and conclusions

In this paper, we have been able to derive a non-trivial bound to *every* overlap integral $a_{i\beta}^2$. The lower bound (72) has been given previously by Weinberger (1960) and has been widely used. The upper bounds (70) in the general case $\beta = s$ (all s) and when $s = m$ ($\beta \geq m$) have also been derived by Cohen and Feldmann (1976). The upper bound (69) has not appeared elsewhere.

Both the lower bound (72) and the upper bound (70) are 'improvable' in the sense that they are refined by increasing the size of the basis set employed, (72) universally and (70) provided $m > \beta$. On the other hand, the upper bound (69) is *not* improvable, but is attained with a minimal basis set. In this case, the number of products appearing in the bound to $a_{s-1,\beta}^2$ is precisely $(s-1-\beta)$.

The intrinsic interest in these bounds is perhaps of less importance than the techniques developed here in obtaining them. We believe that our general method for solving sets of inequalities will prove to have far wider application.

Acknowledgment

The authors wish to thank Dr T Feldmann for helpful and enlightening discussions.

References

Cohen M and Feldmann T 1976 *J. Phys. A: Math. Gen.* **9** 1617–30
Eckart C 1930 *Phys. Rev.* **36** 878–92
Gass S I 1975 *Linear Programming: Methods and Applications* (New York: McGraw-Hill)
Hylleraas E A and Undheim B 1930 *Z. Phys.* **65** 759–72
Löwdin P O 1965 *Phys. Rev.* **139** A357–72
—— 1971 *Int. J. Quant. Chem.* **4** 231–7
MacDonald J K L 1933 *Phys. Rev.* **43** 830–3
Weinberger H F 1960 *J. Res. Natl Bur. Stand.* B **64** 217–25
Weinhold F 1970 *J. Math. Phys.* **11** 2127–38
—— 1972 *Adv. Quant. Chem.* **6** 299–331