Useful extremum principle for the variational calculation of matrix elements

Edward Gerjuoy*
Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

A. R. P. Rau,†‡ L. Rosenberg,§ and Larry Spruch†
Physics Department, New York University, New York, New York 10003

(Received 18 June 1973)

Variational principles for the estimation of the matrix element $W_{nn} \equiv (\phi_n, W\phi_n)$ for an arbitrary operator W are of great interest. The variational estimates are constructed from a trial wave function ϕ_{nt}, an approximation to the nth normalized bound-state eigenfunction ϕ_n, and of a trial auxiliary function L_t, an approximation to L which satisfies $(H - E_n)L = (W_{nn} - W)\phi_n \equiv q(\phi_n)$. Variational-principle applications have been limited by the difficulty of obtaining a reasonable L_t; among other things, one demands that L_t approach L as ϕ_{nt} approaches ϕ_n. The equation $(H - E_{nt})L_t = q(\phi_{nt})$, where $E_{nt} = (\phi_{nt}, H\phi_{nt})$, is known *not* to provide such an L_t. A practical procedure for handling complicated systems given a reasonably accurate Rayleigh-Ritz trial function ϕ_{nt} is called for. This paper provides such a procedure, using techniques developed in the establishment of variational bounds on scattering lengths. Given H and ϕ_{nt}, we define L_t by $AL_t = q(\phi_{nt})$, where A differs from $H - E_n$ in that the influence of states 1 through n has effectively been "subtracted out"; the operator A is non-negative. A functional $M(L_{tt})$ is constructed which is an extremum for $L_{tt} = L_t$. Variational parameters contained in L_{tt} can be determined by extremizing $M(L_{tt})$, thereby providing an approximation to L_t. The method is analogous to the determination of parameters in ϕ_{nt} by the minimization of $(\phi_{nt}, H\phi_{nt})/(\phi_{nt}, \phi_{nt})$. The method is immediately applicable to the variational determination of off-diagonal matrix elements W_{nm} and of diagonal matrix elements of normal and of modified Green's functions.

I. INTRODUCTION

We will be concerned primarily with the approximate evaluation of diagonal matrix elements

$$W_{nn} \equiv \phi_n^\dagger W \phi_n \equiv (\phi_n, W\phi_n) \tag{1.1}$$

of an arbitrary known linear Hermitian operator W with respect to eigenfunctions defined by

$$(H - E_n)\phi_n = 0, \quad \phi_n^\dagger \phi_n = 1. \tag{1.2}$$

We note that the linearity in W of the matrix elements under consideration implies no loss of generality in taking W to be Hermitian; any operator can be written as a sum, with possibly complex coefficients, of two Hermitian operators.

For particular choices of W, one may be able to estimate W_{nn} by the use of a variational bound. If, for example, we have $W = H$ and we are interested in the lowest energy E_n of a state of given symmetry the well-known Rayleigh-Ritz theorem states that

$$E_n \leq \phi_{nt}^\dagger H \phi_{nt} \equiv E_{nt}, \tag{1.3}$$

for any normalized trial function ϕ_{nt} of the given symmetry. Not only is the error of second order in $(\phi_{nt} - \phi_n)$, it is of well-defined sign, so that an optimum choice for the parameter contained in ϕ_{nt} can be obtained by minimizing the variational bound on E_n.

The trial function ϕ_{nt} obtained from a Rayleigh-Ritz calculation may be used in the estimation of W_{nn} for any operator W. The estimate, $\phi_{nt}^\dagger W \phi_{nt}$, is, of course, generally neither a variational estimate of W_{nn} nor a bound. In fact, a trial function ϕ_{nt} which is good for estimating the energy E_n is not necessarily good for estimating $\phi_n^\dagger W \phi_n$. In many cases, therefore, it will be useful to have an alternative procedure. A variational bound on W_{nn} itself has been obtained for W of definite sign,[1] but the usefulness of the formal results has not yet been properly tested. We confine our attention here to variational principles (the error is of second order but its sign is unknown) for matrix elements. The variational principle for W_{nn},[2-5] for example, involves not only the trial function ϕ_{nt} but an auxiliary function L_t which depends on ϕ_{nt}. The parameters in ϕ_{nt} having been obtained from the Rayleigh-Ritz theorem (or a generalization in the case of excited states[6,7]), the problem of determining the parameters in L_t remains. It is this problem which forms the subject of the present paper.

In Sec. II we analyze the formal difficulty which lies at the root of the problem. The point is, roughly speaking, that a singularity in the equation which determines the auxiliary function must be eliminated; the elimination can take place by projecting out the state ϕ_n or by extracting the effects of ϕ_n in some other fashion, but the elimination must be performed in the absence of an

exact knowledge of ϕ_n.

In the remainder of Sec. II we give a brief review of the methods used in the past for overcoming this difficulty. We emphasize the restricted applicability or the restricted utility of such methods. The restricted applicability is associated with the allowable trial functions ϕ_{nt}, while the restricted utility is associated with the difficulty in obtaining adequate approximations to the solutions of the associated equations, the auxiliary functions. In Sec. III we present a method which is not only of unrestricted applicability, allowing the use of an arbitrary ϕ_{nt}, but has the very desirable feature that the parameters contained in the trial auxiliary functions can be determined by minimizing a given functional. While the existence of such a minimum principle is not surprising (it is closely related to the Rayleigh-Ritz property and is known to exist in those special cases alluded to above) it is nonetheless gratifying. It adds greatly to the power of the variational approach.

The mathematical apparatus for the formulation of an extremum principle for the estimation of L_t was suggested by a technique used in the development of a variational bound on the scattering length.[8] This technique is based on the Hylleraas-Undheim method[6] which provides variational bounds on the energies of excited states even when the lower states are not known exactly. In the latest of a series of developments of this idea in scattering theory a variational bound was obtained[9] on the effective potential for the scattering of a particle by a target whose wave function is not known. Although the final estimate of any scattering parameter will not be a bound, the approach, rather similar to the one adopted in the present paper, can be a useful one. In general, extremum principles should be very useful and widely applicable when employed not only in the calculation of parameters of primary interest but in intermediate stages of the calculation as well.[10]

II. THE VARIATIONAL APPROACH

Although there have been a number of quite interesting applications,[2-5,11,12] the full power of the variational approach to estimates of matrix elements is probably very much greater than the applications made thus far might indicate. There are a number of reasons for this. First, some of the derivations contained irrelevant elements.[13] Second, they were unnecessarily restrictive; for example, they sometimes required ϕ_n to be real. We will discuss this point at the end of the present section. Further, in actual applications, they were incorrectly used on a number of occasions; we will return to this point shortly. Finally, a prescription of sufficient power and generality for the practical problem of the estimation of the auxiliary function has been missing. We believe that the prescription given in Sec. III of the present paper remedies that lack. It may be useful, though, to begin with some comments of a general nature.

The variational principle of interest is, for the diagonal case,[2-5]

$$(W_{nn})_{\text{var}} = \phi_{nt}^{\dagger} W \phi_{nt} + L_t^{\dagger}[(H - E_{nt})\phi_{nt}] + [(H - E_{nt})\phi_{nt}]^{\dagger} L_t , \tag{2.1}$$

where

$$E_{nt} = \phi_{nt}^{\dagger} H \phi_{nt} , \quad \phi_{nt}^{\dagger} \phi_{nt} = 1.$$

The auxiliary function L_t is an approximation to L (the n dependence of L_t and L has been suppressed), defined by the requirements that it be quadratically integrable and satisfy

$$(H - E_n) L(\vec{\mathbf{r}}) = (-\mu - W)\phi_n(\vec{\mathbf{r}}), \tag{2.2a}$$

with $\vec{\mathbf{r}}$ denoting the totality of space and spin coordinates; μ is to be determined by the requirement that the right-hand side must vanish when premultiplied by $\phi_n^{\dagger}$, since this is true for the left-hand side. This leads to the result

$$\mu = -\phi_n^{\dagger} W \phi_n . \tag{2.2b}$$

We therefore have the equation

$$(H - E_n) L(\vec{\mathbf{r}}) = [\phi_n^{\dagger} W \phi_n - W]\phi_n(\vec{\mathbf{r}}). \tag{2.3}$$

Equation (2.3) defines L only to within a multiple of ϕ_n. We can therefore impose the condition

$$\phi_n^{\dagger} L = 0. \tag{2.4}$$

(We will see later that this is not necessarily the most convenient choice but it suffices for the purpose of the present discussion.)

Having obtained a normalized approximation ϕ_{nt} to ϕ_n in some fashion (by the Rayleigh-Ritz method, for example), the problem is to obtain L_t in some practical way. Before going any further, it will be necessary to define "order." We will use the term order in the conventional sense. Thus, with $\delta\phi_n \equiv \phi_{nt} - \phi_n$, we assume that $\delta\phi_n = \epsilon_\phi \eta_n$, where ϵ_ϕ is a small real parameter and η_n is a function, normalized to unity and independent of ϵ_ϕ, which satisfies the admissibility conditions for trial bound-state functions. A quantity is of order m in $\delta\phi_n$ if it can be expressed as ϵ_ϕ^m multiplied by a factor which remains finite and nonvanishing for $\epsilon_\phi \to 0$. Correspondingly, with $\delta L \equiv L_t - L$, we assume that we can write $\delta L = \epsilon_L \xi$, with $\epsilon_L \ll 1$ and $\xi^{\dagger}\xi = 1$. A term of first order is one which is linear in the ϵ's. A term of second order is one

which is quadratic or bilinear in the ϵ's.

We return now to the problem of obtaining an approximation L_t to L sufficiently good for our purposes. Suppose, as might seem natural, we adopt the equation

$$(H - E'_{nt})L_t = (-\mu_t - W)\phi_{nt}. \tag{2.5}$$

E'_{nt} is an approximation to E_n which can differ from E_n in at most first order. In particular, E'_{nt} need not be chosen to be the variational estimate E_{nt}. An example of a choice of E'_{nt} which differs from E_n by a quantity of first order in $\delta\phi_n$ is given by

$$E'_{nt} = \frac{1}{2}\,\frac{\phi^\dagger_{nt}[\mathcal{W}H + H\mathcal{W}]\phi_{nt}}{\phi^\dagger_{nt}\mathcal{W}\phi_{nt}},$$

where $\mathcal{W}$ is some appropriate weighting factor, perhaps W. The first thing to note in connection with Eq. (2.5) is that μ_t cannot be determined by the argument given above to determine μ since the homogeneous version of Eq. (2.5) has no solution satisfying the boundary conditions when E'_{nt} is not an eigenvalue of H. Equation (2.5) has a unique solution for *any* value of μ_t. To preserve the variational principle we require μ_t to differ from $-\phi^\dagger_n W\phi_n$ in at most first order in $\delta\phi_n$. (For practical purposes we may assume that the error will appear in first order.) To be specific in the following discussion we make the choice $\mu_t = -\phi^\dagger_{nt}W\phi_{nt}$, which satisfies the criterion just stated, so that Eq. (2.5) becomes

$$(H - E'_{nt})L_t = (\phi^\dagger_{nt}W\phi_{nt} - W)\phi_{nt}. \tag{2.6}$$

A procedure which has (incorrectly) been used on occasion in the past is to choose $E'_{nt} = E_{nt}$ in Eq. (2.6), giving

$$(H - E_{nt})L_t = (\phi^\dagger_{nt}W\phi_{nt} - W)\phi_{nt}. \tag{2.7}$$

To see that (2.7) leads to a breakdown in the variational principle we may project both sides of (2.7) on to $\phi^\dagger_n$. The right-hand side of the projected equation is of first order in $\delta\phi_n$. The left-hand side is $(E_n - E_{nt})\phi^\dagger_n L_t$. Since $E_n - E_{nt}$ is of second order in $\delta\phi_n$, it follows that $\phi^\dagger_n L_t$ is of order $(\delta\phi_n)^{-1}$, or, using (2.4), that $\phi^\dagger_n\delta L$ is of order $(\delta\phi_n)^{-1}$. It follows in turn that δL has a ϕ_n component which is also of order $(\delta\phi_n)^{-1}$, and finally that $J \equiv [(H-E_{nt})\phi_{nt}]^\dagger\delta L$ is of first order in $\delta\phi_n$, contrary to the assumption used in arriving at the equation defining L, or, equivalently, in arriving ultimately at a variational principle; to have a variational principle, J must be of *second* order. [If we write $\delta L = \sum_m d_m\phi_m$, the requirement that (2.1) be variational demands that d_m be of first order for $m \neq n$. Since $[(H_n - E_{nt})\phi_{nt}]^\dagger\phi_n$ is of second order, d_n can be of zeroth order; it cannot however be of order $(\delta\phi_n)^{-1}$.]

The difficulty with Eq. (2.7) is made more apparent by projecting both sides on to $\phi^\dagger_{nt}$; we find

$$\phi^\dagger_{nt}(H - E_{nt})L_t = 0. \tag{2.8}$$

When this result is combined with (2.1) we see that the variational principle has been lost since $(W_{nn})_{var}$ reduces to $\phi^\dagger_{nt}W\phi_{nt}$. The fact that (2.7) is not an allowable approximation has been noted previously.[11]

The above discussion brings out the sensitive dependence on the choice of E'_{nt} in Eq. (2.6); it must be accurate *but not too accurate*. This undesirable feature of an approach based on (2.6) can be traced to the fact that the resolvent operator $(H - E'_{nt})^{-1}$ is almost singular. The inversion of the operator $(H - E_n)$ that appears in Eq. (2.3) defining L causes no trouble, for the effect of the singularity of $(H-E_n)^{-1}$ is eliminated by imposing the orthogonality condition (2.4). On the other hand, the near singularity of $(H - E'_{nt})^{-1}$ that arises in solving for L_t defined by (2.6) cannot be avoided; there is no solution of the homogeneous equation associated with (2.6) and we cannot impose an orthogonality condition that is the analogue of (2.4). In addition, the problem of finding reliable approximations to the solution of (2.6) for a given choice of E'_{nt} still remains. In Sec. III we introduce a different definition of L_t which is very much more convenient than Eq. (2.6) as an aid in generating auxiliary trial functions to be used in (2.1).

The procedure used most frequently in the past, which avoids the above-mentioned difficulty, is based on the introduction of an approximate Hamiltonian $\hat{H}$ and approximate eigenvalue $\hat{E}_n$, such that

$$\hat{H}\hat{\phi}_n = \hat{E}_n\hat{\phi}_n. \tag{2.9}$$

Equation (2.3) is then replaced by

$$(\hat{H} - \hat{E}_n)\hat{L} = (-\hat{\mu} - W)\hat{\phi}_n. \tag{2.10a}$$

Since $\hat{\phi}_n$ is an eigenfunction of $\hat{H}$ we have, in analogy with Eqs. (2.2b) and (2.4),

$$\hat{\mu} = -\hat{\phi}^\dagger_n W\hat{\phi}_n, \tag{2.10b}$$

$$\hat{\phi}^\dagger_n\hat{L} = 0. \tag{2.10c}$$

The condition (2.10c) eliminates any difficulties in the inversion of $\hat{H} - \hat{E}_n$ required for the determination of $\hat{L}$. We assume that $\hat{H}$ is sufficiently close to H so that the difference δL between $\hat{L}$ defined by (2.10) and L defined by (2.3) and (2.4) satisfies the requirement imposed above, namely, that δL can be expressed as $\delta L = \epsilon_L\xi$ with $\epsilon_L \ll 1$ and $\xi^\dagger\xi = 1$. That is, $\hat{L}$ is assumed to be an acceptable trial function.

To consider the usefulness of the approximation just described, it will be helpful to simplify the

discussion by restricting ourselves to the case $n = 1$. We then have

$$(\hat{H} - \hat{E}_1)\hat{L} = [(\hat{\phi}_1^\dagger W \hat{\phi}_1) - W]\hat{\phi}_1, \tag{2.11a}$$

$$\hat{\phi}_1^\dagger \hat{L} = 0. \tag{2.11b}$$

Although for $\hat{H}$ sufficiently close to H the $\hat{L}$ defined by Eqs. (2.11) differs from L only in first order, as required, it will be difficult to obtain a useful approximation to $\hat{L}$ unless $\hat{H} - \hat{E}_1$ is positive definite. (It is well known that this condition is required if one is to construct an extremum principle for the approximate evaluation of $\hat{L}$. We will use this condition later in a different context.) We can normally determine all of the ϕ_n defined by Eq. (2.9), since H is normally chosen such that that is the case; if $n = 1$ represents the ground state, then we know of course that $H - E_1$ is positive definite. In some cases, however, as in the example of (2.12) immediately below, we can readily find only one eigenfunction associated with H.

In fact, when $\hat{\phi}_1$ is the ground-state eigenfunction of $\hat{H}$ the existence of an extremum principle for the estimation of $\hat{L}$, defined by Eqs. (2.11), is to be expected since $\hat{L}$ can be interpreted as the first order correction to the bound state function $\hat{\phi}_1$ due to the perturbation W. The Hylleraas minimum principle for the second-order energy eigenvalue[14] can then be used to provide estimates of $\hat{L}$. The condition that $\hat{H}\hat{\phi}_1 = \hat{E}_1\hat{\phi}_1$ with $\hat{\phi}_1$ the ground-state eigenfunction of $\hat{H}$ is an extremely restrictive one and is met in practice only by the choice $\hat{\phi}_1$ as an appropriate Hartree product, generally a rather poor choice.

Note that for any $\hat{\phi}_1$, we find, on introducing the projection operator

$$\hat{P}_1 = \hat{\phi}_1\hat{\phi}_1^\dagger$$

and choosing[15,10]

$$\hat{H} \equiv (1 - \hat{P}_1)H(1 - \hat{P}_1) + \hat{P}_1 H \hat{P}_1, \tag{2.12}$$

that

$$\hat{H}\hat{\phi}_1 = (\hat{\phi}_1^\dagger H \hat{\phi}_1)\hat{\phi}_1 \equiv \hat{E}_1\hat{\phi}_1;$$

we can therefore avoid near-singularity difficulties for any $\hat{\phi}_1$. However, with $\hat{H}$ defined by (2.12), we are unable to find conditions on $\hat{\phi}_1$ which guarantee that $\hat{H} - \hat{E}_1$ is non-negative, that is, that $\hat{\phi}_1$ is the ground-state wave function of $\hat{H}$, and we cannot therefore readily obtain an approximation to $\hat{L}$.

For completeness we mention that under specified conditions the difficulty in formulating the variational principle which we described in the discussion following Eq. (2.5), can be removed. One writes[2,16]

$$L(\vec{r}) = f(\vec{r})\phi_n(\vec{r}), \tag{2.13a}$$

with $f(\vec{r})$ a function of coordinates. With a square bracket here denoting a commutator, Eq. (2.3) can then be rewritten as

$$[H, f]\phi_n = [(\phi_n^\dagger W \phi_n) - W]\phi_n. \tag{2.13b}$$

If we restrict ourselves to situations for which ϕ_n and L can be taken to be real we can simply replace ϕ_n by ϕ_{nt} and f by a function f_t in Eq. (2.13b) and find as our defining equation for f_t (to within a constant because of its appearance in a commutator)

$$[H, f_t]\phi_{nt} = [(\phi_{nt}^\dagger W \phi_{nt}) - W]\phi_{nt}. \tag{2.14}$$

Premultiplication by $\phi_{nt}^\dagger$ and the use of $H = H^\dagger$ verifies that the equation is consistent, the left- and right-hand sides each vanishing. There still remains the problem of finding approximate solutions to Eq. (2.14). Toward this end Schwartz[2] constructed the functional

$$\begin{aligned} M_s(f_{tt}) = {} & \tfrac{1}{2}\phi_{nt}^\dagger [f_{tt}, [H, f_{tt}]]\phi_{nt} \\ & + \phi_{nt}^\dagger f_{tt}(W - \phi_{nt}^\dagger W \phi_{nt})\phi_{nt} \\ & + \phi_{nt}^\dagger (W - \phi_{nt}^\dagger W \phi_{nt}) f_{tt}\phi_{nt}, \end{aligned} \tag{2.15}$$

which is stationary with respect to arbitrary variations of the function f_{tt} about f_t. In fact, if the potential energy operator is local and we choose a coordinate system for which the kinetic energy can be expressed as

$$T = -\tfrac{1}{2}\hbar^2 \sum_i \nabla_i^2/m_i = \sum_i T_i, \tag{2.16}$$

with the m_i appropriately defined reduced masses, then the error can be put in the form

$$M_s(f_{tt}) - M_s(f_t) = \tfrac{1}{2}\hbar^2 \sum_i \phi_{nt}^\dagger (\vec{\nabla}_i \delta f_t)^2 \phi_{nt}/m_i, \tag{2.17}$$

where $\delta f_t \equiv f_{tt} - f_t$; in arriving at (2.17), we used the relationships

$$\begin{aligned} (2m_i/\hbar^2)[\delta f_t, [T_i, \delta f_t]] & \\ = -[\delta f_t, [\nabla_i^2, \delta f_t]] & = -[\delta f_t, (\nabla_i^2 \delta f_t) + 2(\vec{\nabla}_i \delta f_t)\cdot\vec{\nabla}_i)] \\ = -2[\delta f_t, (\vec{\nabla}_i \delta f_t)\cdot\vec{\nabla}_i] & = 2(\vec{\nabla}_i \delta f_t)^2. \end{aligned}$$

The variational principle is thus seen to be an extremum principle under the above-mentioned circumstances.

While the variational method based on the commutator approach has been successfully applied to a number of problems, the validity of the variational principle based on that approach is rather restricted. The reason lies in the appearance of the nodes in the functions ϕ_n. They appear not only in excited state functions but, in most cases, in ground-state functions as well, the exceptions being limited essentially to the one-body problem and the ground state of He and the associated isoelectronic series in atomic physics. In general,

$f = L/\phi_n$ will be singular at the nodes of ϕ_n. The function f_t, defined by Eq. (2.14), can be expected to be singular at the nodes of ϕ_{nt}. Though the nodes of ϕ_{nt} are known, the strength of the singularity in f_t is (except for certain one-dimensional problems[5]) not known. Therefore, we can expect that δf_t will be singular, thus violating the requirement that it be of first order. More explicitly, with regard to the form shown in Eq. (2.17), we would argue that $\vec{\nabla}_i \delta f_t$ near the node will be proportional to $\vec{\nabla}_i(1/\phi_{nt})$, that is, proportional to $(1/\phi_{nt})^2$. The error term, that is, the right-hand side of (2.17), is then infinite.

III. EXTREMUM PRINCIPLE OF GENERAL APPLICABILITY

We now introduce an extremum principle for the evaluation of the auxiliary function L_t valid for arbitrary ϕ_{nt}; in particular ϕ_{nt} need be neither nodeless nor real, nor need it be possible to obtain a simple associated Hamiltonian. We will begin with diagonal matrix elements considering first the ground-state case and the excited states. In the remaining subsections we consider, briefly, off-diagonal matrix elements and effects of degeneracy and, finally, we remark on the connection with the construction of "Green's functions in the generalized sense."[17]

A. Diagonal matrix elements: ground state

Our objective is to replace Eq. (2.3), with $n = 1$, by a self-consistent equation involving L_t, ϕ_{1t}, and E_{1t}, rather than L and the unknown entities ϕ_1 and E_1, such that L_t approaches L as ϕ_{1t} approaches ϕ_1; further, we want it to be possible to redefine, and thereby more readily estimate, L_t by means of an extremum principle. The problem is very similar to one encountered earlier in the study of variational bounds on scattering lengths.[8] In the latter case all discrete bound states of H had to be "extracted" (their energies lie below the scattering energy) while here it is only the ground state which must be "extracted." This necessitates only minor changes in the derivation of the extremum principle.

As noted earlier, the difficulties that arise in an attempt to approximate Eq. (2.3) defining L by an equation defining L_t have their origins in the fact that the operator $H - E_n$ in (2.3) has the eigenvalue zero, associated with ϕ_n; in the present subsection, we have $n = 1$. This difficulty can be avoided if we can shift the zero eigenvalue. The shift is possible since the ϕ_1 component of L is not defined by (2.3).

We begin by introducing the projection operator P_1 onto the ground state,

$$P_1 = \phi_1 \phi_1^\dagger. \tag{3.1}$$

For later convenience, we introduce the normalized trial function ϕ_t, the trial projection operator P_{1t},

$$P_{1t} = \phi_{1t} \phi_{1t}^\dagger, \tag{3.2a}$$

and

$$E_{1t} = \phi_{1t}^\dagger H \phi_{1t}. \tag{3.2b}$$

We are free to choose L such that

$$\phi_1^\dagger L = c_1, \tag{3.3}$$

with c_1 arbitrary, that is, such that

$$P_1 L = c_1 \phi_1. \tag{3.4}$$

We multiply (3.4) by σ, with σ arbitrary, subtract from (2.3), and find

$$(H - \sigma P_1 - E_1) L = [(\phi_1^\dagger W \phi_1) - \sigma c_1] \phi_1 - W \phi_1. \tag{3.5}$$

As opposed to $H - E_1$, which has eigenvalues 0, $E_2 - E_1, \ldots$, the operator $H - \sigma P_1 - E_1$ has eigenvalues $-\sigma E_1, E_2 - E_1, \ldots$; for $\sigma \neq 0$, the zero eigenvalue (and only the zero eigenvalue) has been shifted. Note too that for $\sigma \neq 0$, L is uniquely defined by (3.5), the solution being

$$L = c_1 \phi_1 - \sum_{m \neq 1} \frac{\phi_m (\phi_m^\dagger W \phi_1)}{E_m - E_1}; \tag{3.6}$$

Eq. (3.4) is an automatic consequence of (3.5).

Assuming, as we shall from now on, that $\sigma \neq 0$, it is simple to find an approximation to (3.5) for which L_t approaches L as ϕ_{1t} approaches ϕ_1. The near-singularity difficulty having been overcome, we need merely write

$$(H - \sigma_t P_{1t} - E_{1t}) L_t = [(\phi_{1t}^\dagger W \phi_{1t}) - \sigma_t c_{1t}] \phi_{1t} - W \phi_{1t}. \tag{3.7}$$

Barring extraordinary circumstances, such as $\phi_{1t} = \phi_n$ and $\sigma_t + E_{1t} = E_n$, L_t is uniquely defined by (3.7). Furthermore, for $\delta\phi_1 = \phi_{1t} - \phi_1$, and for $c_{1t} - c_1$ of first order in $\delta\phi_1$, L_t will differ from L by terms of first order[18] in $\delta\phi_1$. In particular, the eigenvalue $-\sigma E_1$ of $H - \sigma P_1 - E_1$ will be replaced by an eigenvalue that differs from $-\sigma E_1$ by a term that vanishes as $\delta\phi_1$ vanishes, and no near singularities arise when $H - \sigma_t P_{1t} - E_{1t}$ is inverted. (Compare this with the situation that occurs for $\sigma = 0$.) Equation (3.7) therefore provides a prescription, containing in fact two arbitrary parameters, c_{1t} and σ_t, that defines an appropriate approximation L_t of L.

While we now have a sensible defining equation, (3.7), for L_t, it would be difficult to obtain an adequate approximation to L_t. (Fortunately, we need not solve for L_t exactly to preserve our variational principle.) We can, however, obtain a different defining relationship for L_t such that L_t again differs from L in first order in $\delta\phi_1$ *and*

such that L_t can itself be rather readily approximated. If in (3.5) we set $\sigma = E_1$ and note that $P_1 = HP_1H/E_1^2$, we are led to replace (3.5) by

$$(H^{(1)}_{\mathrm{mod},t} - E_{1t})L_t = [(\phi^\dagger_{1t} W \phi_{1t}) - E_{1t} c_{1t}]\phi_{1t} - W\phi_{1t}, \tag{3.8}$$

where we have introduced the modified Hamiltonian

$$H^{(1)}_{\mathrm{mod},t} \equiv H - \frac{HP_{1t}H}{E_{1t}}. \tag{3.9}$$

The merit of (3.8) as opposed to (3.7) lies in the fact that it will normally be possible to prove, as we will show shortly, that

$$\psi^\dagger H^{(1)}_{\mathrm{mod},t}\,\psi - E_{1t} > 0 \tag{3.10}$$

for ψ any normalized function. We have already seen that L_t is uniquely defined by (3.8) and differs from L by first order in $\delta\phi$; if we can prove (3.10), it will also be true that L_t can be characterized as the function which extremizes a particular functional and can therefore be rather readily approximated. We will return to this point after discussing the proof of (3.10).

Taking first the case where more than one bound state exists, and letting ψ represent an arbitrary normalized trial function and E_1 and E_2 the two lowest energy eigenvalues, we have the inequality[9]

$$\psi^\dagger H^{(1)}_{\mathrm{mod},t}\,\psi \geq \frac{E_1}{E_{1t}} E_2. \tag{3.11}$$

We prove a more general inequality, which includes Eq. (3.11), in the Appendix. With the aid of even fairly crude lower bounds on E_1 and E_2, Eq. (3.11) provides a lower bound on $H^{(1)}_{\mathrm{mod},t}$ which, for a sufficiently accurate trial function ϕ_{1t}, will lie well above the ground-state energy level E_1, and, for a reasonable ϕ_{1t} and therefore E_{1t}, above E_{1t}.

Assuming that Eq. (3.11) is satisfied, the condition for the validity of the extremum principle can be stated as follows: If ϕ_{1t} is accurate enough so that

$$E_{1t} < (E_1/E_{1t})E_2, \tag{3.12a}$$

that is

$$E_{1t} < -(E_1E_2)^{1/2}, \tag{3.12b}$$

then Eq. (3.10) is satisfied, i.e.,

$$H^{(1)}_{\mathrm{mod},t} - E_{1t} > 0, \tag{3.13}$$

in the space of quadratically integrable functions. The result (3.13) holds when there is only one bound state of the system provided

$$E_{1t} < -(E_1E_{\mathrm{thr}})^{1/2}. \tag{3.14}$$

E_{thr} is the energy which marks the beginning of the continuum of the spectrum associated with H; the zero reference energy level will be assumed to have been so chosen that $E_{\mathrm{thr}} \leq 0$. [The similarity between (3.14) and (3.12b) is not surprising, since the second eigenvalue is $E_2 = E_{\mathrm{thr}}$ if there is only one bound state; on the other hand, E_2 is associated with a normalized bound state while E_{thr} is not.]

Once the validity of (3.10) has been demonstrated, we are in a position to write down the functional $M_{11}(L_{tt})$ which achieves its extremum for $L_{tt} = L_t$. (The subscripts on M denote the fact that we are here concerned with the diagonal ground state matrix element.) Thus, given $AX_t = q_t$ where A is a known non-negative operator and q_t a known function, we construct the functional

$$M(X_{tt}) = X^\dagger_{tt} A X_{tt} - X^\dagger_{tt} q_t - q^\dagger_t X_{tt}. \tag{3.15a}$$

Writing $X_{tt} = X_t + \delta X_t$, we have

$$M(X_t + \delta X_t) = M(X_t) + \delta X^\dagger_t A\, \delta X_t. \tag{3.15b}$$

$M(X_{tt})$ thus has its minimum value for $X_{tt} = X_t$. Correspondingly, if we set

$$q_t = q_t(c_{1t}, \phi_{1t}) = [(\phi^\dagger_{1t} W \phi_{1t}) - E_{1t} c_{1t}]\phi_{1t} - W\phi_{1t}, \tag{3.16}$$

(3.8) becomes

$$(H^{(1)}_{\mathrm{mod},t} - E_{1t})L_t = q_t, \tag{3.17}$$

and the functional

$$M_{11}(L_{tt}) = L^\dagger_{tt}(H^{(1)}_{\mathrm{mod},t} - E_{1t})L_{tt} - L^\dagger_{tt} q_t - q^\dagger_t L_{tt} \tag{3.18}$$

achieves its minimum value for $L_{tt} = L_t$. More precisely we have

$$M_{11}(L_t + \delta L_t) = M_{11}(L_t) + \delta L^\dagger_t (H^{(1)}_{\mathrm{mod},t} - E_{1t})\delta L_t,$$

and $M_{11}(L_{tt})$ achieves its minimum for $L_{tt} = L_t$, provided δL_t is quadratically integrable. But the quadratic integrability of L_t follows[19] from the nonsingular nature of the Green's function associated with the operator $H^{(1)}_{\mathrm{mod},t} - E_{1t}$ along with the (assumed) quadratic integrability of q_t. It is natural then to choose L_{tt} to be quadratically integrable, in which case δL_t will have the same property.

There are two rather natural choices of c_{1t}. The choice

$$c_{1t} = 0$$

leads to

$$q_t = (\phi^\dagger_{1t} W \phi_{1t})\phi_{1t} - W\phi_{1t},$$

the form which appears on replacing ϕ_1 by ϕ_{1t} in Eq. (2.3) which defines L. The choice

$$c_{1t} = (\phi^\dagger_{1t} W \phi_{1t})/E_{1t} \tag{3.19}$$

reduces q_t to its simplest form, namely, $q_t =$

$-W\phi_{1t}$. Equation (3.18) then reduces to its simplest form,

$$M_{11}(L_{tt}) = L_{tt}^{\dagger}(H^{(1)}_{\mathrm{mod},t} - E_{1t})L_{tt} + L_{tt}^{\dagger}W\phi_{1t} + \phi_{1t}^{\dagger}WL_{tt}. \tag{3.20}$$

B. Diagonal matrix elements: excited states

We now attempt to approximate Eq. (2.3) for $n > 1$, that is, for the case where the state under consideration is not the ground state. Before making any approximations it will be useful, as for the ground-state case, to rewrite Eq. (2.3). Our objective is to replace $H-E_n$ by an operator which does not have zero as an eigenvalue. Furthermore, for the purpose of obtaining an extremum principle, this operator should be of well-defined sign. Both conditions are satisfied by the operator $H^{(n)}_{\mathrm{mod}} - E_n$ where, in terms of the projection operators

$$P_i = \phi_i \phi_i^{\dagger},$$

we define the modified Hamiltonian

$$H^{(n)}_{\mathrm{mod}} \equiv H - \sum_{i=1}^{n} E_i P_i. \tag{3.21a}$$

For later purposes, we note that $H^{(n)}_{\mathrm{mod}}$ can be written as

$$H^{(n)}_{\mathrm{mod}} = H - \sum_{i=1}^{n} \frac{HP_iH}{E_i}. \tag{3.21b}$$

Corresponding to the eigenvalue equation

$$H\phi_i = E_i \phi_i, \quad i = 1, 2, \ldots$$

we have

$$(H^{(n)}_{\mathrm{mod}} - E_n)\phi_i = -E_n\phi_i, \qquad i \leq n,$$

$$(H^{(n)}_{\mathrm{mod}} - E_n)\phi_i = (E_i - E_n)\phi_i, \quad i > n.$$

We see that there is no zero eigenvalue present. It is also evident that $H^{(n)}_{\mathrm{mod}} - E_n$ is positive. Of course, the exact eigenfunctions ϕ_i are not known in practice. The significant feature of $H^{(n)}_{\mathrm{mod}}$ is that given a set of sufficiently accurate trial bound-state functions ϕ_{it}, $i = 1$ to n, it is possible to approximate $H^{(n)}_{\mathrm{mod}} - E_n$ by an operator $H^{(n)}_{\mathrm{mod},t} - E_{nt}$ which preserves the positivity property. We define

$$H^{(n)}_{\mathrm{mod},t} = H - \sum_{i=1}^{n} \frac{HP_{it}H}{E_{it}}, \tag{3.22a}$$

with

$$P_{it} = \phi_{it}\phi_{it}^{\dagger}. \tag{3.22b}$$

The trial functions are assumed to satisfy

$$\phi_{it}^{\dagger}\phi_{jt} = \delta_{ij}, \tag{3.23a}$$

$$\phi_{it}^{\dagger}H\phi_{jt} = E_{it}\delta_{ij}, \tag{3.23b}$$

$$E_{nt} < E_{\mathrm{thr}} \tag{3.23c}$$

for $1 \leq i, j \leq n$; with the E_{it} ordered, E_{nt} is assumed to lie below the threshold of the continuum. Since we are assuming the existence of at least n bound states it is always possible in principle to find such a set of functions. In practice we would construct the $n \times n$ matrix of H with n orthonormal functions. Diagonalization of this matrix leads to the functions ϕ_{it} and the eigenvalues E_{it} of Eqs. (3.23). According to the Hylleraas-Undheim theorem,[6] these eigenvalues satisfy $E_{it} \geq E_i$, $i = 1$ to n. Thus, the parameters contained in the original trial functions can be systematically improved by minimizing the eigenvalues E_{it}. A proof that $H^{(n)}_{\mathrm{mod},t} - E_{nt}$ is positive with respect to quadratically integrable functions for sufficiently accurate trial bound state functions ϕ_{it} is given in the Appendix.

The above discussion serves as motivation for the following transformation of Eq. (2.3). Subtracting $\sum_{i=1}^{n} E_i P_i L$ from both sides, Eq. (2.3) becomes

$$(H^{(n)}_{\mathrm{mod}} - E_n)L = -W\phi_n + P_nW\phi_n - \sum_{i=1}^{n} E_iP_iL. \tag{3.24}$$

We are at liberty to make a choice for P_nL since it is undetermined by Eq. (2.3). To simplify the right-hand side let us take

$$P_nL = \frac{1}{E_n} P_nW\phi_n. \tag{3.25}$$

The components P_iL, $i < n$, are fixed; from (2.3), we have

$$P_iL = \frac{1}{E_n - E_i} P_iW\phi_n, \quad i < n. \tag{3.26}$$

For computational purposes [see (3.32) below] it is useful to deal with the simplest possible form for the inhomogeneous term. We therefore introduce a new function, L', satisfying

$$(H^{(n)}_{\mathrm{mod}} - E_n)L' = -W\phi_n. \tag{3.27}$$

Since $P_iH^{(n)}_{\mathrm{mod}} = 0$ for $i \leq n$ we see that

$$P_iL' = \frac{1}{E_n} P_iW\phi_n, \quad i \leq n. \tag{3.28}$$

Furthermore, since $P_iH^{(n)}_{\mathrm{mod}} = P_iH = E_iP_i$ for $i > n$, we have

$$P_iL' = P_iL, \quad i > n. \tag{3.29}$$

We conclude from Eqs. (3.26)–(3.28) that

$$L = L' + \sum_{i=1}^{n-1} \left(\frac{1}{E_n - E_i} - \frac{1}{E_n} \right) P_iW\phi_n. \tag{3.30}$$

We now define a function L_i' as the solution of

$$(H^{(n)}_{\mathrm{mod},t} - E_{nt})L'_t = -W\phi_{nt}. \tag{3.31}$$

Once this function is determined suitable multiples of the functions ϕ_{it}, $i = 1$ to $(n-1)$ could be added to it to form a function L_t which has the property that $L_t \to L$ as the $\phi_{it} \to \phi_i$. This step is unnecessary, however, since, as a consequence of Eqs. (3.23), the components $P_{it}L_t$, $i \leq n$ make no contribution to the variational expression, Eq. (2.1). It suffices, therefore, to find an approximation, L'_{tt}, to the solution of Eq. (3.31). This can be accomplished, in a way which allows for systematic improvement of the approximation, by minimization of the functional

$$M_{nn}(L'_{tt}) = L'^{\dagger}_{tt}(H^{(n)}_{\mathrm{mod},t} - E_{nt})L'_{tt} + L'^{\dagger}_{tt}W\phi_{nt} + \phi^{\dagger}_{nt}WL'_{tt}. \tag{3.32}$$

C. Off-diagonal matrix elements

The variational principle for $(\phi_n, W\phi_m)$ requires the estimation of two auxiliary functions[20,21] which we will denote by L_{nm} and L_{mn}. The equation defining L_{nm}, in the present off-diagonal case, is of the form

$$(H - E_n)L_{nm} = q_{nm}, \tag{3.33}$$

where q_{nm} depends on both ϕ_m and ϕ_n, and is orthogonal to ϕ_n, as required for consistency. It is clear that the method described above for finding approximate solutions to Eq. (2.3) with the aid of an extremum principle can be applied directly to Eq. (3.33) since, aside from the orthogonality condition, the precise form of the inhomogeneous term plays no role in the discussion. Thus, we have a well-defined procedure for estimating each of the auxiliary functions.

D. Degeneracy

We have been assuming that the state or states under consideration are nondegenerate, but the formalism can easily be extended to allow for degeneracy. Assume that there are ω_n states associated with the energy E_n and let ϕ_{nr}, with $1 \leq r \leq \omega_n$ and with the ϕ_{nr} orthonormal, be a representation of these states. To be specific, we consider the off-diagonal matrix element $\phi^{\dagger}_{nr}W\phi_{ms}$, with $n \neq m$. (The argument is essentially identical for $n = m$.) The variational expression for this matrix element will again involve two auxiliary functions, which we will denote by L_{nrms} and L_{msnr}. The equation satisfied by L_{nrms} is of the form

$$(H - E_n)L_{nrms} = q_{nrms},$$

where

$$\phi^{\dagger}_{np}q_{nrms} = 0, \quad 1 \leq p \leq \omega_n.$$

An extremum principle for generating accurate approximations for the auxiliary functions can now be developed along the lines already described for the nondegenerate case. If we generalize Eq. (3.22b) to

$$P_{it} = \sum_{r=1}^{\omega_i} \phi_{irt}\,\phi^{\dagger}_{irt},$$

the formalism can be taken over intact. The fact that degeneracy can be treated in a fashion essentially as simple as that for the nondegenerate case is understandable; potential difficulties arising from vanishing energy denominators have already been disposed of.

E. Green's functions in the generalized sense

We have pointed out that our procedure for estimating solutions to equations of the type shown in (2.3) is applicable for any form of inhomogeneous term, subject to the usual consistency requirement that the inhomogeneous term be orthogonal to each solution of the homogeneous equation. Of course, equations of this type are frequently encountered, outside the context of the variational application emphasized here. The class of equations of the form (2.3) can be discussed in terms of the Green's function G associated with the operator $H - E_n$. Owing to the existence of a solution of the homogeneous equation this is a Green's function in the generalized sense. The equation satisfied by G,

$$(H - E_n)G = P_n - 1, \tag{3.34}$$

is obtained from Eq. (2.3) by replacing $W\phi_n$ by 1. Making a particular choice for the undefined component of G proportional to P_n we replace $W\phi_n$ by 1 in (3.30) and write

$$G = G' + \sum_{i=1}^{n-1}\left(\frac{1}{E_n - E_i} - \frac{1}{E_n}\right)P_i, \tag{3.35}$$

where, replacing $W\phi_{nt}$ in (3.31) by 1, G' is to be approximated by a solution to

$$(H^{(n)}_{\mathrm{mod},t} - E_{nt})G'_t = -1.$$

[The sum in (3.35) is to be approximated by using trial functions ϕ_{it} satisfying (3.23).] The associated extremum principle for G'_t is based on an equation of the type shown in (3.32). Since the trial function G'_{tt} is a function of two variables rather than one, the functional is to be thought of as an operator in configuration space. The extremum principle then applies to diagonal matrix elements of this operator taken with respect to square integrable functions.

APPENDIX: POSITIVITY OF $H^{(n)}_{\mathrm{mod},t} - E_{nt}$

Consider a set of $n+1$ orthonormal functions $\phi_{1t}, \phi_{2t}, \ldots, \phi_{nt}, \psi$. Suppose that the ϕ_{it} have

been chosen to satisfy Eqs. (3.23). The $(n+1)$-dimensional Hamiltonian matrix constructed from this set of functions has ordered eigenvalues which we denote by $E_i^{(n+1)}$, $i=1, 2, \ldots, n+1$. Now the determinant of a matrix is equal to the product of its eigenvalues. Furthermore, the n dimensional matrix that arises when we eliminate the last row and last column of the $(n+1)$ dimensional matrix is diagonal, with elements $(\phi_{it}, H\phi_{it}) = E_{it}$; the E_{it} arise in a study of n functions, and it will be convenient in this Appendix to write E_{it} as $E_i^{(n)}$. Thus, expanding the determinant of the $(n+1)$-dimensional Hamiltonian matrix, we have

$$\left(\prod_{i=1}^{n} E_i^{(n)}\right)\psi^\dagger H\psi - \sum_{i=1}^{n}(\psi^\dagger H\phi_{it})(\phi_{it}^\dagger H\psi)\prod_{j\neq i}^{n} E_j^{(n)} = \prod_{i=1}^{n+1} E_i^{(n+1)}. \tag{A1}$$

We can rewrite this as

$$\psi^\dagger H_{\mathrm{mod},t}^{(n)}\psi = \prod_{i=1}^{n}(E_i^{(n+1)}/E_i^{(n)})E_{n+1}^{(n+1)}. \tag{A2}$$

The system has at least n bound states with energies $E_i < E_{\mathrm{thr}}$ for $i=1$ to n. By assumption the functions ϕ_{it} are sufficiently accurate so that $E_i^{(n)} < E_{\mathrm{thr}}$. From the Hylleraas-Undheim theorem[6] we know that the $E_i^{(n+1)}$, for $i=1$ to n, satisfy

$$E_i \leq E_i^{(n+1)} \leq E_i^{(n)}. \tag{A3a}$$

Furthermore, if an $(n+1)$th bound state exists we have $E_{n+1}^{(n+1)} \geq E_{n+1}$; otherwise $E_{n+1}^{(n+1)} \geq E_{\mathrm{thr}}$. Since $E_{\mathrm{thr}} \leq 0$, it follows from (A3a) that

$$0 \leq E_i^{(n+1)}/E_i^{(n)} \leq E_i/E_i^{(n)}. \tag{A3b}$$

Suppose now that more than n bound states exist. Then

$$\psi^\dagger H_{\mathrm{mod},t}^{(n)}\psi \geq \prod_{i=1}^{n}(E_i^{(n+1)}/E_i^{(n)})E_{n+1} \geq \prod_{i=1}^{n}(E_i/E_i^{(n)})E_{n+1}. \tag{A4}$$

The inequality

$$\psi^\dagger[H_{\mathrm{mod},t}^{(n)} - E_n^{(n)}]\psi \geq 0 \tag{A5}$$

then follows, provided that

$$E_n^{(n)} \leq \prod_{i=1}^{n}(E_i/E_i^{(n)})E_{n+1}. \tag{A6}$$

This is always possible for sufficiently accurate trial bound state functions. If there are only n bound states the inequality (A5) follows, provided that

$$E_n^{(n)} \leq \prod_{i=1}^{n}(E_i/E_i^{(n)})E_{\mathrm{thr}}. \tag{A7}$$

This completes the proof that $H_{\mathrm{mod},t}^{(n)} - E_{nt}$, with $E_{nt} \equiv E_n^{(n)}$, is non-negative in the space of normalizable functions orthogonal to the trial bound-state functions ϕ_{it}, $i=1$ to n, provided these trial functions are sufficiently accurate to satisfy (A6) or (A7), whichever is appropriate.

Actually, the orthogonality requirement can be dropped. To see this we first introduce the projection operator $Q_t^{(n)}$ which projects onto the space orthogonal to the space spanned by the ϕ_{it},

$$Q_t^{(n)} \equiv 1 - \sum_{i=1}^{n} P_{it} \equiv 1 - P_t^{(n)}. \tag{A8}$$

The inequality then takes the form

$$Q_t^{(n)}(H_{\mathrm{mod},t}^{(n)} - E_{nt})Q_t^{(n)} \geq 0 \tag{A9}$$

in the space of quadratically integrable functions; due to the presence of $Q_t^{(n)}$ the orthogonality requirement need not be explicitly imposed. We then have

$$Q_t^{(n)}(H_{\mathrm{mod},t}^{(n)} - E_{nt})Q_t^{(n)} = H_{\mathrm{mod},t}^{(n)} - E_{nt}Q_t^{(n)} = H_{\mathrm{mod},t}^{(n)} - E_{nt} + E_{nt}P_t^{(n)}, \tag{A10}$$

since $P_{it}H_{\mathrm{mod},t}^{(n)} = H_{\mathrm{mod},t}^{(n)}P_{it} = 0$ for $i=1$ to n. Since $E_{nt}P_t^{(n)}$ is negative, it follows from (A9) and (A10) that $H_{\mathrm{mod},t}^{(n)} - E_{nt}$ is positive definite, for the ϕ_{it} sufficiently accurate, for *any* quadratically integrable function.

This more general version of the theorem is actually not essential in the applications considered here, although it allows us to eliminate the orthogonality condition on the auxiliary trial function and thereby to simplify the calculational procedure. The theorem used in the development of variational bounds on scattering lengths[8] asserts that if there are n bound states and $E_{\mathrm{thr}} = 0$ then $H_{\mathrm{mod},t}^{(n)} \geq 0$. (The more general case, $E_{\mathrm{thr}} < 0$ is treated by letting $H \to H - E_{\mathrm{thr}}$.) This result is contained in the present derivation.

---

*Work supported in part by the Advanced Research Project Agency under Contract No. DA-31-124-ARO-D-440, and by the National Aeronautics and Space Administration under Contract No. NGL-39-011-035.

†Work supported by the U. S Army Research Office, Durham, under Grant No. DA-ARO-D-31-124-72-G92 and by the Office of Naval Research, under Grant No. N00014-67-A-0467-0007.

‡Present address: Theoretical Physics Group, Tata Institute of Fundamental Research, Colaba, Bombay-5,

India.

§Supported by the National Science Foundation.

[1]R. Blau, A. R. P. Rau, and L. Spruch, Phys. Rev. A 8, 131 (1973).

[2]C. Schwartz, Ann. Phys. (N.Y.) 2, 156 (1959); 2, 170 (1959).

[3]A. Dalgarno and A. L. Stewart, Proc. R. Soc. Lond. A 257, 534 (1960); J. C. Y. Chen and A. Dalgarno, Proc. Phys. Soc. Lond. 85, 399 (1965).

[4]L. M. Delves, Proc. Phys. Soc. Lond. 92, 55 (1967).

[5]J. O. Hirschfelder, W. B. Brown, and S. T. Epstein, Adv. Quantum Chem. 1, 255 (1964).

[6]E. A. Hylleraas and B. Undheim, Z. Phys. 65, 759 (1930).

[7]A discussion of the Hylleraas-Undheim theorem (and a review of the methods developed for obtaining bounds on zero-energy scattering parameters) can be found in L. Spruch, in *Lectures in Theoretical Physics, Boulder, 1961,* edited by W. E. Britten, W. B. Downs, and J. Downs (Wiley, New York, 1962), Vol. IV. pp. 161–263. (For the theorem, see especially pp. 166–171.)

[8]L. Rosenberg, L. Spruch, and T. F. O'Malley, Phys. Rev. 118, 184 (1960).

[9]W. Monaghan and L. Rosenberg, Phys. Rev. A 6, 1076 (1972).

[10]S. Aranoff and J. Percus, Phys. Rev. 166, 1255 (1968).

[11]J. B. Krieger and V. Sahni, Phys. Rev. A 6, 919 (1972); Phys. Rev. A 6, 928 (1972).

[12]R. O. Mueller, A. R. P. Rau, and L. Spruch (unpublished).

[13]$\phi_n^\dagger W\phi_n$ can be thought of as the first-order energy correction of the perturbation W for W sufficiently small. The perturbation theoretic approach gives one a very good insight into the physical significance of the auxiliary function, but the perturbation theoretic derivation of the variational principle for W_{nn} obscures the fact that variational principles can be constructed in a number of contexts and even for cases where there is no simple perturbation interpretation.

[14]E. A. Hylleraas, Z. Phys. 65, 209 (1930).

[15]S. Borowitz and E. Gerjuoy, University of Colorado JILA Report No. 36, 1965 (unpublished).

[16]This method of approximation of the solution of inhomogeneous equations had arisen earlier in the context of perturbation theory, and had been analyzed by P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw- Hill, New York, 1953) and A. Dalgarno and J. T. Lewis, Proc. Roy. Soc. Lond. A 233, 70 (1955).

[17]We use the terminology of R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1965), Vol. I. p. 356.

[18]$L_t - L$ is of order $\delta\phi_1$ even if $\sigma_t - \sigma$ is not of first order in $\delta\phi_1$; this is not surprising since, as is obvious from (3.6), L is independent of σ. Note too that if $c_{1t} - c_1$ is not of first but of zeroth order in $\delta\phi_1$, and $\sigma_{1t} - \sigma_1$ is of zeroth order in $\delta\phi_1$, then, writing $L - L_t = \sum a_n\ \phi_n$, we find that a_1 is of zeroth order in $\delta\phi_1$, but a_n for $n \neq 1$ is of first order in $\delta\phi_1$. Since the ϕ_{1t} component of L_t makes no contribution to the variational expression of (2.1), we do indeed have a variational principle for W_{nn}, for any choice $\sigma_t \neq 0$ and c_{1t}. The above remarks follow from the manner in which (3.7) was arrived at. They can be verified explicitly by solving (3.7) for L_t explicitly; to do so, one can introduce $1 = \sum_m \varphi_{mt}\varphi_{mt}^\dagger$, where $(H - \sigma_t P_{1t})\varphi_{m\tau} = \mathcal{E}_{mt}\varphi_{mt}$, and invert $(H - \sigma_t P_{1t} - E_{1t})$. It is now simple to compare L_t and L.

[19]See, for example, A. Messiah, *Quantum Mechanics* (Wiley, New York, 1966), p. 713.

[20]L. M. Delves, Nucl. Phys. 45, 313 (1963).

[21]E. Gerjuoy, A. R. P. Rau, and L. Spruch, Phys. Rev. A 8, 662 (1973).