

Variation Principles for an Arbitrary Operator. III

By L. M. Delves

Abstract. The methods, described in two previous papers, for generating variation principles for the matrix elements of Hermitian operators are extended here in several ways. The method is first extended to cover inhomogeneous equations. A defect of the original formulation, that it involved two trial functions, is removed by rewriting the principle so that one only appears. Finally, variation-iteration schemes are proposed.

1. Introduction. In two previous papers [1], [2], hereafter called I and II, the following problem was considered. Given two states $|\alpha\rangle, |\beta\rangle$, which are solutions of the homogeneous equations

$$(1) \quad \begin{aligned} (H_\alpha - E_\alpha)|\alpha\rangle &= 0, \\ (H_\beta - E_\beta)|\beta\rangle &= 0, \end{aligned}$$

find a variation principle for the matrix element $\langle\alpha|W|\beta\rangle$ of an arbitrary Hermitian operator W . We here extend these results in several ways. We first extend the method to cover inhomogeneous equations of the type $A|\gamma\rangle = |b\rangle$, where $|b\rangle$ is a given vector, and derive variation principles for the number $\langle\gamma|W|\gamma\rangle$.

Second, we are able to simplify the results somewhat. The principles derived involve two trial functions for each state $|\alpha\rangle$; in Section 3 we show that an alternative principle may be given involving a single trial function. This principle requires the inversion of the operator $W - E_1$, where E_1 is a constant; for many operators W of practical interest, this inversion is trivial.

Finally, in Section 4 we give a suitable variation-iteration scheme for use with these principles, and thus a practical way of improving the results.

We use throughout the notation of finite matrices. Thus, the states $|\alpha\rangle, |\beta\rangle$ will be written α, β ; and the matrix element $\langle\alpha|W|\beta\rangle$ will be written $\alpha^+W\beta$. This notation is chosen for its brevity and ease of printing; the results are directly applicable also to differential operators H_α , providing the boundary conditions are such that these are Hermitian. This requirement precludes scattering states from our discussion; the modifications to the method required to include these were given in II, and similar modifications go through for the cases treated here.

2. Inhomogeneous Equations. We consider a system represented by the state vector x_0 satisfying the (linear) inhomogeneous equation

$$(2) \quad Hx_0 = b,$$

where the matrix H is hermitian; suppose that we wish to find a variation principle for the quantity $\langle W \rangle = x_0^+Wx_0$, for some Hermitian operator W . Then, following the procedures of I and II, we consider the associated equation

$$(H + \lambda W)x_\lambda = b,$$

which leads to

$$(3) \quad \begin{aligned} Hx_1 &= -Wx_0, \\ x_1 &= \left. \frac{dx_\lambda}{d\lambda} \right|_{\lambda=0}. \end{aligned}$$

Then it is easy to show, following again the arguments of I, that a variational estimate of $\langle W \rangle$ is $[W]$, where

$$(4) \quad [W] = x_0^+ Hx_1 + x_1^+ Hx_0 - b^+ x_1 - x_1^+ b + x_0^+ Wx_0.$$

Similarly for off-diagonal matrix elements: given two systems x_0, y_0 :

$$(5) \quad \begin{aligned} H_1x_0 &= b, \\ H_2y_0 &= c, \end{aligned}$$

we can define x_1, y_1 through

$$(6) \quad \begin{aligned} H_1x_1 &= -Wy_0, \\ H_2y_1 &= -Wx_0, \end{aligned}$$

and find as a variation principle for $y_0^+ Wx_0$:

$$(7) \quad \begin{aligned} \delta[y_0^+ Wx_0] &= 0, \quad \text{where} \\ [y_0^+ Wx_0] &= y_0^+ Wx_0 + y_0^+ H_2y_1 + x_1^+ H_1x_0 - c^+ y_1 - x_1^+ b. \end{aligned}$$

A variation principle for $x_0^+ W y_0$ follows from taking the conjugate transpose of (7).

An Example. As a simple sample of the use of these principles, we consider a set of linear equations (2) with

$$(8) \quad \begin{aligned} H &= \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix}, \\ b^+ &= (10, 1, 5), \end{aligned}$$

and suppose we are interested in the number $\sum x_{0i}^2 = x_0^+ W_1 x_0 = \langle W_1 \rangle$, where the operator $W_1 = I$, the unit operator. The exact solution to (2) and (3) in this case is

$$(9) \quad x_0^+ = (1, 2, 3); \quad x_1^+ = \left(-\frac{2}{3}, 2, \frac{2}{3}\right), \quad \langle W_1 \rangle = 14,$$

and we assume we have found approximate solution x_{0t}, x_{1t} :

$$(9a) \quad \begin{aligned} x_{0t}^+ &= (0.9, 2.2, 3.3), \\ x_{1t}^+ &= (-1.9, 1.9, 0.5). \end{aligned}$$

The errors in these approximate solutions are about 10%. The vector x_{0t} gives, as a first approximation to $\langle W_1 \rangle$:

$$\langle W_1 \rangle_t = x_{0t}^+ W_1 x_{0t} = 16.54,$$

while the variational principle (4) gives, with this x_{0t}, x_{1t} :

$$[W_1] = 14.18$$

with an error of about 1.3% compared to 18% in $\langle W_1 \rangle_t$.

This is likely to be a particularly favourable example because of the regular form of W chosen. However, we can repeat the calculation for a singular W_2 . We choose

$$(10) \quad W_2 = \begin{pmatrix} 0 & & \\ & 1 & \\ & & 0 \end{pmatrix}$$

for which $x_1^+ = (-1, 1, 1)$ and $\langle W_2 \rangle \equiv x_0^2 = 4$, so that the variation principle for this choice of W gives an approximation to a single component of x_0 .

With the same approximation as before for x_{0t} and $x_{1t}^+ = (-0.9, 0.9, 1.1)$, we find $\langle W_2 \rangle_t = 4.84$, $[W_2] = 4.24$.

The improvement in this case, although marked, is less dramatic.

3. An Alternative Principle. The variation principles developed so far in these papers have required the use of two trial functions, x_{0t} and x_{1t} , for each state x_0 of interest. It is possible to rewrite the equations so that only one of these functions, x_1 , appears, while retaining the variational character of the results, and we do this here. The new principle then involves the operator $(E_1 - W)^{-1}$; but when this operator can be evaluated, the new form may be easier to use than the old.

We consider here the eigenvalue equation treated in I. Suppose, for Hermitian H and eigenvalue E_0 ,

$$(H - E_0)x_0 = 0, \quad x_0^+x_0 = 1,$$

and that we are interested in

$$(11) \quad \langle W \rangle = x_0^+Wx_0.$$

We shall assume E_0 is known (as discussed in I, a variational approximation to E_0 is sufficient to retain the variational character of our results), and we shall write $H - E_0 = L$. Then we define x_1 as a solution of the equation

$$(12) \quad Lx_1 = (E_1 - W)x_0,$$

so that

$$E_1 = x_0^+Lx_1 + x_0^+Wx_0 = \langle W \rangle.$$

Then x_1 also satisfies the equation

$$(13) \quad \begin{aligned} Ax_1 &= 0, \\ A &= L(E_1 - W)^{-1}L. \end{aligned}$$

The operator A is Hermitian, since L and W are Hermitian.

Then it is easy to show that a variation principle for $\langle W \rangle$ is $[W]$, where

$$(14) \quad \begin{aligned} [W] &= E_1 + x_1^+Ax_1, \\ \delta[W] &= 0. \end{aligned}$$

This principle does not involve x_0 explicitly; methods of using (14) in practice are

similar to those discussed in I. It may appear that an approximation x_{0t} to x_0 is still necessary, first to give an approximation E_{1t} to E_1 through

$$E_{1t} = x_{0t}^+ W x_0,$$

and second, to determine the normalization of x_1 through (12). This is not so, however; for a given form of x_1 , which may itself contain a number of parameters, $[W]$ is a function of E_1 ; the best value of E_1 is then that for which $\partial[W]/\partial E_1 = 0$.

Further, the normalization of x_1 may be specified by imposing the condition

$$x_1^+ L(E_1 - W)^{-2} L x_1 = 1,$$

which, from (12), is equivalent to $x_0^+ x_0 = 1$, but which does not require a trial function for x_0 .

The same procedure can be carried through for off-diagonal elements of W , but the results are unfortunately much more complicated. We assume we have

$$(15) \quad \begin{aligned} L_1 x_0 &= 0, & x_0^+ x_0 &= 1, \\ L_2 y_0 &= 0, & y_0^+ y_0 &= 1, \end{aligned}$$

for Hermitian L_1, L_2 ; and we are interested in $y_0^+ W x_0$ for some operator W .

We shall assume

$$(16) \quad y_0^+ W x_0 = x_0^+ W y_0$$

so that the matrix element is real. Then we introduce functions x_1, y_1 satisfying

$$(17) \quad \begin{aligned} L_1 x_1 &= E_1 x_0 - W y_0, \\ L_2 y_1 &= E_2 y_0 - W x_0, \end{aligned}$$

for which $E_1 = x_0^+ W y_0, E_2 = y_0^+ W x_0 = E_1$ from condition (16). Then it is possible to show that a variation principle for $y_0^+ W x_0$ is

$$(18) \quad \begin{aligned} \delta[y_1^+ W x_0] &= 0, \\ [y_0^+ W x_0] &= E_1 - \frac{1}{2}[x_1^+ L_1 E_1 B L_1 x_1 + y_1^+ L_2 E_1 B L_2 y_1 \\ &\quad + x_1^+ L_1 W B L_2 y_1 + y_1^+ L_2 W B L_1 x_1], \end{aligned}$$

$$(18a) \quad B = (W^2 - E_1^2)^{-1}.$$

Equation (18) reduces to equation (14) when $L_1 = L_2$.

As for the diagonal elements, it is necessary to fix the normalization of x_1, y_1 in some way. One possible way, which does not involve a knowledge of a trial function for x_0, y_0 , is through the relations

$$(19a) \quad \begin{aligned} y_1^+ L_2 E_1 B^2 W L_1 x_1 + y_1^+ L_2 E_1^2 B^2 L_2 y_1 \\ + x_1^+ L_1 W E_1 B^2 L_2 y_1 + x_1^+ L_1 W^2 B^2 L_1 x_1 = 1, \end{aligned}$$

$$(19b) \quad \begin{aligned} x_1^+ L_1 E_1 B^2 W L_2 y_1 + x_1^+ L_1 E_1^2 B^2 L_1 x_1 \\ + y_1^+ L_2 W E_1 B^2 L_1 x_1 + y_1^+ L_2 W^2 B^2 L_2 y_1 = 1. \end{aligned}$$

Equations (19a), (19b) reduce to the conditions $x_0^+ x_0 = 1, y_0^+ y_0 = 1$, for exact solutions x_1, y_1 of (17).

It is interesting that together they imply

$$(19c) \quad x_1^+ L_1 B L_1 x_1 = y_1^+ L_2 B L_2 y_1,$$

which may be used to simplify (18), since E_1 is a constant.

The complicated nature of (18) and (19) would seem to rule them out as a practical means of calculation; the original principle of II would appear easier to use.

Inhomogeneous Equations. The method carries through also for inhomogeneous equations of the type considered in Section 2. We give here only the results for the diagonal elements. With the same notation as Section 2, a variation principle for $\langle W \rangle$ is

$$(20) \quad [W] = -(x_1^+ b + b^+ x_1 + x_1^+ H W^{-1} H x_1),$$

for which $\delta[W] = 0$, where x_1 satisfies (3) and (2).

An Example. If we repeat the example of Section 2 for $W = W_1 = I$, and with the same approximation x_{1t} to x_1 , we find, from (20),

$$[W_1] = 13.64.$$

Thus in this case, equation (20) gives less accurate results than equation (7), for the trial functions considered. Equation (20) cannot be used to calculate $\langle W_2 \rangle$, since W_2 has no inverse.

4. A Variation-Iteration Scheme. So far we have assumed that approximations x_{0t} and x_{1t} to x_0 and x_1 are available. We give here one method by which a sequence of such approximations may be generated. For simplicity, we again limit the discussion to diagonal matrix elements of W , although off-diagonal elements may be considered similarly.

We consider the equations

$$(21a) \quad (H - E_0)x_0 = 0, \quad x_0^+ x_0 = 1,$$

$$(21b) \quad (H - E_0)x_1 = (E_1 - W)x_0, \quad E_1 = x_0^+ W x_0,$$

and suppose that we have approximations $x_1^{(n)}, E_1^{(n)}, x_0^{(n)}, E_0^{(n)}$ to x_1, E_1, x_0, E_0 .

Then we can define $(n + 1)$ st approximations in several ways:

(1) Possible iterative schemes for x_0 are

$$(22a) \quad x_0^{(n+1)} = A H x_0^{(n)},$$

or

$$(22b) \quad x_0^{(n+1)} = A (E_1^{(n)} - W)^{-1} (H - E_0^{(n)}) x_1^{(n)},$$

where in each case the multiplier A is determined from the condition

$$x_0^{(n+1)+} x_0^{(n+1)} = 1.$$

(2) The obvious iterative scheme for E_0 is

$$(23) \quad E_0^{(n+1)} = x_0^{(n+1)+} H x_0^{(n+1)}.$$

(3) Possible iterative schemes for x_1 are

$$(24a) \quad (H - E_0^{(n+1)})x_1^{(n+1)} = (E_1^{(n)} - W)x_0^{(n+1)},$$

or

$$(24b) \quad (H - E_0^{(n+1)})x_1^{(n+1)} = (\langle W \rangle_0^{(n+1)} - W)x_0^{(n+1)},$$

or

$$(24c) \quad x_1^{(n+1)} = \frac{1}{E_0^{(n+1)}} [Hx_1^{(n)} + (W - E_1^{(n)})x_0^{(n+1)}],$$

or

$$(24d) \quad x_1^{(n+1)} = \frac{1}{E_0^{(n+1)}} [Hx_1^{(n)} + (W - \langle W \rangle_0^{(n+1)})x_0^{(n+1)}].$$

(4) Finally, a nonvariational estimate of $E_1 = x_0^+ W x_0$ is

$$(25a) \quad E_1^{(n+1)} = \langle W \rangle_0^{(n+1)} \equiv x_0^{(n+1)+} W x_0^{(n+1)},$$

while the variational estimates we have been discussing may be written

$$(25b) \quad E_1^{(n+1)} = E_1^{(n)} + x_1^{(n+1)+} (H - E_0^{(n+1)}) (E_1^{(n)} - W)^{-1} (H - E_0^{(n+1)}) x_1^{(n+1)},$$

or

$$(25c) \quad \begin{aligned} E_1^{(n+1)} &= \langle W \rangle_0^{(n+1)} \\ &+ x_1^{(n+1)+} (H - E_0^{(n+1)}) (\langle W \rangle_0^{(n+1)} - W)^{-1} (H - E_0^{(n+1)}) x_1^{(n+1)}, \end{aligned}$$

or

$$(25d) \quad \begin{aligned} E_1^{(n+1)} &= \langle W \rangle_0^{(n+1)} + x_0^{(n+1)+} (H - E_0^{(n+1)}) x_1^{(n+1)} \\ &+ x_1^{(n+1)+} (H - E_0^{(n+1)}) x_0^{(n+1)}. \end{aligned}$$

In each of the alternative forms (25), any of the forms (24) may be used for x_1 . We shall consider here several representative overall schemes, which we define as follows: First, (22b) appears to have no advantages over (22a), unless (22a) should not converge. We therefore consider only (22a), and define several approximations to $\langle W \rangle$ in an obvious notation as follows:

$$\begin{aligned} \langle W \rangle_0 &= E_1 & (25a), \\ \langle W \rangle_1 &= E_1 & (24c, 25c), \\ \langle W \rangle_2 &= E_1 & (24c, 25d), \\ (26) \quad \langle W \rangle_3 &= E_1 & (24b, 25d), \\ \langle W \rangle_4 &= E_1 & (24b, 25c), \\ \langle W \rangle_5 &= E_1 & (24a, 25b), \\ \langle W \rangle_6 &= E_1 & (24a, 25d). \end{aligned}$$

The forms using (24a), (24b) are at first sight attractive, since these generate directly the product $(H - E_0)x_1$, which is all that is required by (25); further, we do not then need to estimate E_0 separately. However, it is easy to see that they are never a useful sequence. In fact, we have

$$(27) \quad \begin{aligned} \langle W \rangle_3^{(n+1)} &= \langle W \rangle_4^{(n+1)} = 2 \langle W \rangle_0^{(n)} - \langle W \rangle_0^{(n+1)}, \\ \langle W \rangle_i^{(n+1)} &= 2 \langle W \rangle_i^{(n)} - \langle W \rangle_0^{(n+1)}, \quad i = 5, 6. \end{aligned}$$

Thus, if the sequence $\langle W \rangle_0$ converges, the sequences $\langle W \rangle_5$ and $\langle W \rangle_6$ in general do not; further, while $\langle W \rangle_3$ and $\langle W \rangle_4$ converge, they will give at each stage a worse approximation than $\langle W \rangle_0$.

An Example. As an example, we consider a simple eigenvalue problem that has been used previously [3]. We take

$$(28) \quad H = \begin{pmatrix} -23 & 11 & 1 \\ 11 & -3 & -2 \\ 1 & -2 & 1 \end{pmatrix}, \quad W = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix},$$

and first approximations $x_0^{(1)} = x_1^{(1)} = (1, 0, 0)$.

The results for $x_0^{(n)}, x_1^{(n)}, E_0^{(n)}, E_1^{(n)}$ are listed in Table I.

TABLE I.
Variation iteration results for the matrix element $\langle W \rangle$ defined by equation (28).
The $\langle W \rangle_i$ are from different approximate schemes defined in equation (27).

n	$x_0^{(n)}$	$x_1^{(n)}$	$E_0^{(n)}$	$\langle W \rangle_0^{(n)}$	$\langle W \rangle_1^{(n)}$	$\langle W \rangle_2^{(n)}$	$\langle W \rangle_3^{(n)}$	$\langle W \rangle_5^{(n)}$
1	1, 0, 0	1, 0, 0	-23	1	1	1	1	1
2	0.902, - .431, - .039	0.823, - .378, - .033	-27.93	1.188	0.876	1.168	0.812	0.812
3	.913, - .404, - .062	0.833, - .355, - .051	-27.97	1.171	1.165	1.172	1.205	0.453
∞	.912, - .406, - .060	—	-27.97	1.172	1.172	1.172	1.172	—

It is seen that, as expected, $\langle W \rangle_5$ does not converge at all, while $\langle W \rangle_3$ is a very poor approximation. In this example, $\langle W \rangle_2$ is a better approximation than $\langle W \rangle_1$; this is due to the circumstance that the matrix $W - E_1$ is nearly singular.

As expected, the variational estimate $\langle W \rangle_2$ is a better approximation than the nonvariational $\langle W \rangle_0$. In fact, to the number of figures retained, $\langle W \rangle_2^{(3)}$ is already equal to $\langle W \rangle_0^\infty$. This example is a testing one since $x_0^{(n)}$ converges very rapidly to x_0 ; the improvement to be expected in general from the variation principle is greater than shown here.

Acknowledgements. Thanks are due to Professor J. M. Blatt and to Dr. G. H. Derrick for their interest in this work.

University of Sussex
Falmer, Brighton
England

1. L. M. DELVES, "A variation principle for arbitrary operators," *Nuclear Phys.*, v. 41, 1963, pp. 497-503. MR 26 #7317.
2. L. M. DELVES, *Nuclear Phys.*, v. 45, 1963, pp. 313-320.
3. D. R. HARTREE, *Numerical Analysis*, Clarendon Press, Oxford, 1952; 2nd ed., 1958, p. 201. MR 14, 690.