A unified formulation of the construction of variational principles

E. Gerjuoy

Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

A. R. P. Rau

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803

Larry Spruch

Physics Department, New York University, New York, New York 10003

The use of variational principles as a calculational tool is reviewed, with special emphasis on methods for constructing such principles. In particular, it is shown that for a very wide class of problems it is possible to construct a variational principle (VP) for just about any given quantity Q of interest, by routine procedures which do not require the exercise of ingenuity; the resultant VP will yield an estimate of Q correct to second order whenever the quantities appearing in the VP are known to first order. The only significant requirement for application of the routine procedures is that the entities which enter into the definition of Q be uniquely specified by a given set of equations; the equations may involve difference or differential or integral operators, they may be homogeneous or inhomogeneous, linear or nonlinear, self-adjoint or not, and they may or may not represent time-reversible systems. No numerical calculations are presented, but procedures for the construction of VP's are illustrated for numerous quantities Q of physical interest, particularly those Q arising in quantum-mechanical scattering and transition probability calculations. For pedagogical purposes VP's are also derived for several problems in classical and (simple) mathematical physics which the authors hope will prove instructive and perhaps even amusing. The quantum-mechanical quantities Q whose VP's are examined include various matrix elements and the quantum-mechanical eigenfunctions themselves. Topics examined include some points which have not always been appreciated in the literature, such as the necessity for properly specifying the phase when complex eigenfunctions are involved, and the importance of avoiding, wherever possible, formulations requiring the inversion of singular operators. The basic element of the technique is the recognition that the defining equations of a system can be incorporated into the VP as constraints through the use of (generalized) Lagrange multipliers; these can be constants, scalar, vector, or tensor functions of one or more variables, operators, etc. In the typical problem, these Lagrange multipliers $\mathscr{L}$ serve as a new set of adjoint functions or entities, and the construction of the VP simultaneously provides well-defined equations for the $\mathscr{L}$. Moreover, these $\mathscr{L}$ often have ready physical significance themselves; for example, they often may be regarded as generalized Green's functions. The construction of the VP also readily yields so-called variational identities for the quantity Q of interest; these identities give explicit (if formal) expressions for the error in the variational estimate of Q. In some cases this error can be shown to have a definite sign, so that the VP actually is an extremum principle, that is, that it yields an upper or lower variational bound for Q; however, our routine procedures for constructing VP's do not routinely yield extremum principles.

CONTENTS

I. INTRODUCTION

A. Background

In terms of the high accuracy of the results obtainable and the enormous range of applicability, variational principles are easily one of the most effective approximation methods in all of mathematical physics. Nevertheless, there has been little emphasis—and that relatively recent—placed on the development of a method which will enable one, in systematic fashion, to construct a variational principle (VP) for the quantity of interest. Variational principles in different areas are normally developed differently, using a wide range of tricks and approaches. A systematic approach has only recently been developed. Furthermore, recent progress in methods for using VP's makes it appear reasonably likely that these principles will prove more effective in a wider domain in the near future than they have in the past. For these reasons, a review of the formulation of VP's seems particularly appropriate at this time.

If the researcher starts with approximate solutions—or trial guesses—of the quantities of interest and possibly of related quantities, these VP's yield expressions that automatically lead to estimates good to second order, even though the trial guesses are, by definition, good only to first order. (Since the output is then generally much superior to the input, a VP can be thought of as an "intelligence amplifier.") Thereby such VP's provide a powerful iterative procedure for obtaining successively better estimates of the quantity of interest starting from some reasonably good trial approximation. An alternative way of using such principles is to have "open" variational parameters in the trial choice which are finally varied to seek a stationary value for the estimate.

The literature on VP's is vast, but many physicists seem to be under the impression that special simplifying circumstances (linearity, self-adjointedness, involvement of the Hamiltonian operator itself as opposed to an arbitrary operator, time reversibility, etc.) must apply if one is to be able to obtain a VP, and even then only if one exercises ingenuity. Over the last fifteen years or so, the work of several groups, proceeding largely independently in different branches of physics, has gone a considerable way towards dispelling this pessimistic view. It now seems possible to construct systematically a VP for just about any given quantity Q of interest, provided that the entities which enter into the definition of Q are uniquely defined by a set of equations; the equations may involve difference or differential or integral operators, they may be homogeneous or inhomogeneous, linear or nonlinear, self-adjoint or not, and they may or may not represent time-reversible systems.[1] This possibility has not gained the wide recognition it warrants. In particular, in the area of quantum mechanics, where the work of the authors and their coworkers has mostly been concentrated, no unified review of the formalism has been presented. The present paper is designed to fulfill this need. Particularly in the areas of atomic and nuclear physics, what is often of interest is the matrix element of some operator between two eigenfunctions of a Hamiltonian. The standard procedure in bound-state problems in these areas is to obtain approximations to the eigenfunctions separately through the Rayleigh-Ritz energy principle and then to sandwich the operator in between, a procedure which is not variational; recall that the energy principle gives a variationally accurate energy value but that the eigenfunction thereby obtained is not itself variationally good. Apart from the Rayleigh-Ritz principle, only in the area of scattering theory are a few VP's for matrix elements widely used. We survey at some length how our general procedure can be applied to write VP's directly for any matrix element of interest or for the eigenfunctions themselves.

Basic to the technique is the recognition that a variational principle is provided by the incorporation of the defining equations of a system as constraints through the use of (generalized) Lagrange multipliers—which can be constants, scalar, vector, or tensor functions of one or more variables, operators, etc. In a general problem, these Lagrange multipliers serve as a new set of adjoint functions or entities and the construction itself provides well-defined equations for them. These adjoint functions often

[1]For an illustration of a VP for a nonlinear system, see Appendix C; for the construction of a VP for the energy eigenvalue when the Hamiltonian is non-Hermitian and when the boundary conditions explicitly involve the eigenvalue, see Appendix D.

have ready physical significance themselves, and provide insight into the nature of the system. They typically may be regarded as generalized Green's functions. A by-product of our technique is the prescription for constructing them for integral and nonlinear operators where they are less well known. The exact meanings of the terms "adjoint" and "generalized" Green's functions will become clear as we proceed. Approximate solutions of the original defining set and this adjoint set serve as inputs to the desired variational estimate of the quantity of interest. In the course of the construction, we also get, equally simply, so-called "variational identities" for the quantity of interest; these give explicit if formal expressions for the error in the variational estimate. These errors can sometimes be bounded so that the VP is actually an extremum principle. We give careful attention to some points which have not always been appreciated in the literature, including the proper specification of the phase when complex functions are involved, and the importance of avoiding wherever possible formulations which require the inversion of singular operators.

The basic construction can be understood by a reading of Sec. I.B alone, where several illustrative examples of old and new VP's are looked at from the unified point of view, the examples being drawn from a wide variety of problems while at the same time being relatively free of some of the formal difficulties that can arise. These formal questions are taken up in Secs. II and III. The remaining sections deal with VP's for other quantum-mechanical problems. In Sec. IV, VP's are constructed for eigenfunctions and matrix elements of arbitrary operators. In Sec. V we consider at some length VP's in scattering theory because of their central role in atomic, molecular, and nuclear physics. A particularly subtle and complicated scattering theory VP, usable when the target ground-state wave function is only imprecisely known, is treated separately in Sec. VI. Time-dependent problems are considered in Sec. VII. A survey of the literature is given in Sec. VIII. Four appendixes further elaborate on certain aspects of the unified view of VP's.

Our main purpose is to estimate some functional F of ϕ, where ϕ designates a set of one or more real quantities which are only approximately known, but which are precisely defined by relations of the form

$$B_i(\phi)=0, \quad i=1,2,\ldots . \tag{1.1}$$

As one example, the desired functional might be the diagonal element of a given Hermitian operator W,

$$F(\phi)=\phi^{\dagger}W\phi , \tag{1.2}$$

where ϕ is a normalized bound-state eigenfunction of a Hermitian Hamiltonian H, that is, where

$$B_1(\phi)\equiv(H-E)\phi=0 , \tag{1.3a}$$

with

$$B_2(\phi)\equiv\phi^{\dagger}\phi-1=0 . \tag{1.3b}$$

In Eqs. (1.2) and (1.3), we use the dagger to denote the adjoint. (The transpose would do here, but we will later be dealing with complex entities.) Matrix notation, including summation over all discrete dummy indices and integration over all continuous dummy variables, is assumed. For example, with spinless particles $1,2,\ldots,n$, using the coordinate representation, we have (remembering that ϕ is real)

$$\phi^{\dagger}W\phi=\int d\mathbf{r}\,\phi(\mathbf{r})W\phi(\mathbf{r}) , \tag{1.4}$$

where $\mathbf{r}$ stands for the collection $\mathbf{r}_1,\ldots\mathbf{r}_n$, and $d\mathbf{r}$ represents $d\mathbf{r}_1 d\mathbf{r}_2\cdots d\mathbf{r}_n$. As Eq. (1.3b) illustrates, the defining relations in Eq. (1.1) need not be linear operations on ϕ.

Usually it is difficult to find the exact ϕ described by Eq. (1.1); certainly H of Eq. (1.3a) usually is so complicated that neither ϕ nor its corresponding eigenvalue E is known exactly. Under these circumstances, it often is advantageous to estimate $F(\phi)$ from a VP, as has long been recognized. We now begin consideration of the construction of VP's.

B. Some illustrative examples

Section II contains a formal approach to the construction of VP's. Many readers will prefer to begin with a number of illustrative examples, sufficiently complicated to illuminate some interesting points but sufficiently simple and/or well known to avoid any notational or conceptual obscurity.

1. Power dissipation in a passive resistive network

For a two-terminal passive resistive network, with the various resistances R_i and the input current I given, we consider the problem of estimating the power loss P of the system. A well-known theorem states that if we approximate the branch currents I_i by trial branch currents I_{it} such that current conservation is satisfied at each junction, the expression $\Sigma_i I_{it}^2 R_i$ will represent a VP for P; in fact, the expression represents an extremum (minimum) principle for P (Smythe, 1968, p. 252).

If one did not know that theorem, the question of how to go about estimating P would naturally arise. The generation of a VP for P turns out to be a particularly simple and illuminating example of our general approach.

To make the problem as absolutely simple as possible, we consider the case of two resistors, R_1 and R_2, in parallel, with a known input current I. We then have

$$P=I_1^2R_1+I_2^2R_2 , \tag{1.5}$$

where I_1 and I_2, which here play the role of the ϕ's, are to be determined from the constraints

$$B_1(I_1,I_2)\equiv I_1+I_2-I=0 , \tag{1.6}$$

$$B_2(I_1,I_2)\equiv I_1R_1-I_2R_2=0 . \tag{1.7}$$

Choosing trial approximations I_{it} which differ from I_i by first-order quantities $\delta I_i=I_{it}-I_i$, introducing a Lagrange

multiplier to account for each constraint—that is, for Eq. (1.6) and Eq. (1.7)—and including an approximation to P [defined by Eq. (1.5)], we come to the starting expression for our VP for P,

$$P_v \equiv \langle P(I_{1t},I_{2t},\lambda_{1t},\lambda_{2t})\rangle_{\rm var} = I_{1t}^2R_1+I_{2t}^2R_2+\lambda_{1t}(I_{1t}+I_{2t}-I) +\lambda_{2t}(I_{1t}R_1-I_{2t}R_2)\ . \tag{1.8}$$

The λ_{it} may be functions of the I_{it}; the only restriction on the I_{it} is that they be good approximations—by definition "to first order"—to the I_i.

The expression for P_v in Eq. (1.8) is a natural starting point, reducing as it does to the exact result for the power P when the I_{it} coincide with the exact I_i, independent of the values of λ_{it}. Furthermore, when the I_{it} have first-order errors, we can arrange to have the error in P_v of second order by choosing the λ_{it} to differ in first order from the λ_i, where the λ_i are defined by the requirement that $\delta P=0$ with respect to independent variations of I_{1t} and I_{2t}. The requirement that $\delta P=0$ gives

$$2I_1R_1+\lambda_1+\lambda_2R_1=0 \tag{1.9}$$

and

$$2I_2R_2+\lambda_1-\lambda_2R_2=0\ , \tag{1.10}$$

respectively. This pair of equations constitutes a pair of adjoint equations to Eqs. (1.6) and (1.7), and the λ_i are therefore referred to as adjoint functions. It will be noticed that the coefficients of the first-order quantities $\delta\lambda_1$ and $\delta\lambda_2$ on variation of P_v are automatically zero. [On varying P_v, we retain only first-order terms. The coefficients of $\delta\lambda_1$ and $\delta\lambda_2$ are thus I_1+I_2-I and $I_1R_1-I_2R_2$, respectively, which vanish by Eqs. (1.6) and (1.7).] This behavior, the vanishing of the coefficient of the first-order quantities $\delta\mathscr{L}$, is characteristic. Equations (1.9) and (1.10) have as their solution

$$\lambda_1=-I_1R_1-I_2R_2,\quad \lambda_2=0\ . \tag{1.11}$$

The λ_i are indeed functions of I_i. Here is an example of a problem where the equations for the exact λ's can be explicitly solved for. (The solution is, as often happens, only formal, since it involves the unknown I_i.) It then is natural, though not necessary, to choose for the trial solutions λ_{it} the values $\lambda_i(I_{it})$, so that

$$\lambda_{1t}=-I_{1t}R_1-I_{2t}R_2,\quad \lambda_{2t}=0\ .$$

The insertion of these trial values into Eq. (1.8) gives

$$P_v=I(I_{1t}R_1+I_{2t}R_2)-I_{1t}I_{2t}(R_1+R_2)\ , \tag{1.12}$$

which is indeed a VP. [Thus writing $I_{it}=I_i+\delta I_i$, one finds that $P_v=P-\delta I_1\delta I_2(R_1+R_2)$, that is, that P_v differs from P by a second-order term.] Note that if we choose I_{1t} and I_{2t} such that $I_{1t}+I_{2t}=I$, then the value of λ_{1t} to be inserted into Eq. (1.8) is irrelevant—indeed, with such I_{1t},I_{2t} the term proportional to $I_{1t}+I_{2t}-I$ in Eq. (1.8) was not needed from the very start—and P_v reduces to $I_{1t}^2R_1+I_{2t}^2R_2$. [The connection between Eq. (1.12) and the extremum principle quoted in the first paragraph of this subsection will be discussed later in Appendix B.]

The above result can readily be recast in matrix form; once the constraints analogous to Eqs. (1.6) and (1.7) have been given, the extension of the result in Eq. (1.12) to an arbitrary two-terminal passive resistive network is trivial.

2. Newton's method for the location of roots

Newton's method for the location of the roots of an equation is another trivial but illustrative example (suggested to us by Professor L. Rosenberg) of our procedure for constructing VP's. The problem is that of finding a VP for the root x_0 (in some neighborhood $-a\le x_0\le a$) of the equation

$$f(x_0)=0\ , \tag{1.13}$$

in the single one-dimensional real variable x, $-a\le x\le a$, where $f(x)$ is a known function. This illustration is so simple that casting it in the formalism of our general procedure may appear artificial and forced. Nevertheless, we note that in the present case x_0 represents the unknown quantity ϕ, which is now merely a single number, the desired functional $F(\phi)$ is x_0 itself, and the relation (1.13) defining x_0 is the sole constraint. Therefore, our starting point is

$$\langle x_0\rangle_{\rm var}=x_{0t}+\lambda_t f(x_{0t})\ . \tag{1.14a}$$

In Eq. (1.14a), x_{0t} is a trial estimate of x_0, the quantities $\langle x_0\rangle_{\rm var}$, x_{0t}, and $f(x_{0t})$ all are numbers, and λ_t is a number, a trial approximation to a numerical Lagrange multiplier.

With $x_{0t}=x_0+\delta x_0$ and $\lambda_t=\lambda+\delta\lambda$, the first variation of Eq. (1.14a) takes the form

$$\delta F_v=\langle x_0\rangle_{\rm var}-x_0 =\delta x_0+\lambda f'(x_0)\delta x_0+(\delta\lambda)f(x_0)\ .$$

The last term on the right-hand side of the above equation vanishes by Eq. (1.13); alternatively, we may say that requiring Eq. (1.14a) to be stationary with respect to arbitrary variations of λ_t about its exact value λ retrieves the constraint Eq. (1.13), as we expect. Putting the coefficient of δx_0 equal to zero in our expression for δF_v yields

$$1+\lambda f'(x_0)=0\ ,$$

which specifies λ. Once again, as in all cases when the multiplier is a number, it can be solved for exactly by a simple algebraic operation. Also, as in the previous problem, it is given in terms of the unknown ϕ (here x_0). A reasonable trial estimate of $\lambda=-1/f'(x_0)$ is therefore

$$\lambda_t=-1/f'(x_{0t})\ ,$$

and a reasonable estimate is all that is needed. (One need only know the Lagrange multipliers to first order to obtain a VP; our approach would be useless if one were required to determine these multipliers exactly.) The insertion of the above expression for λ_t in Eq. (1.14a) yields Newton's method (Courant, 1937),

$$\langle x_0\rangle_{\rm var}=x_{0t}-f(x_{0t})/f'(x_{0t}) \tag{1.14b}$$

for estimating the root x_0 of Eq. (1.13).

That the $\langle x_0\rangle_{\rm var}$ just derived is indeed a variational estimate of the root x_0 defined by $f(x_0)=0$ follows by construction. If we wish directly to confirm the variational property of $\langle x_0\rangle_{\rm var}$, we neglect second-order terms, use $f(x_0)=0$ to write $f(x_{0t})=f'(x_{0t})\delta x_0$, write $x_{0t}=x_0+\delta x_0$, and find

$$\langle x_0\rangle_{\rm var}=x_0+\delta x_0-\delta x_0=x_0 \ .$$

In the future, we will not bother to confirm the variational character of entities constructed to be variational. As this example and the power dissipation example above indicate, it is usually rather simple and direct to verify that the estimate constructed by our general procedure is in fact variational.

Evidently the above derivation of the variational estimate (1.14b) for x_0 could have been made much shorter. However, we wanted to make unmistakably clear our general procedure of incorporating the defining equations as constraints through Lagrange multipliers and determining the exact Lagrange multipliers by equating to zero all first-order variations; we shall not give so many details in succeeding examples.

If a VP for the root $\mathbf{r}_0=(x_0,y_0)$ satisfying $f_1(x_0,y_0)=f_2(x_0,y_0)=0$ is sought, the corresponding Lagrange multipliers λ_1 and λ_2 (associated with f_1 and f_2, respectively) involve the inverse of the Jacobian of f_1 and f_2. Similar results hold for the roots of sets of equations in higher dimensions. A more general treatment of Newton's rule is included in a paper by Aranoff and Percus (1968). For completeness, we add that Newton's method can be generalized (Rosenberg and Tolchin, 1973) to the solution of much more complicated equations than Eq. (1.13) [or than sets of equations like $f_1(x_0,y_0)=f_2(x_0,y_0)=0$], even to nonlinear integral equations; moreover, under appropriate circumstances the absolute value of the error made by using Newton's method can be bounded (Kantorovich and Akilov, 1964).

3. The inverse of an operator

As a third simple example, consider the variational determination of the inverse A^{-1} of a known operator A. We will assume that both the left and right inverses of A exist, are unique, and are identical. Thus the exact A^{-1} is defined by

$$AA^{-1}-1=0 \ , \tag{1.15a}$$

or, equivalently, by

$$A^{-1}A-1=0 \ . \tag{1.15b}$$

We shall also assume throughout that the associative law holds.

Because Eqs. (1.15) are equivalent under our present assumptions, they really do not provide independent constraints on the quantity A^{-1} regarded as an operator function of the known operator A. Therefore (as is further discussed in Sec. III), it should not be necessary in deriving a VP for A^{-1} to include both the constraints (1.15). Consequently, we shall take the starting point for our VP to be

$$\langle A^{-1}\rangle_{\rm var}=A_t^{-1}+\Lambda_t(AA_t^{-1}-1) \ , \tag{1.16}$$

that is, we employ Eq. (1.15a), but not Eq. (1.15b).

Of course, as can readily be verified, our analysis would be only trivially modified if we had employed Eq. (1.15b) in place of Eq. (1.15a). In Eq. (1.16), the quantity A_t^{-1} denotes the trial estimate of A^{-1}, *not* the reciprocal of an estimate A_t of A. For the sake of economy of notation, we are not using what would otherwise be a more explicit representation, namely, $(A^{-1})_t$. The single Lagrange multiplier or adjoint function being estimated by Λ_t is denoted by Λ because it obviously must be of the same functional character as A_t^{-1}, that is, it must be an operator; this notation (see Appendix A) distinguishes the present Lagrange multiplier Λ from the simpler types λ and L, where λ, encountered previously, represents a constant and L will represent a function.

The requirement that Eq. (1.16) be stationary yields

$$\delta A^{-1}+\delta\Lambda(AA^{-1}-1)+\Lambda A\delta A^{-1}=0 \ , \tag{1.17a}$$

where δA^{-1} is defined by[2]

[2]Since notation plays a significant role in our considerations, note that for any entity Q we always use δQ to denote the difference between Q evaluated with trial quantities and Q evaluated with the corresponding exact quantities. This does *not* imply that δQ always contains nonvanishing first-order quantities. In fact, it will be our purpose to construct F_v so that δF_v will contain only quantities which are of second order. In this connection note that the set of equations $B_i(\phi)=0$ which prescribe the set of quantities ϕ cannot always be derived from a Lagrangian—that is, the set cannot always be derived from a variational principle analogous to Hamilton's principle in classical mechanics. The inability to construct such a variational formulation of $B_i(\phi)=0$ can be completely consistent with the ability to construct, via our routine procedure, a variational principle $\langle F(\phi)\rangle_{\rm var}$ for any given desired $F(\phi)$ depending on the very same ϕ prescribed by $B_i(\phi)=0$. In other words, not all authors define the term "variational principle" as we do; failure to recognize this point could be a source of confusion in reading the literature. For example, Finlayson (1972) opens his Chap. 9 with the statement "Variational principles do not exist for many heat and mass transfer problems of interest." The material following this assertion of Finlayson's clearly indicates that he is concerned solely with the possibility of regarding the set $B_i(\phi)=0$ arising in heat and mass transfer problems as a set of Euler equations, as this term is defined in the calculus of variations (Morse and Feshbach, 1953, Sec. 3.1); one should *not* infer that variational principles (as we use the term) cannot be constructed for the many heat and mass transfer problems which cannot be formulated via Euler equations. Appendix C illustrates these remarks.

$$\delta A^{-1} \equiv A_t^{-1} - A^{-1} . \tag{1.17b}$$

[Since A is given, there is no such entity as δA and, in the present problem, one should *not* try to relate δA^{-1} to δA via $(\delta A)A^{-1} + A\delta A^{-1} = 0$ which would hold when A and A^{-1} both were permitted to vary, subject to the constraint (1.15a).] In Eq. (1.17a), the coefficient of $\delta\Lambda$ vanishes by virtue of Eq. (1.15a), and we deduce that

$$\Lambda A + 1 = 0 , \tag{1.18a}$$

implying that the exact adjoint function Λ is

$$\Lambda = -A^{-1} . \tag{1.18b}$$

Once again, the exact Lagrange multiplier is explicitly obtained but involves the unknown functions which we are seeking (here A^{-1}). The natural choice is

$$\Lambda_t = -A_t^{-1} , \tag{1.18c}$$

so that Eq. (1.16) becomes

$$\langle A^{-1} \rangle_{\text{var}} = A_t^{-1} - A_t^{-1}(AA_t^{-1} - 1) . \tag{1.19a}$$

Equation (1.19a) provides the desired variational estimate of A^{-1} in terms of a trial A_t^{-1}, where A_t^{-1} is presumably accurate to first order; we need *not* know the reciprocal of A_t^{-1}. If the reciprocal of A_t^{-1} *is* known, the VP in Eq. (1.19a) can be rewritten in a perhaps more usual and more easily grasped notation by setting $A_t^{-1} = B^{-1}$, where both B and B^{-1} are known, and, of course, $BB^{-1} = B^{-1}B = 1$. Thus Eq. (1.19a) becomes

$$\left\langle \frac{1}{A} \right\rangle_{\text{var}} = \frac{1}{B} + \frac{1}{B}(B - A)\frac{1}{B} . \tag{1.19b}$$

The VP in Eq. (1.19b) also can be obtained from the well-known identity

$$\frac{1}{A} = \frac{1}{B} + \frac{1}{B}(B - A)\frac{1}{A} , \tag{1.20}$$

where B is now any operator whose (left or right) inverse B^{-1} exists. This identity is obvious by inspection. When B is a first-order approximation to A, replacing $1/A$ in the last term on the right of Eq. (1.20) by $1/B$ causes an error of second order, since the factor $(B - A)$ in that term already is of first order. Therefore, the right side of Eq. (1.19b) is an estimate of $1/A$ accurate to second order when B is a first-order estimate of A. This essentially one-step derivation of the VP (1.19b), starting from the identity Eq. (1.20), is of the sort we characterized earlier as involving ingenuity, in contrast to our somewhat longer but routine derivation starting from Eq. (1.16).

We add that the identity in Eq. (1.20) is no more than Eq. (1.17a) itself; this can be seen by inserting Eqs. (1.15a), (1.17b), and (1.18b) and manipulating a bit. This fact—that quite generally the equation that defines the adjoint function, namely, the one where all first-order terms are set equal to zero, is an identity—is an illustration of the identities related to VP's discussed by Gerjuoy *et al.* (1972) and Gerjuoy, Rau *et al.* (1975).

The variational expression for A^{-1} derived in Eq. (1.19a) can be used in a variety of ways. If we set $A = H - E$, where H is the Hamiltonian and E the energy (with an infinitesimal imaginary part $i\epsilon$ if its real part is positive), A^{-1} is a Green's function. Equation (1.19a) could then be used to obtain a variational estimate of the Green's function. For A an $n \times n$ matrix, Eq. (1.19a) can be used to obtain a variational estimate of the inverse matrix. One would use a computer to obtain an estimate A_t^{-1} of A^{-1}. The estimate might be insufficiently accurate for the purposes at hand. A better (variational) estimate of A^{-1} could then be obtained from Eq. (1.19a) using the same computer, since round-off error in the multiplication of matrices will normally be much less than round-off error in the inversion of a matrix.

Finally, it is amusing to apply Eq. (1.19b) to something seemingly trivial, such as a variational estimation of $\frac{1}{11}$. (This example was suggested to us by L. Rosenberg.) With $\frac{1}{10}$ as a guessed trial estimate, we have $A = 11$ and $B = 10$ in Eq. (1.19b) so that the estimate we get for $\frac{1}{11}$ is 0.09. The initial fractional error of $\frac{1}{10}$ in the trial estimate has been reduced to a fractional error of $\frac{1}{100}$ in the variational result. Further, we can recognize that in this admittedly trivial case our VP is actually a bound by writing an alternative form for the identity in Eq. (1.20), namely,

$$\frac{1}{A} = \frac{1}{B} + \frac{1}{A}(B - A)\frac{1}{B} . \tag{1.21}$$

When we substitute this into Eq. (1.20), the error beyond the variational expression in Eq. (1.19b), that is, $1/A - \langle 1/A \rangle_{\text{var}}$, is of second order in $B - A$ and takes the explicit form

$$\frac{1}{B}(B - A)\frac{1}{A}(B - A)\frac{1}{B} .$$

In the special case that A is a real 1×1 matrix, i.e., a real number, this is of a well-defined sign, indeed is always ≥ 0 when $A > 0$; thus (1.19b) yields a lower bound to $1/A$ in the present case that A is a positive number, as the variational estimate 0.09 for $\frac{1}{11}$ illustrates. More generally, suppose A is a positive definite real symmetric matrix, so that $1/A$ also is a positive definite real symmetric matrix. Then if A_t^{-1} also is a real symmetric matrix, but not necessarily positive definite, the quantity $X^{\dagger}\langle A^{-1} \rangle_{\text{var}} X$—with X an arbitrary vector and $\langle A^{-1} \rangle_{\text{var}}$ given by Eq. (1.19a)—provides a variational lower bound, not merely a variational principle, for $X^{\dagger}A^{-1}X$.

4. Frequency eigenvalues

As another example, consider the determination of the eigenfrequencies of vibration ω_n on a stretched string of length s and variable density $\rho(x)$, held fixed at its end points, $x = 0$ and $x = s$. The tension T is assumed constant, and gravitational forces are neglected. Then for small displacements the differential equation for the wave amplitude $\phi_n(x)$ in the nth mode is (Rayleigh, 1877)

$$\frac{d^2\phi_n}{dx^2} + \frac{\omega_n^2 \rho(x)}{T}\phi_n = 0 , \tag{1.22a}$$

subject to the boundary conditions

$$\phi_n(0)=\phi_n(s)=0\ . \tag{1.22b}$$

We will assume $\phi_n(x)$ to be real, in view of its physical interpretation. However, even for real ϕ_n, Eqs. (1.22) do not suffice to specify the solution completely since the overall normalization of ϕ_n still has been left arbitrary. Our starting point for a variational estimate of ω_n^2, with the question of normalization temporarily ignored, will be

$$\langle\omega_n^2\rangle_{\rm var}=\omega_{nt}^2+\int_0^s dx\, L_{nt}(x)\left[\frac{d^2\phi_{nt}}{dx^2}+\frac{\omega_{nt}^2\rho}{T}\phi_{nt}\right]\ . \tag{1.23}$$

We use $L_{nt}(x)$ for our trial Lagrange multiplier, since it is now a *function* (rather than a constant or an operator), and we refer to $L_{nt}(x)$ as the trial adjoint function; ω_{nt}^2 is the trial ω_n^2, and $\phi_{nt}(x)$ is a trial estimate of a solution $\phi_n(x)$. The trial ϕ_{nt}, like ϕ_n itself, is required to satisfy Eq. (1.22b); the possibility of relaxing this requirement will be discussed in Sec. III.C.2. [A remark about ω_{nt}^2 is in order. We are treating it here as a direct estimate of ω_n^2 which is good to first order and independent of ϕ_{nt}. An alternative would be to regard it as a functional of ϕ_{nt}, since multiplication of Eq. (1.22a) by ϕ_n and integration affords an expression for ω_n^2 in terms of ϕ_n. A similar case will arise in a quantum-mechanical context; we defer a detailed discussion to Secs. I.B.5 and III.A.2.]

It then can be seen that Eq. (1.23) will be stationary for arbitrary variations of ω_{nt}^2 and $\phi_{nt}(x)$ about corresponding exact ω_n^2 and $\phi_n(x)$ if

$$1+\frac{1}{T}\int_0^s dx\, L_n(x)\rho(x)\phi_n(x)=0 \tag{1.24}$$

and also

$$\frac{d^2L_n}{dx^2}+\frac{\omega_n^2\rho}{T}L_n=0\ , \tag{1.25a}$$

together with the boundary conditions

$$L_n(0)=L_n(s)=0\ . \tag{1.25b}$$

Equation (1.24) is obtained by equating to zero the coefficient of $\delta\omega_n^2$ in the variation of Eq. (1.23); Eq. (1.25a) guarantees, after integration by parts, that terms proportional to $\delta\phi_n(x)$ vanish in the variation of Eq. (1.23); end-point contributions are eliminated by Eqs. (1.22b) and (1.25b). Equations (1.22) and (1.25) imply

$$L_n(x)=c_n\phi_n(x)\ , \tag{1.26a}$$

where c_n is some constant. Inserting Eq. (1.26a) into Eq. (1.24) yields

$$c_n=-\left[\int_0^s dx\,\rho\phi_n^2/T\right]^{-1}\ . \tag{1.26b}$$

An obvious first-order estimate L_{nt} of the exact $L_n(x)$ is an expression obtained from Eqs. (1.26) with ϕ_n replaced by ϕ_{nt}. As a result, we have the VP

$$\langle\omega_n^2\rangle_{\rm var}=-\left[\int_0^s dx\,\phi_{nt}\frac{d^2\phi_{nt}}{dx^2}\right]\left[\int_0^s dx\,\rho\phi_{nt}^2/T\right]^{-1}\ , \tag{1.27}$$

a result due to Rayleigh, though obtained of course quite differently. Historically the result in Eq. (1.27) probably represents the first truly useful variational principle from the computational point of view.

There can be no objection to the fact that we ignored normalization in deriving the VP for ω_n^2 given by Eq. (1.27); that Eq. (1.27) does indeed represent a VP is clear from its construction. Physically, the point is the somewhat surprising but very well-known fact that the frequency is independent of the amplitude. Had we specified the normalization, we of course again would have obtained a VP; see the following subsections.

5. Energy eigenvalues

Another instructive example, closely related to the one just given, is the determination of bound-state energy eigenvalues in quantum mechanics. The eigenfunction ϕ associated with the eigenvalue E obeys the Schrödinger equation (1.3a), where H is Hermitian. In general, H as well as ϕ may be multicomponented and complex; for instance, H may be the Dirac Hamiltonian. In the present section, however, we shall assume H and ϕ to be purely real; the extension to problems involving complex ϕ and/or H is given in Sec. III.B. On the other hand, in this section we do permit both H and/or ϕ to be multicomponented, provided that all these components are purely real.

From the results of Sec. I.B.4 we can infer that for the purpose of constructing a VP for E it is not necessary to specify the normalization of ϕ; the correctness of this inference can be verified (see below). Nevertheless, now we will include the normalization condition (1.3b) in the specification of ϕ, to illustrate the role of the Lagrange multiplier associated with the normalization constraint. We further remark that for notational convenience in this subsection we have dropped the subscript n (carried in Sec. I.B.4) identifying the eigenvalue. Since H may have many discrete eigenvalues, in employing the VP to estimate the energy it is, of course, necessary that the trial function ϕ_t employed correspond to an appropriate ϕ; for example, the choice of $\phi_t(\mathbf{r})$ should be made with the quantum numbers of the exact ϕ associated with the desired E in mind. To construct the desired VP, however, restrictions on ϕ and ϕ_t delimiting E to a specific unique value are irrelevant; the resultant VP has the same form for every bound state E.

With the foregoing considerations in mind, we start from

$$\langle E\rangle_{\rm var}\equiv E_v=E_t+L_t^\dagger[(H-E_t)\phi_t]+\lambda_t(\phi_t^\dagger\phi_t-1)\ . \tag{1.28}$$

Integration over $\mathbf{r}$ is implied in the scalar product $\phi_t^\dagger\phi_t$; and, when ϕ is multicomponented, summation over the discrete indices of the components of ϕ also is implied. The term involving L_t in Eq. (1.28) is similarly interpreted; in particular, when $\phi(\mathbf{r})$ is multicomponented and regarded as a column vector, L_t also is a column vector whose elements are ordinary functions of $\mathbf{r}$. Whether or not ϕ is multicomponented, λ must be merely an ordinary (real) number in Eq. (1.28), because E is merely a number.

Varying Eq. (1.28) and using Eqs. (1.3), we obtain the requirement

$$\delta E_v=\delta E+L^\dagger[(H-E)\delta\phi]-\delta E(L^\dagger\phi)+\lambda[\phi^\dagger\delta\phi+(\delta\phi)^\dagger\phi]=0 \ , \tag{1.29}$$

where $\delta E=E_t-E$. To obtain the condition for the coefficient of $\delta\phi$ to vanish in Eq. (1.29), the operation $(H-E)$ on $\delta\phi$ must be transferred to L. In other words, we must assume that it is legitimate to write

$$L^\dagger[(H-E)\delta\phi]=[(H-E)L]^\dagger\delta\phi \ . \tag{1.30}$$

Equation (1.30) represents a boundary condition.

It is conceivable that there are problems for which one cannot obtain generalized Lagrange multipliers which satisfy the differential equations and relations such as (1.30), which arise in the course of the development of a VP. If so, our procedure for obtaining a VP will be useless. We ourselves have not encountered a problem for which the generalized Lagrange multipliers do not exist.

In Eq. (1.29) it also is true that

$$(\delta\phi)^\dagger\phi=\phi^\dagger\delta\phi \tag{1.31}$$

in view of the definition of the scalar product implied by the notation. Thus, using Eqs. (1.30) and (1.31), the conditions for the vanishing of δE_v given by Eq. (1.29) become

$$(H-E)L+2\lambda\phi=0 \ , \tag{1.32a}$$

$$1-L^\dagger\phi=0 \ , \tag{1.32b}$$

after equating to zero the coefficients of $\delta\phi$ and δE, respectively. Multiplying Eq. (1.32a) on the left by $\phi^\dagger$, we find, using Eq. (1.3a), that

$$\lambda=0 \ . \tag{1.33a}$$

With $\lambda=0$ we see from Eq. (1.32a) that L is a solution of the Schrödinger equation corresponding to the same bound-state energy E as ϕ itself; moreover, it will be made apparent below that to avoid difficulties with the boundary conditions L should be quadratically integrable. In other words, L should be restricted to the quadratically integrable eigenfunctions of the Schrödinger equation at energy E. If E is nondegenerate, we therefore infer that

$$L=c\phi \ , \tag{1.33b}$$

where c is a constant whose value must be

$$c=1 \ , \tag{1.33c}$$

by virtue of Eq. (1.32b) and the normalization condition (1.3b). If E is degenerate, Eq. (1.33b) continues to be a valid solution of Eq. (1.32a) with $\lambda=0$, but is no longer the most general solution. For degenerate E, projections on the other eigenfunctions belonging to E can be added to the right-hand side of Eq. (1.33b); however, these additional projections merely complicate the VP without altering it in any essential way (see below).

The solutions (1.33) for the exact L and λ in the nondegenerate case suggest the trial estimates

$$\lambda_t=0 \ , \tag{1.34a}$$

$$L_t=\phi_t \ . \tag{1.34b}$$

Substituting Eqs. (1.34) into Eq. (1.28) yields

$$\langle E\rangle_{\rm var}=E_t(1-\phi_t^\dagger\phi_t)+\phi_t^\dagger H\phi_t \ , \tag{1.35}$$

which is the desired VP. We mention that the results in Eqs. (1.33) are not typical; as we shall see, in VP's for other $F(\phi)$ than the eigenvalues of H, the exact Lagrange multiplier associated with the normalization constraint generally does not vanish, nor does one normally have $L=\phi$.

So far we have not specified the normalization of ϕ_t. The constraint (1.3b) on the exact ϕ does not require ϕ_t to be normalized, but does imply that if ϕ_t differs from ϕ by first order, then, to that order,

$$\phi_t^\dagger\phi_t-1=(\phi+\delta\phi)^\dagger(\phi+\delta\phi)-1=\phi^\dagger\delta\phi+(\delta\phi)^\dagger\phi \ . \tag{1.36}$$

In other words, Eq. (1.3b) implies that $\phi_t^\dagger\phi_t$ should differ from unity by a quantity of first order. Using Eq. (1.36), one can easily verify that Eq. (1.35) indeed is a variational estimate of E, that is, that it differs from the exact E by quantities of second order when E_t and an unnormalized ϕ_t in Eq. (1.35) are accurate to first order. Often, when ϕ is constrained by Eq. (1.3b), one also requires that

$$\phi_t^\dagger\phi_t-1=0 \ , \tag{1.37a}$$

in which event, according to Eq. (1.36), we have

$$\phi^\dagger\delta\phi+(\delta\phi)^\dagger\phi=0 \ . \tag{1.37b}$$

Substituting Eq. (1.37a) into Eq. (1.35) gives the usual well-known form of the VP for E, namely,

$$\langle E\rangle_{\rm var}=\phi_t^\dagger H\phi_t \ , \tag{1.38}$$

with ϕ_t constrained by Eq. (1.37a). [When Eq. (1.37a) is satisfied, a more consistent procedure might be to start with Eq. (1.28), but with the λ_t term omitted. One again arrives at Eq. (1.38).]

6. Generalizations of variational principles for energy

The above derivation of Eq. (1.38) obviously is much more complicated than usual or necessary, but it does illustrate our assertion that our routine procedure will lead

to the desired VP, though it sometimes may do so rather clumsily. The derivation also indicates various possible generalizations of Eqs. (1.35) and (1.38). For instance, if Eq. (1.37a) is not required, and λ_t differs from zero by first order, then by virtue of Eq. (1.36) the last term in Eq. (1.28) makes a contribution of second order. This is consistent with the requirement that Eq. (1.28) remain a VP for any trial estimates L_t and λ_t of the auxiliary functions differing by first order from their exact values in Eqs. (1.33). Similarly, it can be seen that when E is degenerate the use of more general quadratically integrable solutions to Eqs. (1.32) than Eqs. (1.33b) and (1.33c) for L maintains Eq. (1.28) as a VP, provided only that the corresponding L_t differs by first order from the more general L. Suppose, for instance, Eqs. (1.33b) and (1.33c) are replaced by

$$L=\phi+b\psi \ , \tag{1.39a}$$

where b is any constant and ψ is a normalized eigenfunction of the same energy E orthogonal to ϕ; evidently Eq. (1.39a), taken together with Eq. (1.33a), continues to satisfy Eqs. (1.32). With Eq. (1.39a), the obvious generalization of Eq. (1.34b) would be

$$L_t=\phi_t+b_t\psi_t \ , \tag{1.39b}$$

where ψ_t is a trial estimate of ψ, and b_t is a trial constant. Substituting Eq. (1.39b) into Eq. (1.28) and retaining Eq. (1.34a) lead to

$$\langle E\rangle_{\rm var}=E_t(1-\phi_t^\dagger\phi_t)+\phi_t^\dagger H\phi_t+b_t\psi_t^\dagger(H-E_t)\phi_t \ . \tag{1.40}$$

But the b_t term in (1.40) can be seen to make a contribution of second order if we remember that the exact ψ satisfies the Schrödinger equation (1.3a) and that ψ is orthogonal to the exact ϕ. In other words, the more general solution (1.39a) of Eqs. (1.32) yields essentially the same VP (1.35) as did the solutions (1.33) employed earlier. Actually, the above argument starting from Eq. (1.39a) is merely a specific illustration of the readily verifiable fact that Eqs. (1.30) and (1.32) really are sufficient to guarantee that Eq. (1.28) is a VP, in which event it must be the case that $\langle E\rangle_{\rm var}$ is a second-order estimate of E when L_t and λ_t differ by first order from any corresponding L and λ satisfying Eqs. (1.32).

Returning to the nondegenerate case, we already have mentioned that omitting the normalization condition (1.3b) on ϕ does lead to a VP for E. In particular, with this omission, and after dropping the λ_t term in Eq. (1.28) and finding that Eq. (1.33c) is replaced by $c=(\phi^\dagger\phi)^{-1}$, we are led to the well-known generalization of Eq. (1.38) for unnormalized ϕ_t, namely,

$$\langle E\rangle_{\rm var}=\frac{\phi_t^\dagger H\phi_t}{\phi_t^\dagger\phi_t} \ . \tag{1.41}$$

Note that Eq. (1.41) is not identical in form with Eq. (1.35), if only because in Eq. (1.35) some first-order estimate E_t of E is required. More importantly, because the numerator and denominator of (1.41) happen to contain the same number of ϕ_t factors, Eq. (1.41) holds for a ϕ_t of unrestricted normalization, that is, for ϕ_t close to ϕ in form but not necessarily close to ϕ in normalization. In Eq. (1.35), as in most forms of VP's, it is assumed that ϕ_t differs by first order from a normalized ϕ, so that Eq. (1.36) holds. For ϕ_t consistent with Eq. (1.36), the VP (1.41) reduces to

$$\langle E\rangle_{\rm var}=\phi_t^\dagger H\phi_t+(\phi_t^\dagger H\phi_t)(1-\phi_t^\dagger\phi_t) \ , \tag{1.42}$$

after we write

$$\phi_t^\dagger\phi_t=1-(1-\phi_t^\dagger\phi_t)$$

in the denominator of Eq. (1.41), and then expand in powers of $(1-\phi_t^\dagger\phi_t)$. Equation (1.42) also is not identical in form with Eq. (1.35). Nevertheless, both Eqs. (1.35) and (1.42), being VP's, must yield estimates of E accurate to second order, that is, Eqs. (1.35) and (1.42) differ by at most second-order terms. Finally, had we, in writing Eq. (1.28), considered E_t not to be an independent first-order estimate of E but as the value given by $\phi_t^\dagger H\phi_t$, the subsequent analysis would have led to Eq. (1.42), as can be readily verified.

The energy eigenvalue VP's of this section and of Sec. I.B.5 can be further generalized to non-Hermitian Hamiltonians and to more complicated boundary conditions, as Appendix D illustrates.

7. Electric field

Our final illustration in this section will be concerned with a case in which the basic equations for the adjoint function are rather more difficult to solve than heretofore. Suppose that we seek a VP for the electric field at some given point $\mathbf{r}_0$ between the two conductors of a condenser with the plates held at some known potentials. In other words, we shall consider the problem of variationally estimating $\nabla\phi(\mathbf{r}_0)$, with $\phi(\mathbf{r})$, the potential, defined by $\nabla^2\phi(\mathbf{r})=0$ together with given values of ϕ at the boundaries. Then our starting point for a variational estimate of $\nabla\phi(\mathbf{r}_0)$, assuming $\phi_t(\mathbf{r})$ satisfies the boundary conditions on $\phi(\mathbf{r})$, is

$$\langle\nabla\phi(\mathbf{r}_0)\rangle_{\rm var}=\nabla\phi_t(\mathbf{r}_0)+\int d\mathbf{r}\,\mathbf{\Lambda}_t(\mathbf{r}_0,\mathbf{r})\nabla^2\phi_t(\mathbf{r}) \ , \tag{1.43a}$$

where (consistent with the notation described in Appendix A) we now denote the adjoint function by $\mathbf{\Lambda}$, because $\mathbf{\Lambda}$ obviously is a function of two variables, namely, $\mathbf{r}_0$ and $\mathbf{r}$. The quantity $\mathbf{\Lambda}$ and its estimate $\mathbf{\Lambda}_t$ are written as vectors because Eq. (1.43a) clearly requires that $\mathbf{\Lambda}_t$ (and therefore also $\mathbf{\Lambda}$) be an ordinary three-dimensional vector, like $\nabla\phi_t$ and $\nabla\phi$; in other words, $\mathbf{\Lambda}$ actually denotes a collection of three adjoint functions, the components of $\mathbf{\Lambda}$. To avoid notational confusion, it is convenient to concentrate on one component of $\nabla\phi_t(\mathbf{r}_0)$—the z component, say—where the z direction will be quite arbitrary.

In Eq. (1.43a) it is further convenient to write

$$\begin{aligned}\nabla\phi_t(\mathbf{r}_0)&=\int d\mathbf{r}\,\delta(\mathbf{r}-\mathbf{r}_0)\nabla\phi_t(\mathbf{r})\\&=-\int d\mathbf{r}[\nabla\delta(\mathbf{r}-\mathbf{r}_0)]\phi_t(\mathbf{r}) \ ,\end{aligned} \tag{1.43b}$$

so that both terms on the right-hand side of Eq. (1.43a) are represented as integrals over $\mathbf{r}$ with integrands containing $\phi_t(\mathbf{r})$.

Skipping details of the many manipulations involved, we find that the adjoint function is defined by

$$\nabla^2\Lambda_z(\mathbf{r}_0,\mathbf{r})=\frac{\partial}{\partial z}\delta(\mathbf{r}-\mathbf{r}_0) \tag{1.44a}$$

in the volume, together with

$$\Lambda_z(\mathbf{r}_0,\mathbf{r})=0 \tag{1.44b}$$

for $\mathbf{r}$ on the conductor surfaces (and at infinity); of course, the derivatives in Eq. (1.44a) are with respect to $\mathbf{r}$, not $\mathbf{r}_0$. We note that this is an example where the exact Lagrange multipliers do not depend on the functions ϕ. That Λ_z in this problem is a Green's function is evident from Eqs. (1.44).

Therefore, the desired VP for the z component of $\nabla\phi(\mathbf{r}_0)$ is given by the z component of Eq. (1.43a) with $\Lambda_{zt}(\mathbf{r}_0,\mathbf{r})$ a first approximation to the unique $\Lambda_z(\mathbf{r}_0,\mathbf{r})$ satisfying Eqs. (1.44). Naturally, we are not likely to know Λ_z exactly; usually we are no more able to solve Eqs. (1.44) for the exact Λ_z than we are able to solve for the exact ϕ. However, it is possible to deduce some further information about the exact Λ_z from Eqs. (1.44), thereby somewhat easing the problem of finding a reasonable trial estimate Λ_{zt} of Λ_z, but we will not go into this.

Had we been interested in a VP for the potential $\phi(\mathbf{r}_0)$ at the point $\mathbf{r}_0$, rather than for the electric field at $\mathbf{r}_0$, we would have started from

$$\langle\phi(\mathbf{r}_0)\rangle_{\text{var}}=\phi_t(\mathbf{r}_0)+\int\Lambda_t(\mathbf{r}_0,\mathbf{r})\nabla^2\phi_t(\mathbf{r})d\mathbf{r} \ ,$$

with Λ_t now a single function rather than a collection of three functions. We will not complete the derivation of this VP, but note that Sec. IV.A treats the very similar problem of finding a VP for the Schrödinger wave function $\phi(\mathbf{r}_0)$ at the point $\mathbf{r}_0$. Further, in a spherical condenser, where $\phi(\mathbf{r})$ is known to have spherical symmetry and the electric field is known to be radial, $\nabla^2\phi=0$ can be replaced by the one-dimensional equation

$$\frac{1}{r^2}\frac{d}{dr}\left[r^2\frac{d\phi}{dr}\right]=0 \ , \tag{1.45}$$

where the origin is at the condenser center. In the case of a spherical condenser, therfore, we would expect that there exists a less complicated derivation (than the above) of a VP for the electric field. Such a derivation, based on Eq. (1.45), has been carried out explicitly in Gerjuoy, Rau, and Spruch (1972).

II. GENERAL CONSTRUCTION OF VARIATIONAL PRINCIPLES FOR REAL QUANTITIES

A. Introduction

In this section, we restrict ourselves to real quantities, and give a somewhat formal statement of the general construction of VP's. Even for solely real quantities, the functional character of the quantities and equations of interest can be so diverse and so complicated that it would be extremely difficult, and perhaps impossible, to write down compactly a set of equations valid for all cases. We discuss the construction of VP's under circumstances much more general than any considered above. The discussion may appeal to the more formally minded reader; a rereading of some of the illustrative examples of Sec. I.B should help bring out the basic features of the general procedure. Other readers could very well glance over this material cursorily, since it is not required for an understanding of the material in later sections. To an extent, the same applies to some of the formal material in Sec. III.

B. Basic equations

Let the set of quantities ϕ_t be a trial estimate of the exact but unknown set ϕ satisfying Eq. (1.1). In the most familiar variational principles, such as the Rayleigh-Ritz principle for finding the eigenvalues E of Eq. (1.3a), the ϕ_t are often the only trial quantities. However, there is no reason why the variational estimate of an arbitrary $F(\phi)$ should not involve trial quantities other than the ϕ_t. In particular, we shall let the set $\mathscr{L}_t$ of "auxiliary functions" denote trial estimates of a set $\mathscr{L}$ which is still to be specified, possibly (but not necessarily) independent of ϕ and of one another. We will consistently refer to these auxiliary functions as adjoint functions, a terminology also adopted by Finlayson (1972), Atherton and Homsy (1975), and others. The notation $\mathscr{L}$ anticipates the Lagrange undetermined multiplier aspect of each member of the set. Then the functional $F_v \equiv F_v(\phi_t,\mathscr{L}_t)\equiv\langle F(\phi)\rangle_{\text{var}}$ is a VP for the desired $F(\phi)$ if the following two conditions hold:

(i) $F_v(\phi_t,\mathscr{L}_t)$ coincides with $F(\phi)$ whenever the trial ϕ_t coincide with the exact ϕ, whether or not the trial $\mathscr{L}_t$ coincide with the exact $\mathscr{L}$.

(ii) in the error

$$\delta F_v=F_v(\phi_t,\mathscr{L}_t)-F(\phi) \ , \tag{2.1a}$$

the total first-order error, composed of terms proportional to $\delta\phi$ and $\delta\mathscr{L}$, vanishes.

Here, we are defining

$$\delta\phi\equiv\phi_t-\phi, \ \ \delta\mathscr{L}\equiv\mathscr{L}_t-\mathscr{L} \ . \tag{2.1b}$$

We come now to the basic problem, the construction of F_v. For simplicity, not necessity, we begin with a restricted but nonetheless instructive and easily generalized case, which will be followed by a more general and formal consideration. Specifically, we assume (a) that $\phi(\mathbf{r})$ is a single function, (b) that the F under consideration is a single numerical quantity independent of $\mathbf{r}$, as, for example, F of Eq. (1.2), and (c) that there are precisely two constraints, of particular kinds which we now describe. The first type

of constraint provides a *single* restriction on ϕ, that is, $B(\phi)$ is in the space of (real) numbers; this constraint is written in the form

$$B_1(\phi)=0 \ . \tag{2.2a}$$

An example of this sort of constraint is the normalization condition Eq. (1.3b); another example is the requirement that $\phi(\mathbf{r})=0$ at the point $\mathbf{r}=0$. The second type of constraint, with $B(\phi)$ in the space of functions, provides one restriction on ϕ for each value of $\mathbf{r}$ in some allowed *range* of $\mathbf{r}$, but the restrictions for different values of $\mathbf{r}$ are essentially independent; in other words, such a constraint, which we write in the form

$$B_2(\phi)=0 \ , \tag{2.2b}$$

really represents a continuous set of $\mathbf{r}$-dependent restrictions on ϕ, one restriction for each value of $\mathbf{r}$. An example is Eq. (1.3a), the Schrödinger equation; at each point $\mathbf{r}$, Eq. (1.3a) provides a purely real restriction on $\phi(\mathbf{r})$ which is essentially independent of the restrictions provided at other points $\mathbf{r}'\neq\mathbf{r}$. Note that although Eq. (1.3a) happens to be linear in ϕ, we do *not* assume that B_2 must be expressible in the form $\theta\phi$, where θ is a linear operator.

Now consider the expression

$$F_v=F(\phi_t)+\lambda_t B_1(\phi_t)+L_t^{\dagger}B_2(\phi_t) \ , \tag{2.3}$$

where λ_t and L_t are approximations to exact quantities λ and L whose defining equations still must be ascertained, and where the last term of Eq. (2.3) utilizes the matrix notation introduced in Sec. I. It is clear from the construction of Eq. (2.3) that λ and L are analogs of the Lagrange multipliers customarily employed when treating problems involving constraints—for example, in extremum problems in elementary calculus and in the calculus of variations (Courant and Hilbert, 1937, pp. 164 and 165 and 216–222) or in applications of Hamilton's principle to problems in mechanics (Goldstein, 1980, p. 47). We construct F_v of Eq. (2.3) by first adding all constraints—multiplied by their respective Lagrange multipliers—to the original $F(\phi)$, after which we replace all unknown quantities by their trial estimates. As will be seen in a moment, the matrix notation in Eq. (2.3) expresses precisely the sum we require.

Because $F(\phi_t)$ has been assumed to be a single numerical quantity independent of $\mathbf{r}$, while Eq. (2.2a) provides a single restriction on ϕ, the Lagrange multiplier in Eq. (2.3) associated with the constraint Eq. (2.2a) must itself be a single numerical quantity independent of $\mathbf{r}$. For this reason, instead of employing the generic Lagrange multiplier symbol $\mathscr{L}$ introduced earlier, we denote the Lagrange multiplier associated with Eq. (2.2a) by λ, to emphasize its close analogy with the numerical Lagrange undetermined multipliers used in elementary calculus. On the other hand, since Eq. (2.2b) provides a continuous set of $\mathbf{r}$-dependent restrictions on ϕ, it must be associated in Eq. (2.3) with a correspondingly continuous set of Lagrange multipliers, that is, a function, which we denote by $L(\mathbf{r})$ (again instead of the generic $\mathscr{L}$) to emphasize the functional dependence of L on $\mathbf{r}$. Thus in Eq. (2.3) the contribution made by the constraint Eq. (2.2b) is a sum, or rather an integral, over constraints multiplied by their associated Lagrange multipliers. The integral is over the range of allowed variables $\mathbf{r}$—that is,

$$L_t^{\dagger}B_2(\phi_t)=\int d\mathbf{r}\, L_t(\mathbf{r})\{[B_2(\phi_t)]_{\mathbf{r}}\} \ .$$

The notation for our generalized Lagrange multipliers $\mathscr{L}$ is further discussed in Appendix A.

Because of Eqs. (2.2), the formula (2.3) for $\langle F\rangle_{\rm var}$ evidently obeys condition (i) quoted at the beginning of this section. Therefore, Eq. (2.3) will be a VP for $F(\phi)$ if the Lagrange multipliers λ and L can be chosen so that condition (ii) above is satisfied. Generally, it is not possible to find λ and L exactly, but it is possible to ascertain the relations specifying the exact λ and L. Note that because F_v in Eq. (2.3) is linear in the $\mathscr{L}$ by construction, the second-order difference between F_v and $F(\phi)$ will contain terms of order $(\delta\phi)^2$ and $(\delta\phi)(\delta\mathscr{L})$, but not $(\delta\mathscr{L})^2$; this assertion is valid in more complicated problems as well, because we always insert the Lagrange multipliers linearly into the analog of Eq. (2.3).

To determine the needed relations specifying λ and L of Eq. (2.3), it first is necessary to obtain and manipulate appropriately the first variation of Eq. (2.3). We have by definition (see footnote 2 above):

$$\delta[B_1(\phi)]=B_1(\phi_t)-B_1(\phi) \ . \tag{2.4a}$$

In the variation δF_v defined by Eq. (2.1a) we have, therefore,

$$\begin{aligned}\lambda_t B_1(\phi_t)&=(\lambda+\delta\lambda)\{B_1(\phi)+\delta[B_1(\phi)]\}\\&=\lambda\delta[B_1(\phi)]+(\delta\lambda)\delta[B_1(\phi)] \ .\end{aligned}$$

In the last step we used Eq. (2.2a). Note that $\delta\lambda$ appears only in a second-order term. Similar relations hold for the variations of the term in Eq. (2.3) involving the constraint Eq. (2.2b). Therefore, neglecting terms of second order, we obtain

$$\delta F_v=\delta F+\lambda\delta[B_1(\phi)]+L^{\dagger}\delta[B_2(\phi)] \ , \tag{2.4b}$$

where, of course,

$$\delta F=F(\phi_t)-F(\phi) \ . \tag{2.4c}$$

Note that Eq. (2.4b) also follows from a direct first-order variation of Eq. (2.3).

Each of the terms on the right-hand side of Eq. (2.4b) can be expressed in terms of $\delta\phi$, the only quantity we can arbitrarily vary in F and the B's. If we are to succeed in obtaining a VP $\langle F(\phi)\rangle_{\rm var}$, it must be possible to combine these first variations so that Eq. (2.4b) takes the form

$$\delta F_v=K^{\dagger}\delta\phi\equiv\int d\mathbf{r}\, K(\mathbf{r})\delta\phi(\mathbf{r}) \ , \tag{2.5a}$$

where $K(\mathbf{r})$ must be a function and the operation involved in Eq. (2.5a) be only one of multiplication of $\delta\phi$ by K; it must not, for example, be a derivative. Only if Eq. (2.4b) has been reduced to the form Eq. (2.5a), with $K(\mathbf{r})$ a function, is it possible to infer a condition, namely,

$$K(\mathbf{r})=0 \ , \tag{2.5b}$$

sufficient to guarantee that δF_v vanishes for all allowed $\delta\phi(\mathbf{r})$, that is, sufficient to guarantee that Eq. (2.3) really is a VP. Moreover, although the allowed trial functions ϕ_t are not wholly unrestricted [for example, it may have been presumed *a priori* that $\phi_t(\mathbf{r})$ are continuous and continuously differentiable], and although expressing Eq. (2.4b) in the form (2.5a) typically implies that $\delta\phi(\mathbf{r})$ must satisfy some boundary conditions (see immediately below), nevertheless, in general the allowed $\phi_t(\mathbf{r})$ remain sufficiently unrestricted that $\delta\phi(\mathbf{r})$ can be regarded as essentially arbitrary at every $\mathbf{r}$ (see Secs. III.B and III.C for more details). Therefore, for the purpose of guaranteeing that δF_v vanishes for all allowed $\delta\phi(\mathbf{r})$, Eq. (2.5b) generally is necessary, not merely sufficient; that is, only if Eq. (2.5b) can be satisfied will it generally be true that Eq. (2.3) is a VP.

There remains the problem of putting Eq. (2.4b) into the form Eq. (2.5a). To do so, it normally is necessary to perform manipulations which are valid only when $\delta\phi$ and the Lagrange multipliers in Eq.(2.4b) satisfy certain "boundary conditions"; simultaneously, these manipulations normally will cause $K(\mathbf{r})$ to involve operations on λ and L, so that Eq. (2.5b) will be, for example, a differential equation for $L(\mathbf{r})$ involving λ as a parameter. To illustrate these assertions, consider for instance that Eq. (2.2a) is the normalization condition (1.3b), while $B_2(\phi)=\theta\phi$, where θ is a linear operator [such as $(H-E)$] independent of ϕ. In this event, the term in Eq. (2.4b) involving the variation of B_1 immediately makes a contribution of the required form (2.5a) to δF_v, but the term involving B_2 makes the contribution

$$L^\dagger\delta[\theta\phi]=L^\dagger[\theta\delta\phi]$$

to δF_v. Then, assuming there are no manipulation problems with δF, it will be possible to put Eq. (2.4b) into the form Eq. (2.5a)—with $K(\mathbf{r})$ a function, not an operator—if and only if

$$L^\dagger[\theta\delta\phi]=[\hat{\theta}L]^\dagger\delta\phi \ , \tag{2.6a}$$

that is, in less condensed notation, if and only if

$$\int d\mathbf{r}\,L(\mathbf{r})[\theta\delta\phi(\mathbf{r})]=\int d\mathbf{r}[\hat{\theta}L(\mathbf{r})]\delta\phi(\mathbf{r}) \ , \tag{2.6b}$$

where $\hat{\theta}$ denotes some other operator, the real adjoint of θ. Having replaced the left-hand side of Eq. (2.6a) by its right-hand side, we see that the term involving B_2 contributes a term $\hat{\theta}L$ to $K(\mathbf{r})$ in Eq. (2.5b). We term Eqs. (2.6a) or (2.6b) "boundary conditions" because, as will be seen, in many common circumstances (when θ is $H-E$, for instance) the restrictions Eqs. (2.6) impose on $\delta\phi(\mathbf{r})$ and/or $L(\mathbf{r})$ are boundary conditions of the usual type (at the origin and at infinity, for instance). More specifically, in typical cases the analogs of Eqs. (2.6) amount to the assertion that orders of integration can be inverted, or that boundary terms resulting from integration by parts can be ignored. Presumably the trial functions $\phi_t\neq\phi$ cannot satisfy all the constraints (1.1) as well as the boundary conditions on ϕ, often not explicitly included in Eq. (1.1). Usually, but not always, the trial ϕ_t are required to satisfy the boundary conditions on ϕ and some of the more readily attainable constraints (1.1). For instance, when the exact ϕ in Eq. (1.2) is required to satisfy Eqs. (1.3a) and (1.3b), the trial ϕ_t normally would be required to satisfy the normalization condition (1.3b); if the wave function ϕ describes zero angular momentum scattering by a central potential, the trial $r\phi_t$ might be required to behave like r at the origin, because the exact $r\phi$ is known to behave so at the origin. At any rate, once the class of allowable trial functions ϕ_t is specified for any given problem, the set of equations comprising Eqs. (2.6) plays the role of boundary conditions, in that this set delimits (and may make unique) the set of adjoint functions $\mathscr{L}$ that are the solutions of Eq. (2.5b). We shall elaborate on the remarks of this paragraph in Sec. III.C.

It now should be quite apparent how we proceed to derive a VP for $F(\phi)$ when (still assuming all quantities are real) the assumptions (a),(b),(c) quoted in the paragraph immediately following Eq. (2.1b) fail to hold, as, for example, when F is not restricted to be a single numerical quantity independent of $\mathbf{r}$, or when there are a large number of constraints $B_i(\phi)=0$ not necessarily of the forms of Eqs. (2.2). We first construct the analog of Eq. (2.3) by adding to $F(\phi_t)$ a sum $\sum\mathscr{L}_{it}^\dagger B_i(\phi_t)$ over all constraints B_i, where the functional character (constant, function of one variable, function of two variables, collection of functions, etc.) of the Lagrange multiplier $\mathscr{L}_i$ will follow immediately from[3] the functional character of $F(\phi)$ and $B_i(\phi)$; correspondingly, for any given B_i the term $\mathscr{L}_{it}^\dagger B_i(\phi_t)$ will represent an integral over all allowed ranges of the free variables in B_i. Next, using only for the present paragraph the notation $(2.XX)'$ to indicate the analog of one of the equations $(2.XX)$, we compute Eq. (2.4b)', from which the variations $\delta\mathscr{L}_i$ again will disappear, because of Eq. (1.1). Finally, we recast Eq. (2.4b)' into Eq. (2.5a)', wherein $K^\dagger\delta\phi$ will now denote summation and/or integration as appropriate. In this way, we will derive Eq. (2.5b)', which again will be a necessary and sufficient condition for the desired VP, but which now may represent a collection of equations. The solutions to this Eq. (2.5b)' will determine the set of exact Lagrange multipliers whose estimates[4] $\mathscr{L}_t$ appear in Eq. (2.3)', and

[3]The $\mathscr{L}_i$ map the space of entities $B_i(\phi)$ onto the space of entities $F(\phi)$; $\mathscr{L}_i:B_i(\phi)\rightarrow F(\phi)$. Thus, in the example considered above, when $F(\phi)$'s are real and $B_1(\phi)$'s are also real, the corresponding Lagrange multiplier, λ, is also in the space of reals, but when $F(\phi)$'s are real and $B_2(\phi)$'s are functions of one variable, the Lagrange multiplier $L(\mathbf{r})$ is also in the space of functions of one variable.

[4]Note that implicit in our use of trial estimates $\mathscr{L}_t$ is the assumption that the space of $\mathscr{L}_i$'s defined by $\mathscr{L}_i:B_i(\phi)\rightarrow F(\phi)$ admits a topology so that there is the concept of a neighborhood of entities $\mathscr{L}_{it}$ around the $\mathscr{L}_i$.

these solutions will be subject to a set of boundary conditions (2.6)′ on $\mathscr{L}$ and $\delta\phi$, which arise during the course of the manipulations yielding Eq. (2.5a)′ from Eq. (2.4b)′.

In the above discussion it was assumed that one can obtain approximations L_t's to the L's, good to first order, once the defining equations for the L's have been obtained. In special cases it may not be simple to do so. A familiar example for which the determination of the L_t's becomes particularly crucial is the one of VP's for calculating matrix elements of operators between solutions of the Schrödinger equation; only relatively recently has the resolution of all the complications that arise, particularly in dealing with complex functions, been achieved (Gerjuoy *et al.*, 1974). We note that the above remarks relating to the construction of VP's make no restrictions on the nature of ϕ, $F(\phi)$, and $B(\phi)$, and that they therefore are applicable when complex quantities, nonlinear functionals, or integral operators are involved (Rau *et al.*, 1978). It is also of interest that in the case of nonlinear problems, the method makes contact (Rau, 1976) with methods of functional analysis and invariant imbedding (Kalaba, 1959; Bellman and Kalaba, 1959; Bellman and Wing, 1975).

With the preceding remarks, we have completed the formal presentation of our general procedure; the remaining sections of this paper illustrate how easily it can be applied to a variety of problems. As these illustrations and those in Sec I.B show, in practice it is not difficult to translate Eqs. (2.5) and (2.6) into actual differential or integral equations for $\mathscr{L}$ together with appropriate boundary conditions on $\mathscr{L}$. In fact, in many of the illustrative examples considered, it further turns out that the equations for $\mathscr{L}$ can be solved exactly. In later sections (especially IV.C) we will consider cases where this is not true and point out prescriptions for how approximate solutions, $\mathscr{L}_t$, are to be obtained.

Our illustrative derivations of VP's in Sec. I and in later sections proceed so routinely that it should be manifest that the above formal presentation suffices for successful application of our general procedure. Nevertheless, our presentation up to this juncture has glossed over a number of questions bearing on the precise implications of the boundary conditions imposed on the auxiliary functions, $\mathscr{L}$, the necessity versus the sufficiency of Eqs. (2.5b) and (2.6) or their analogs, and the need for complete specification of the wave function, especially the phase of the wave function. These and related topics will be examined in Sec. III.

III. COMPLEX QUANTITIES AND FORMAL EXTENSIONS

In Sec. III.A we generalize the discussion in Sec. II.B so that it can be applied to problems involving complex quantities. The succeeding subsections III.B–III.D then examine a number of questions concerning our formal presentation which were glossed over in Sec. II, especially concerning the specification of the wave function. In particular, Sec. III.B examines the relationship between the usual specifications of the exact wave function ϕ and the constraint Eqs. (1.1) which are incorporated into the starting expression (2.3) (or its analogs) for the VP, Sec. III.C discusses the permissible restrictions on the trial wave function ϕ_t, and Sec. III.D is concerned with methods of specifying the phase of the wave function, a subject which did not arise in Secs. I and II, where we dealt with purely real functions. To our knowledge, many of these and other matters considered in Secs. III.B–III.D were not fully discussed in the VP literature until rather recently (Gerjuoy, Rosenberg, and Spruch, 1975; Gerjuoy *et al.*, 1973, 1974; Gerjuoy, Rau *et al.*, 1975), but they are crucial for establishing the VP. Sections III.B–III.D also reexamine a few of the comparatively simple examples treated in Sec. I, after they have been slightly generalized (as, for example, by relaxing end-point restrictions on ϕ) so as to illustrate some of the points of present interest; the considerations of Sec. III.D on specifying the phase are illustrated in Secs. IV.A and IV.B.

A. Constructing the variational principle

When one is generalizing the formalism of Sec. II to problems involving complex quantities, appropriate modification of our previous notation is the most immediately evident need. In particular, with complex ϕ in constraints such as the normalization condition (1.3b),

$$\phi^\dagger\phi - 1 = 0 \,, \tag{3.1}$$

the dagger denotes the adjoint (complex conjugate transpose), and as before matrix notation is presumed; $\phi(\mathbf{r})$ is a column vector with components $\phi_j(\mathbf{r})$, while $\phi^\dagger(\mathbf{r})$ is a row vector whose components are $\phi_j^*(\mathbf{r})$. Correspondingly, the diagonal element of a Hermitian operator W now is denoted not by $F(\phi)$ but by

$$F(\phi^\dagger,\phi) = \phi^\dagger W\phi \,, \tag{3.2}$$

where, for particles with spin, Eq. (1.4) is replaced by

$$\phi^\dagger W\phi = \sum_{j,k} \int d\mathbf{r}\, \phi_k^*(\mathbf{r}) W_{kj} \phi_j(\mathbf{r}) \tag{3.3}$$

summed over all wave-function components (indexed by k in $\phi^\dagger$ and j in ϕ). In order not to complicate the notation further, we allow ϕ to represent other unknowns in addition to the conventional wave functions.

When a $B_i(\phi)$ is complex, its real and imaginary parts generally impose independent conditions on ϕ. Generally, therefore, to a real constraint of the form $B(\phi)=0$ for ϕ real, there will correspond for ϕ complex two independent real restrictions, which, as usual, are more conveniently expressed as

$$B(\phi^\dagger,\phi) = 0 \tag{3.4}$$

and

$$B^\dagger(\phi^\dagger,\phi) = 0 \,. \tag{3.5}$$

It is also more convenient to regard (each component of) ϕ and $\phi^\dagger$, and $\delta\phi$ and $\delta\phi^\dagger$, as independent, rather than the

real and imaginary parts of ϕ, and of $\delta\phi$.

When $F(\phi^\dagger,\phi)$ is real, as, for example, when F is given by Eq. (3.2) with W Hermitian, we anticipate that the exact Lagrange multipliers associated with the VP will obey some simple relationships. Thus, if there is a constraint $B_1(\phi^\dagger,\phi)=0$ with $B_1(\phi^\dagger,\phi)$ a number, there will be terms in the starting point of the VP of the form

$$\lambda_{at}B_1(\phi_t^\dagger,\phi_t)+\lambda_{bt}B_1^\dagger(\phi_t^\dagger,\phi_t)\ , \tag{3.6}$$

and we expect that

$$\lambda_a=\lambda_b^* \tag{3.7}$$

to ensure the reality of the contribution of the terms in (3.6) to $F_v(\phi_t^\dagger,\phi_t)$. Similarly, if there is a constraint $B_2(\phi^\dagger,\phi)=0$ with $B_2(\phi^\dagger,\phi)$ a function, the starting point of the VP will contain the terms

$$L_{at}^\dagger B_2(\phi_t^\dagger,\phi_t)+B_2^\dagger(\phi_t^\dagger,\phi_t)L_{bt}\ , \tag{3.8}$$

and we expect that

$$L_a(\mathbf{r})=L_b(\mathbf{r})\ . \tag{3.9}$$

However, there is no need to presume the relationships (3.7) and (3.9) at this juncture; if these relationships hold, we should be able to deduce them from the requirement that F_v be stationary. More generally, it will often be possible to anticipate such relations between "paired" left-occurring and right-occurring Lagrange multipliers (Appendix A) from the nature of the problem.

1. Making the starting expression stationary

In an analysis of δF_v, it is important to note that from the fundamental definitions of δB and $\delta B^\dagger$, we have

$$\begin{aligned}\delta(B^\dagger)&\equiv B^\dagger(\phi^\dagger+\delta\phi^\dagger,\phi+\delta\phi)-B^\dagger(\phi^\dagger,\phi)\\&=[B(\phi^\dagger+\delta\phi^\dagger,\phi+\delta\phi)-B(\phi^\dagger,\phi)]^\dagger\equiv(\delta B)^\dagger\ ,\end{aligned}$$

whereas, in general, one has

$$[B(\phi^\dagger,\phi)]^\dagger\equiv B^\dagger(\phi^\dagger,\phi)\neq B(\phi,\phi^\dagger)\ .$$

As a concrete example, consider $B(\phi^\dagger,\phi)=\phi^\dagger A\phi$, where $A^\dagger\neq A$.

If one is to obtain a VP, it must be possible, after one has expressed δF_v in terms of $\delta\phi$ and $\delta\phi^\dagger$, to recombine the terms so that δF_v takes the form

$$\delta F_v=K_a^\dagger\delta\phi+\delta\phi^\dagger K_b \tag{3.10a}$$

$$\equiv\sum_j\int d\mathbf{r}\,[K_{aj}^*(\mathbf{r})\delta\phi_j(\mathbf{r})+\delta\phi_j^*(\mathbf{r})K_{bj}(\mathbf{r})]\ , \tag{3.10b}$$

where the components $K_{aj}(\mathbf{r})$ of K_a and $K_{bj}(\mathbf{r})$ of K_b must be functions, and must not, for example, contain differential operations on the quantities they are adjacent to in Eqs. (3.10). Equations (3.10) are the present generalization of Eq. (2.5a). Equation (3.10b) implies, for every j, that the conditions

$$K_{aj}^*(\mathbf{r})=0,\quad K_{bj}(\mathbf{r})=0 \tag{3.11a}$$

will guarantee that δF_v in Eq. (3.10) vanishes for all allowed $\delta\phi_j(\mathbf{r})$ and $\delta\phi_j^*(\mathbf{r})$; Eqs. (3.11a) are the generalization of Eq. (2.5b). In a condensed notation, Eq. (3.11a) is

$$K_a^\dagger=0,\quad K_b=0\ . \tag{3.11b}$$

However, putting δF_v in the form Eq. (3.10) will require manipulations which are valid only when $\delta\phi,\delta\phi^\dagger$ and the Lagrange multipliers in the starting expression for F_v satisfy certain boundary conditions. Frequently these boundary conditions can be stated in pairs, because terms in $B_i^\dagger$ as well as in B_i now occur in the starting expression for F_v, and because terms in both $\delta\phi$ and $\delta\phi^\dagger$ now must be cast into the correct form. In typical cases, relations such as Eq. (2.6a) and those that arise in the present context of complex ϕ amount to the assertion that orders of integration can be inverted, or that boundary terms resulting from integration by parts can be ignored. In other words, they are typically boundary conditions of the usual type.

2. Starting expression without explicit functionals

At this point it may be useful to elaborate slightly on a remark made at the end of Secs. I.B.4 and I.B.5 regarding alternative treatments of ω_{nt}^2 and E_t—whether to regard them as direct guesses or as functionals of ϕ_t. In making a variational estimate of some entity Q, it often is unnecessary to write down explicitly the functional F such that

$$F(\phi^\dagger,\phi)=Q\ .$$

This remark is important because there are many problems for which the explicit form of the functional $F(\phi^\dagger,\phi)$ is unknown or extremely complicated. The use of $F(\phi^\dagger,\phi)=Q$ can then be avoided by replacing the trial estimate $F(\phi_t^\dagger,\phi_t)$ in the starting point for F_v by Q_t, where Q_t is a direct guess at Q, not necessarily involving ϕ_t; the starting point for F_v is then Q_t plus the usual Lagrange terms incorporating the various constraints. The derivation of the VP then proceeds in essentially the same way as when F_v contained the term $F(\phi_t^\dagger,\phi_t)$. These assertions are illustrated by the VP's in Secs. I.B.4 and I.B.5; other illustrations will be encountered below in Secs. III.B and V.B.1.

With these remarks, the procedure for deriving a VP for an arbitrary complex F with arbitrary complex constraints $B_i(\phi^\dagger,\phi)$ should be quite apparent, even though we have not given explicit equations valid for all F and B_i. Illustrative combinations of dummy and free variables (and/or indices) which can occur are described in Appendix A; this appendix also describes types of $\mathscr{L}_i$ which can occur, and summarizes the notation for Lagrange multipliers employed in this work.

B. Energy eigenvalues with nonreal wave functions

1. Basic variational principle for the energy

To illustrate the procedure presented in Sec. III.A, we return to the energy eigenvalues problem of Sec. I.B.5 where, however, the ϕ and/or H no longer need be purely real. The Hamiltonian H remains Hermitian, and the eigenfunction ϕ associated with the eigenvalue E continues to obey the Schrödinger equation (1.3a). Because one of our major objectives in this section is to contrast our procedures for complex quantities with our previous procedures for real quantities in Sec. I.B.5, we will continue to specify ϕ via a normalization condition [namely Eq. (3.1)], but for the moment will not attempt explicit specification of the wave-function phase. In Sec. III.C.1 we explain that it should not be necessary to specify explicitly the wave-function phase when constructing VP's for phase-independent quantities such as energy eigenvalues (more generally, any diagonal matrix element); it is of course clear that the phase was implicitly specified in Sec. I.B.5 by the assumption that $\phi(\mathbf{r})$ was purely real. The complications induced in the VP for the energy eigenvalue by insisting on explicitly specifying the phase are discussed in Gerjuoy, Rau *et al.* (1975).

With the foregoing considerations in mind, the analog of Eq. (2.3) in the present case obviously is

$$\langle E\rangle_{\rm var}=E_t+L_{at}^{\dagger}[(H-E_t)\phi_t]+[(H-E_t)\phi_t]^{\dagger}L_{bt}+\lambda_t(\phi_t^{\dagger}\phi_t-1) \tag{3.12}$$

because the normalization constraint equals its own adjoint, so that the $\lambda_{at}(\phi_t^{\dagger}\phi_t-1)$ and $\lambda_{bt}(\phi_t^{\dagger}\phi_t-1)^{\dagger}$ terms can be combined (that is, λ_t replaces $\lambda_{at}+\lambda_{bt}$). Equation (3.12) is to be compared with Eq. (1.28). Since the eigenvalue E is real—and since any reasonable variational estimate of E will be real—it seems obvious that setting

$$L_{at}=L_{bt} \tag{3.13}$$

in Eq. (3.12) right from the start still should permit derivation of a VP—that is, it seems obvious that right from the start we can assume that the exact adjoint functions satisfy $L_a=L_b$ consistent with Eq. (3.9). However, to verify that it is not necessary (though it is convenient) to impose the simplifying relation (3.9) at the very outset, we will proceed here as if L_a and L_b are unrelated and let the result that they are equal follow from the derivation. Similarly, we expect to infer that λ_t is real. We will, however, assume that δE is real.

Varying Eq. (3.12) we obtain

$$\delta E_v=\delta E+L_a^{\dagger}[(H-E)\delta\phi]+[(H-E)\delta\phi]^{\dagger}L_b-(\delta E)L_a^{\dagger}\phi-(\delta E)\phi^{\dagger}L_b+\lambda[\phi^{\dagger}\delta\phi+(\delta\phi^{\dagger})\phi]=0\ , \tag{3.14}$$

which is our present counterpart of Eq. (1.29). In Eq. (3.14), to obtain the conditions for the coefficients of $\delta\phi$ and $\delta\phi^{\dagger}$ to vanish, we must transfer the operation $(H-E)$ on $\delta\phi$ in two terms. Thus we will assume that it is legitimate to write

$$L_a^{\dagger}[(H-E)\delta\phi]=[(H-E)L_a]^{\dagger}\delta\phi\ , \tag{3.15a}$$

$$[(H-E)\delta\phi]^{\dagger}L_b=\delta\phi^{\dagger}[(H-E)L_b]\ . \tag{3.15b}$$

As in the case of Eq. (1.30), the validity of Eqs. (3.15) will have to be determined after the fact. We continue to proceed in very much the same way as in the case of the real wave function of Sec. II.B. When we use Eqs. (3.15), the conditions for the vanishing of δE_v given by Eq. (3.14) become

$$(H-E)L_b+\lambda\phi=0, \tag{3.16a}$$

$$[(H-E)L_a]^{\dagger}+\lambda\phi^{\dagger}=0\ , \tag{3.16b}$$

$$1-L_a^{\dagger}\phi-\phi^{\dagger}L_b=0\ , \tag{3.16c}$$

after equating to zero the coefficients of $\delta\phi^{\dagger}$, $\delta\phi$, and δE, respectively. The adjoint of Eq. (3.16b) is

$$[(H-E)L_a]+\lambda^{*}\phi=0\ . \tag{3.16d}$$

Multiplying Eq. (3.16a) on the left by $\phi^{\dagger}$, and assuming that the operation $(H-E)$ on L_b can be transferred to ϕ (again to be verified after the fact), we find, using Eq. (1.3a), that

$$\lambda=0\ . \tag{3.17a}$$

Therefore we see from Eqs. (3.16a) and (3.16d) that L_a and L_b are both eigenfunctions of H corresponding to the same eigenenergy E as ϕ itself. If E is nondegenerate, we therefore infer

$$L_a=c_a\phi,\quad L_b=c_b\phi\ , \tag{3.17b}$$

where c_a and c_b are constants. If E is degenerate, Eq. (3.17b) continues to be a valid solution of Eqs. (3.16) but is no longer the most general solution. For degenerate E, multiples of the other eigenfunctions belonging to E can be added to the right-hand side of Eq. (3.17b); however, these additional projections merely complicate the VP without altering it in any essential way (see below in Sec. III.B.2). Substituting Eqs. (3.17) in Eq. (3.16c) and using the normalization constraint (3.1) yield the added condition

$$c_a^{*}+c_b=1\ . \tag{3.17c}$$

The solutions (3.17) for the exact L_a, L_b, and λ suggest the trial estimates

$$\lambda_t=0,\quad L_{at}=c_{at}\phi_t,\quad L_{bt}=c_{bt}\phi_t\ , \tag{3.18a}$$

with

$$c_{at}^{*}+c_{bt}=1\ . \tag{3.18b}$$

Substituting Eqs. (3.18) into Eq. (3.12) yields

$$\langle E\rangle_{\rm var}=E_t(1-\phi_t^{\dagger}\phi_t)+\phi_t^{\dagger}H\phi_t\ . \tag{3.19}$$

Because c_{at} and c_{bt} do not appear explicitly in Eq. (3.19), obviously we could just as well have chosen originally $c_a=c_b=\frac{1}{2}$, which would make Eqs. (3.17) consistent with Eq. (3.9) and would correspond to making relations (3.18)

consistent with Eq. (3.13), all in accordance with our original *a priori* expectations. Similarly, Eqs. (3.17a) and (3.18a) do not disagree with our *a priori* expectation that λ_t should turn out to be real. As remarked following Eq. (1.35), the result in Eq. (3.17a) that λ is not merely real, but actually vanishes, is not typical; the exact Lagrange multiplier associated with the normalization constraint generally does not vanish and would not vanish, had we been considering here the diagonal matrix element of some operator other than the Hamiltonian.

So far we have not specified the normalization of ϕ_t. If ϕ_t differs from ϕ by first order, then to that order, using Eq. (3.1), we have

$$\phi_t^\dagger\phi_t-1=(\phi+\delta\phi)^\dagger(\phi+\delta\phi)-1$$
$$=\phi^\dagger\delta\phi+(\delta\phi^\dagger)\phi\ . \tag{3.20}$$

Using Eq. (3.20), we can easily verify that Eq. (3.19) is a variational estimate of E. When ϕ_t is strictly constrained by Eq. (3.1)—that is, when

$$\phi_t^\dagger\phi_t-1=0\ , \tag{3.21}$$

Eq. (3.20) implies

$$\phi^\dagger\delta\phi+(\delta\phi^\dagger)\phi=0\ , \tag{3.22}$$

and Eq. (3.19) gives the usual well-known form of the VP for E, namely,

$$\langle E\rangle_{\rm var}=\phi_t^\dagger H\phi_t\ . \tag{3.23}$$

2. Generalizations to degenerate E or unnormalized ϕ

The generalization of the above VP for energy eigenvalues of complex wave functions when there is degeneracy, or when the trial ϕ_t are unnormalized, goes through exactly as in Sec. I.B.6, where we considered real wave functions. There is little point in pursuing this further; in fact, the assertion that the extension to problems involving complex quantities often consists of no more than a reinterpretation of the dagger symbol is no exaggeration. In particular, it can be easily verified that

$$\langle E\rangle_{\rm var}=(\phi_t^\dagger H\phi_t)/(\phi_t^\dagger\phi_t) \tag{3.24}$$

is the correct generalization of Eq. (3.23) for unnormalized ϕ_t and that

$$\langle E\rangle_{\rm var}=\phi_t^\dagger H\phi_t+(\phi_t^\dagger H\phi_t)(1-\phi_t^\dagger\phi_t) \tag{3.25}$$

is valid for ϕ_t consistent with Eq. (3.20)—in other words for ϕ_t a first approximation (not normalized) to an exactly normalized ϕ.

C. Constraints on the exact ϕ versus its specifications

1. The different types of constraints which can arise

When constructing a VP for a given $F(\phi^\dagger,\phi)$ it is reasonable to expect the ϕ to be specified uniquely; otherwise, the problem may not be well defined. On the other hand, it is reasonable to expect that unique specification of ϕ may not be necessary when the value of $F(\phi^\dagger,\phi)$ is not affected by this lack of uniqueness in ϕ. Thus, because the eigenfrequencies ω_n on a stretched string do not depend on the normalization of the eigenfunctions $\phi_n(x)$, we are not surprised that in Sec. I.B.4 we were able to construct a VP for ω_n without specifying the normalization of $\phi_n(x)$. Similarly, we anticipate that specification of the phase of ϕ is unnecessary in the variational evaluation of diagonal matrix elements $\phi^\dagger W\phi$, but expect that at least the relative phase of ϕ_1 and ϕ_2 must be specified in the variational evaluation of off-diagonal matrix elements $\phi_1^\dagger W\phi_2$ (see Sec. IV.B).

The complete specification of ϕ is provided by a totality of what may be termed "partial specifications," as, for example, each of the constraints

$$B_i(\phi^\dagger,\phi)=0\ , \tag{3.26}$$

and the continuity conditions on ϕ. Let us suppose now that we are concerned with an $F(\phi^\dagger,\phi)$ which is not itself completely specified unless ϕ is completely specified; in other words, we are supposing that the VP being constructed for F is expected to require unique specification of ϕ. The question arises whether in this circumstance we expect to have to require that each of the partial specifications be incorporated into the starting expression for F_v via constraints of the form (3.26). The answer is clearly no, if only because one can have constraints which are not quantitative equations in ϕ of the form (3.26). As an example, it is often implicit that the functions ϕ are everywhere continuous and continuously differentiable. Such partial specifications helping to determine ϕ uniquely cannot be incorporated into the analogs of the starting expressions (2.3) or its analog for complex ϕ, although they usually must be kept in mind when the VP ultimately is employed for actual numerical estimation of F. Other partial specifications [for example, boundary conditions on surfaces or end points such as Eq. (1.22b)] are quantitative, but involve $\phi(\mathbf{r})$ at a set of points of measure zero compared to the measure of the entire range of $\mathbf{r}$ under consideration. Such partial specifications can be omitted from the starting expression without significantly modifying the possibility of obtaining a VP, because this possibility depends on one's being able to make first-order terms in δF_v vanish at essentially all $\mathbf{r}$ in its range, and because omission of constraints at a few specific values of $\mathbf{r}$ is usually not serious; in the event of such omission, these partial specifications—like the nonquantitative ones discussed above—usually will have to be kept in mind when actually employing the VP for computation. On the other hand, such partial specifications also can be included in the starting expression if we choose, as we shall illustrate in Sec. III.C.2 below.

2. Frequency eigenvalues with end-point constraints

In this section we illustrate the above remarks concerning the possibility of including in the starting expression

for the VP partial specifications which are not expressible in terms of quantitative expressions of the form (3.26). More particularly, we shall reexamine the frequency eigenvalues problem of Sec. I.B.4, starting now from

$$\langle\omega^2\rangle_{\rm var}=\omega_t^2+\int_0^s dx\, L_t(x)\left[\frac{d^2\phi_t}{dx^2}+\frac{\omega_t^2\rho}{T}\phi_t\right]$$
$$+\lambda_{0t}\phi_t(0)+\lambda_{st}\phi_t(s) \tag{3.27}$$

instead of Eq. (1.23); as before, we drop the subscript n. Equation (3.27) regards the boundary conditions (1.22b) as constraints of the form (1.1), to be incorporated into the starting point of the VP; the numerical Lagrange multipliers λ_0 and λ_s are associated with the constraints $\phi(0)=0$ and $\phi(s)=0$, respectively. Of course, we now no longer are requiring the trial ϕ_t to satisfy Eq. (1.22b), as we did in Sec. I.B.4; if the ϕ_t satisfy Eq. (1.22b), replacing Eqs. (1.23) by Eq. (3.27) obviously is pointless.

The requirement that Eq. (3.27) be a VP is

$$2\omega\delta\omega+\int_0^s dx\, L\left[\frac{d^2}{dx^2}\delta\phi+\frac{\omega^2\rho}{T}\delta\phi+\frac{2\omega\rho}{T}\phi\delta\omega\right]$$
$$+\lambda_0\delta\phi(0)+\lambda_s\delta\phi(s)=0\ . \tag{3.28}$$

(As always, we need not include terms which involve variations of the Lagrange multipliers—in this case L, λ_0, and λ_s—since the coefficients of the $\delta\mathscr{L}$—in this case δL, $\delta\lambda_0$, and $\delta\lambda_s$—always automatically vanish.) Collecting the coefficients of $\delta\omega$ again yields Eq. (1.24). Integrating by parts the term in Eq. (3.28) involving $d^2\delta\phi/dx^2$ again leads to Eq. (1.25a), together with the boundary conditions (1.25b) on L. However, because $\delta\phi$ no longer is being assumed equal to zero at the end points, this integration by parts leaves residual nonvanishing terms (proportional to $\delta\phi$ at the end points) which did not remain after integrating by parts the first variation of Eq. (1.23). These new terms combine with the last two terms in Eq. (3.28), also new, to yield

$$\lambda_0+L'(0)=0,\quad \lambda_s-L'(s)=0\ , \tag{3.29a}$$

where it now is convenient to use the prime to denote the derivative with respect to x. In Eq. (3.27), therefore, we can employ the trial estimates

$$\lambda_{0t}=-L_t'(0)=\phi_t'(0)\left[\int_0^s dx\,\rho\phi_t^2/T\right]^{-1},$$
$$\lambda_{st}=L_t'(s)=-\phi_t'(s)\left[\int_0^s dx\,\rho\phi_t^2/T\right]^{-1}. \tag{3.29b}$$

using Eq. (1.26), which remains valid.

Inserting Eqs. (1.26) (with the various exact quantities replaced by trial estimates) and Eqs. (3.29) into Eq. (3.27), we get a new version of the VP for ω^2, replacing Eq. (1.27), namely

$$\langle\omega^2\rangle_{\rm var}=\left[\int_0^s dx\,\rho\phi_t^2/T\right]^{-1}$$
$$\times\left[-\int_0^s dx(\phi_t\phi_t'')+\phi_t(0)\phi_t'(0)-\phi_t(s)\phi_t'(s)\right]. \tag{3.30}$$

The new result (3.30) differs from Eq. (1.27) by the terms in $\phi_t(0)$ and $\phi_t(s)$; these terms are of first order, since it is assumed $\phi_t(0)$ and $\phi_t(s)$ are chosen so that they differ by quantities of first order from $\phi(0)=\phi(s)=0$, whereas $\phi'(0)$ and $\phi'(s)$ are not equal to 0, implying that $\phi_t'(0)$ and $\phi_t'(s)$ are of zeroth order [if $\phi'(0)$ or $\phi'(s)$ were equal to 0 along with $\phi(0)=\phi(s)=0$, the solution $\phi(x)$ to Eq. (1.22a) would be identically zero]. However, although the algebra is a bit tedious, one can verify that Eq. (3.30) indeed is stationary for trial $\phi_t(x)$ differing by first order from the exact ϕ, but not necessarily satisfying Eq. (1.22b). Therefore the VP in Eq. (3.30)—derived starting from the form Eq. (3.27)—is a nontrivial generalization of the VP (1.27), and is valid under less severe restrictions on the trial functions than is Eq. (1.27), all in accordance with assertions made in Sec. III.C.1 above.

3. Restrictions on the trial functions ϕ_t

If the unknown exact ϕ is uniquely specified by the totality of partial specifications discussed in Sec. III.C.1—namely, the constraints (3.26) together with those requirements not being expressed in the form (3.26)—then (as remarked in Sec. III.C.1) the trial functions $\phi_t\neq\phi$ surely cannot satisfy all these partial specifications. Do the trial functions have to satisfy any of these partial specifications? To answer this question we note that for the purposes of insertion into a VP of the form (2.3) or its analogs, the only explicit requirement on ϕ_t is that it must satisfy all conditions [for example, the conditions implied by Eq. (2.6a)] needed to guarantee that δF_v will be stationary for all allowed $\delta\phi$; if such conditions coincide with any of the partial specifications uniquely determining ϕ, then of course ϕ_t must satisfy them. For instance, in Sec. I.B.5 we saw that $\phi_t(\mathbf{r})$, like $\phi(\mathbf{r})$ itself, had to be continuous and continuously differentiable. In Sec. III.B.2, however, the trial functions ϕ_t employed in the VP (3.19) did not have to be normalized, although the normalization condition (3.1) is required for complete specification of ϕ.

On the other hand, there is always the implicit requirement on ϕ_t that it differ from the exact ϕ by at most first order. As a result, even those partial specifications on ϕ which ϕ_t does not have to satisfy cannot be wholly ignored; in Sec. III.B.2, $\phi_t^\dagger\phi_t$ had to differ from unity by at most first order, as Eq. (3.20) showed. In practice, it often is not obvious whether those partial specifications on ϕ which ϕ_t does not have to satisfy are failing by at most first order; correspondingly, it often is not obvious whether the $\phi_t(\mathbf{r})$ selected is so close to $\phi(\mathbf{r})$ that $\phi_t(\mathbf{r})$ justifiably can be termed a "first-order" approximation to $\phi(\mathbf{r})$. In practice, therefore, the ϕ_t often are chosen to satisfy partial constraints on ϕ not needed to guarantee that δF_v will be stationary, on the assumption that confining attention to such additionally restricted classes of ϕ_t improves the likelihood of selecting a particular ϕ_t [for insertion into $F_v(\phi_t^\dagger,\phi_t,\mathscr{L}_t)$] which actually is a first-order approximation to ϕ. Such additional restrictions on ϕ_t also may be desirable to ensure that a sequence of successively improved ϕ_t, for example, via the introduction

of a successively increased number of arbitrary parameters, actually is converging to the desired ϕ; for example, ϕ_t should have the same symmetry and number of nodes as dictated by the known quantum numbers of the desired ϕ.

This brings us back to the role of the constraint terms in $\langle F\rangle_{\rm var}$. For simplicity, we will suppose for the moment that ϕ, F, and all constraints B_i are real and that the distinction between left- and right-occurring $\mathscr{L}_i$ can be ignored. Suppose first that there are two constraints and that they are incorporated into the starting expression as follows:

$$\langle F\rangle_{\rm var}=F_v(\phi_t,\mathscr{L}_{1t},\mathscr{L}_{2t}) =F(\phi_t)+\mathscr{L}_{1t}^{\dagger}B_1(\phi_t)+\mathscr{L}_{2t}^{\dagger}B_2(\phi_t)\ , \tag{3.31}$$

and that we make this $\langle F\rangle_{\rm var}$ stationary by canceling the $\delta\phi$ terms. To isolate $\delta\phi$, we bring the operators acting on it to the left, thereby obtaining the defining equation

$$\frac{\delta F}{\delta\phi}+\left[\left[\frac{\delta B_1}{\delta\phi}\right]^{\dagger}\mathscr{L}_1\right]^{\dagger}+\left[\left[\frac{\delta B_2}{\delta\phi}\right]^{\dagger}\mathscr{L}_2\right]^{\dagger}=0\ . \tag{3.32}$$

In some cases the presence of the term in $\mathscr{L}_2$ is absolutely necessary for the above equation to be consistent. A specific example is the case for which $F=\phi^{\dagger}W\phi$, $B_1=(H-E)\phi$, and $B_2=\phi^{\dagger}\phi-1$, so that Eq. (3.32) takes the form

$$2W\phi+(H-E)L_1+2\lambda_2\phi=0\ . \tag{3.33}$$

As is immediately observed on multiplying from the left by $\phi^{\dagger}$, this equation would be meaningless unless $\lambda_2=-\langle W\rangle$.

Now suppose we do not incorporate $B_2(\phi)=0$ as a constraint in the starting expression but merely write

$$\langle F\rangle_{\rm var}=F(\phi_t)+\mathscr{L}_{1t}^{\dagger}B_1(\phi_t) \tag{3.34}$$

and choose to work only with ϕ_t's of the class that satisfy $B_2(\phi_t)=0$. For such a class the two expressions (3.31) and (3.34) are obviously equivalent and we would expect Eq. (3.34) to be an acceptable stationary expression. But at first sight the equation for $\mathscr{L}_1$ that would follow from the starting point (3.34) would seem to be Eq. (3.32) without the term in $\mathscr{L}_2$, and this clearly would not define $\mathscr{L}_1$ properly. This seeming paradox is, however, resolved, because an equation such as Eq. (3.32) follows from the fact that in making $\langle F\rangle_{\rm var}$ stationary one sets the inner product of the left-hand side of Eq. (3.32) with $\delta\phi$ equal to zero, and, for arbitrary variations, Eq. (3.32) then follows. Now, however, if the ϕ_t's are not arbitrary but belong to the restricted class that satisfy $B_2(\phi_t)=0$, the variations are also no longer arbitrary and the only requirement is that the terms multiplying $\delta\phi$ be orthogonal to such $\delta\phi$'s. As a result, the Lagrange multiplier $\mathscr{L}_2$ makes its appearance at this stage, leading to the same Eq. (3.32) as before, even though the term in $\mathscr{L}_2$ is not introduced into the starting expression. Thus, when we are working with normalized real ϕ_t, the requirement on the term multiplying $\delta\phi$ is that it be proportional to ϕ [recall Eqs. (1.31) and (1.37)], so that we again have Eq. (3.33). Another example of this situation is considered in Sec. III.C.5. In practice, the convenience of being able to estimate F variationally from Eq. (3.34) rather than from the more complicated Eq. (3.31) is tempered by the inconvenience of having to find ϕ_t obeying $B_2(\phi_t)=0$.

4. Identities satisfied by ϕ_t

The preceding remarks lead naturally to some questions concerning the implications of finding identities satisfied by ϕ, that is, relations satisfied by ϕ which are not necessary for its unique specification. To be specific, suppose, as before, that $B_1(\phi)=0$, $B_2(\phi)=0$ (together with other partial specifications of the type discussed in Sec. III.C.1) suffice to specify ϕ uniquely, but that in addition there exists a known identity,

$$B_3(\phi)=0\ , \tag{3.35}$$

derivable from $B_1(\phi)=0, B_2(\phi)=0$. Then, what are the consequences of treating Eq. (3.35) as an independent constraint—that is, what are the consequences of attempting to derive a VP starting from

$$\langle F\rangle_{\rm var}=F(\phi_t)+\mathscr{L}_{1t}^{\dagger}B_1(\phi_t)+\mathscr{L}_{2t}^{\dagger}B_2(\phi_t) +\mathscr{L}_{3t}^{\dagger}B_3(\phi_t) \tag{3.36}$$

instead of the previous starting expression (3.31)? Correspondingly, what are the consequences of imposing the requirement $B_3(\phi_t)=0$ on the trial ϕ_t, either when estimating F from Eq. (3.31) or when employing the starting expression (3.36)? Of course, Eq. (3.36) will be the starting expression if Eq. (3.35) is being regarded as a truly independent third constraint, because the possibility of deriving Eq. (3.35) from other constraints has not been recognized.

The questions raised above are not difficult to answer. If Eq. (3.31) is a satisfactory starting point, then Eq. (3.36) obviously is at least equally satisfactory, because Eq. (3.36) reduces to Eq. (3.31) with the perfectly possible choice $\mathscr{L}_{3t}=\mathscr{L}_3=0$; in other words, if Eq. (3.32) inferred from (3.31) has solutions $\mathscr{L}_1,\mathscr{L}_2$, then the equation [corresponding to Eq. (3.32)] inferred from Eq. (3.36) surely has at least one set of solutions, namely $\mathscr{L}_3=0$ and $\mathscr{L}_1,\mathscr{L}_2$, as in Eq. (3.32). In essence, therefore, Eq. (3.36) merely offers a somewhat wider choice of possible variational expressions than does Eq. (3.31). Whether advantage can be taken of this wider choice—that is, whether enough information about the classes of exact $\mathscr{L}_1,\mathscr{L}_2,\mathscr{L}_3$ associated with Eq. (3.36) can be garnered to yield VP's (3.36) significantly different from Eq. (3.31)—is another matter, of course.

The considerations of the present section are readily extended to problems involving complex F, ϕ, B_i, etc., as will be seen; such considerations are relevant to the problem of diagonal and off-diagonal matrix elements discussed in Sec. IV.B. To some extent, the energy eigenvalue problems in Secs. I.B.5 and III.B.1 also illustrate the

considerations of the present section; in fact, for the purely real case of Sec. I.B.5, the Schrödinger equation (1.3a) and the normalization constraint (1.3b) play, in Eq. (1.28), precisely the roles played by B_1 and B_2, respectively, in Eq. (3.31). This energy eigenvalue illustration is flawed, however, by the fact that according to Eq. (3.17a) the exact value of λ [corresponding to $\mathscr{L}_2$ in Eq. (3.31)] turns out to be zero, so that Eq. (3.31) reduces automatically to Eq. (3.34); in other words, the energy eigenvalue problem is so simple that it is almost irrelevant whether or how one employs the normalization condition corresponding to B_2 in Eq. (3.31). In the literature, nontrivial illustrations of the considerations of this section mainly have involved the application of extra restrictions on the trial ϕ_t to quantum-mechanical calculations of atomic and molecular quantities. For instance, Rasiel and Whitman (1965) urge the use (in the conventional Rayleigh-Ritz VP for the ground-state energy) of an approximate wave function constrained to yield the known experimental value of, say, the molecular dipole moment. A related procedure, in so-called variation-perturbation calculations, has been proposed by Kirtman (1971), and applied to computations of atomic polarizability by Kirtman and Mowery (1971), and by Scott and Kirtman (1972). Note that these situations differ slightly from those discussed above and those to be noted immediately below, in that we are concerned not with some theoretical relations derivable from the basic defining relations, but with experimentally determined numbers. For quantum-mechanical bound-state problems, several authors (Epstein and Hirschfelder, 1961; Hirschfelder and Coulson, 1962; Hirschfelder *et al.*, 1964) have emphasized the possibility of "optimizing" the trial wave function by requiring it to satisfy various so-called hypervirial theorems. Correspondingly, Heaton and Moiseiwitsch (1971) and McWhirter and Moiseiwitsch (1972) have employed identities of the virial theorem type to optimize the trial ϕ_t in calculations of atomic scattering cross sections. Another example would be the variational calculation of oscillator strengths where constraints may be imposed so that alternative formulations ["length," "velocity," "acceleration"; see, for example, Bethe and Salpeter (1957), Sec. 59β] yield the same result.

5. An example of restrictions on ϕ_t: power dissipation

As a concrete example of the possibility of accounting for a constraint by introducing an appropriate term into the VP *or* by choosing the ϕ_t to satisfy the constraint (as discussed formally in Sec. III.C.3), we return to the treatment of power losses given in Sec. I.B.1; our discussion will be confined to the very simple case, involving only two currents, worked out explicitly in that section. There, we took I_{1t} and I_{2t} to be arbitrary, and introduced two Lagrange multipliers, λ_{1t} and λ_{2t}, to account for the constraints (1.6) and (1.7), respectively. Alternatively, we can choose I_{1t} to be arbitrary but choose

$$I_{2t}=I-I_{1t} .$$

Equation (1.6) is then automatically satisfied by the trial currents as well as the exact I_1, I_2 and we can now write

$$P_v \equiv \langle P(I_{1t},\lambda_t)\rangle = I_{1t}^2 R_1+(I-I_{1t})^2 R_2+\lambda_{2t}[I_{1t}R_1-(I-I_{1t})R_2] . \tag{3.37a}$$

Setting $\delta P_v=0$ leads to $\lambda_2=0$; the choice $\lambda_{2t}=0$ then gives

$$P_v = I_{1t}^2 R_1+(I-I_{1t})^2 R_2 . \tag{3.37b}$$

P_v of Eq. (3.37b) is a second VP, different from the P_v of Eq. (1.12). It will be useful to rederive the new result along lines closer to those used in Sec. III.C.3. Thus, having decided to choose the trial I_{1t} and I_{2t} so that they satisfy Eq. (1.6), we need not actually eliminate I_{2t}. Introducing $\hat{\lambda}_{2t}$ to distinguish it from the λ_{2t} that appears in Eq. (3.37a), we rewrite Eq. (3.37a) as

$$P_v \equiv \langle P(I_{1t},I_{2t},\hat{\lambda}_{2t})\rangle_{\text{var}} = I_{1t}^2 R_1+I_{2t}^2 R_2+\hat{\lambda}_{2t}(I_{1t}R_1-I_{2t}R_2) . \tag{3.37c}$$

The variation of this P_v is given by

$$\delta P_v=(2I_1R_1+\hat{\lambda}_2R_1)\delta I_1+(2I_2R_2-\hat{\lambda}_2R_2)\delta I_2 . \tag{3.38a}$$

Now however, the desired result $\delta P_v=0$ does not imply that the individual coefficients of δI_1 and δI_2 must vanish, since δI_1 and δI_2 are not independent. Rather, by virtue of the relationship $I_{2t}=I-I_{1t}$, they satisfy $\delta I_2=-\delta I_1$. The use of this relation in the above expression for δP_v gives

$$\delta P_v=[2(I_1R_1-I_2R_2)+\hat{\lambda}_2(R_1+R_2)]\delta I_1 . \tag{3.38b}$$

Since $I_1R_1=I_2R_2$, we must have $\hat{\lambda}_2=0$, and we are led back from Eq. (3.37c) to Eq. (3.37b). Alternatively, writing $\delta I_1+\delta I_2=0$, we can demand that δP_v in Eq. (3.38a) be a multiple of $\delta I_1+\delta I_2$. Taking the multiple to be $-\hat{\lambda}_1$, we arrive at

$$2I_1R_1+\hat{\lambda}_2R_1+\hat{\lambda}_1=0 ,$$
$$2I_2R_2-\hat{\lambda}_2R_2+\hat{\lambda}_1=0 .$$

These equations are identical in form to Eqs. (1.8) and (1.9); those, however, arose from the requirement that $\delta P_v=0$ with I_{1t} and I_{2t} each arbitrary. Thus we have a concrete example of the case discussed in Sec. III.C.3; the explicit satisfaction of a constraint eliminates one term, with a Lagrange multiplier, from the variational expression F_v, but the Lagrange multiplier reappears in the definition of other Lagrange multipliers in the act of making F_v stationary.

The two forms of the VP for P which we have derived, in this section and in Sec. I.B.1, are sufficiently different that one of them leads to a bound for the power dissipation P, while the other does not. It is instructive to pursue this question a bit further, and we do so in Appendix B, though it is a digression from the main focus of this

work. Appendix B also discusses the connection between our VP's for P and the extremum principle quoted at the beginning of Sec. I.B.1. To conclude, it may be instructive to examine a numerical illustration of the alternative VP's in Eqs. (1.12) and (3.37b). Consider a network with $R_1=3\Omega$, $R_2=5\Omega$, and $I=8A$. The exact solution has $I_1=5$, $I_2=3$, and $P=120$ W. Consider the choice $I_{1t}=4$, $I_{2t}=3.5A$. This choice does not satisfy the constraint that the current is conserved, and we have to use Eq. (1.12). Thereby we get our variational estimate $P_v=124$ W. Note that the fractional error is $\frac{1}{30}$, whereas the currents are in error by about $\frac{1}{5}$. Consider instead the choice $I_{1t}=4.5$, $I_{2t}=3.5A$, which does satisfy current conservation. We can now use Eq. (3.37b) and get $P_v=122$ W. This time the fractional error is $\frac{1}{60}$, whereas the error in the currents is approximately $\frac{1}{7}$.

D. Specifying the phase

As noted earlier, when constructing variational principles for certain quantum-mechanical quantities (as, for example, off-diagonal matrix elements $\phi_1^\dagger W\phi_2$), it is to be expected that some specification of the phase of the wave function must be imposed; otherwise, various ambiguities and contradictions can occur. Ways of specifying the phase will be discussed in the present section. The need for phase specification stems from the fact that if $\phi(\mathbf{r})$ satisfies the Schrödinger equation (1.3a), the normalization condition (3.1) and, for example, continuity and differentiability, then $e^{i\alpha}\phi(\mathbf{r})$ also satisfies continuity and differentiability as well as Eqs. (1.3a) and (3.1), where α is any real constant. (A comparable but much simpler question arises in problems studied earlier where ϕ is real—the overall sign of ϕ has to be specified.)

This presence of the arbitrary phase factor $e^{i\alpha}$ makes the need for unique phase specification particularly obvious when a VP for ϕ itself is sought—that is, when the desired functional F is ϕ. By definition, a variational estimate is in error by terms of second order and not of first order. Therefore, if a VP for $F\equiv\phi$ were possible without having to specify the phase, the corresponding variational estimate $\langle F\rangle_{\rm var}$ would have to differ from all members of the set $e^{i\alpha}\phi$ by terms of second order. But for small α the set $e^{i\alpha}\phi$ differs from ϕ by terms of first order, implying that a VP for a ϕ with unspecified phase is not possible; explicitly or implicitly, such a VP somehow must be singling out a particular phase. Similar remarks pertain to the off-diagonal matrix element $F\equiv\phi_1^\dagger W\phi_2$. If we specify ϕ_1 and ϕ_2 only to within phase factors, we can replace ϕ_1 by $e^{i\alpha_1}\phi_1$ and ϕ_2 by $e^{i\alpha_2}\phi_2$, so that F is changed by the factor $\exp i(\alpha_2-\alpha_1)$; it is not then possible to obtain a VP for F.

There are various ways of specifying the phase of ϕ; the choice is a matter of convenience, which unfortunately often reduces to a matter of least inconvenience. When the wave function $\phi(\mathbf{r})$ can be supposed purely real [as, for example, when $\phi(\mathbf{r})$ represents a spinless single-particle s state in a purely real spherically symmetric potential $V(r)$], then the phase is specified by the implicit or explicit decision that $\phi(\mathbf{r})$ indeed will be purely real, that is, by the decision to choose $\phi(\mathbf{r})$ such that

$$\phi(\mathbf{r})-\phi^*(\mathbf{r})=0\ . \tag{3.39}$$

As a matter of fact, the phase was determined by Eq. (3.39) in the energy eigenvalue problem of Sec. I.B.5, although seemingly the question of phase specification was wholly ignored in that section. Similarly, the phase was being determined in other problems examined earlier, as, for example, in Sec. I.B.4, although of course in classical nonquantal problems the decision that the eigenfunction has to be purely real can be based on the physical interpretation of the eigenfunction. We remark that because $e^{2\pi i}=1$, it usually is quite difficult as well as pointless to specify the phase of the function more uniquely than to modulo 2π. But even apart from modulo 2π, the phase of ϕ is not uniquely specified by Eq. (3.39); if ϕ satisfies Eq. (3.39), then so does $e^{i\pi}\phi$. As is further explained a little later in this section, this lack of uniqueness in the phase specified by Eq. (3.39) ordinarily does not cause ambiguities or contradictions in VP's; moreover the ambiguity factor $e^{i\pi}$ can be avoided by imposing in addition to Eq. (3.39) the requirement that $\phi(\mathbf{r}_0)>0$ at some given point $\mathbf{r}=\mathbf{r}_0$ where $\phi(\mathbf{r}_0)\neq 0$.

However, even with real potentials and spinless particles, Eq. (3.39) cannot be postulated for most single-particle wave functions—for example, for a spinless particle in a bound $l=1$ state with a definite magnetic quantum number, $+1$ or -1, or for a scattering state corresponding to the plane wave e^{ikz}. In such circumstances, Eq. (3.39) is unsatisfactory, but one can require that $\phi(\mathbf{r})$ be real at some given point $\mathbf{r}_0$. This requirement, however, fails to fix the phase when $\phi(\mathbf{r}_0)$ happens to vanish, so that some further thought concerning the choice of $\mathbf{r}_0$ is necessary.

The above discussion suggests that there may not be any one criterion that can be useful in fixing the phase for all cases. Many systems of interest, however, will be invariant under time reversal and rotation. Wigner (1932) showed that this important case permits a powerful single prescription for fixing the phase. Writing the wave function as a sum of products of radial functions and standard angular and spin functions, it is always permissible to take the radial functions as purely real. This convention, along with standard phase conventions for the angular and spin parts, serves to fix the phase of ϕ (see Sec. IV.A and Gerjuoy, Rau *et al.*, 1975, for more details).

On the other hand, when the system is not invariant under time reversal and rotation, the simplest and most useful procedure seems to be to fix the phase of ϕ relative to some arbitrary known function χ through a restriction of the sort that $\phi^\dagger\chi$ is either purely real or purely imaginary. Such a prescription is illustrated in Sec. IV.A for a VP for ϕ, along with a similar prescription for off-diagonal matrix elements $\phi_1^\dagger W\phi_2$.

IV. VARIATIONAL PRINCIPLES FOR WAVE FUNCTIONS AND MATRIX ELEMENTS

In this section, we derive VP's—by our direct approach—for bound-state and continuum wave functions and matrix elements of arbitrary operators between such wave functions. Since these derivations are treated fairly extensively in the literature (Gerjuoy *et al.*, 1972 and 1973, and Gerjuoy, Rau *et al.*, 1975), only a sketch of them will be given here. [Many of these results for matrix elements were derived earlier by other approaches (Schwartz, 1959a; Delves, 1963a; Dalgarno *et al.*, 1956).] In Sec. IV.A, we deal with VP's for wave functions and in Sec. IV.B with matrix elements, including a discussion of two alternative VP's. Formal questions concerned with a proper definition of the functions, particularly for systems which may not be invariant under time reversal and rotation, are also considered. Section IV.C sketches the resolution of a principal difficulty associated with these VP's, namely, the development of a systematic procedure for obtaining trial solutions to the Lagrange multipliers that are involved.

A. Variational principles for wave functions

1. Bound states

When ϕ is a real bound-state wave function defined by $(H-E)\phi=0$ and $\phi^\dagger\phi=1$, the construction of ϕ_v, the variational estimate, is straightforward. We write

$$\phi_v(\mathbf{r})=\phi_t(\mathbf{r})+\Lambda_{at}^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}')-E_t]\phi_t(\mathbf{r}')\} +L_{nt}(\mathbf{r})(\phi_t^\dagger\phi_t-1)\ . \tag{4.1}$$

In the second term on the right, the product involves an integration over $\mathbf{r}'$. The subscript n on L_{nt} helps to keep in mind that it arose from the normalization condition, whereas the subscript a on Λ_{at} is to facilitate comparison with a more general case that will follow. Their structures, that L_n is a function of one variable, while Λ_a is a function of two, is obvious on inspection and is evident from the general considerations of Sec. II.B and Appendix A. To derive the equations satisfied by these functions, as usual we have to set equal to zero the first-order terms in Eq. (4.1). Most of these terms are proportional to the real $\delta\phi\equiv\phi_t-\phi$ (we assume that ϕ_t, like ϕ, is real), but there also is a term $-(\Lambda_a^\dagger\phi)\delta E$. For the terms in $\delta\phi$, boundary conditions on Λ_a which will guarantee

$$\Lambda_a^\dagger\{(H-E)\delta\phi\}=\{(H-E)\Lambda_a\}^\dagger\delta\phi\ ,$$

are necessary; also, since we are dealing now with real functions, we have

$$\delta\phi^\dagger\phi=\phi^\dagger\delta\phi\ . \tag{4.2}$$

With these manipulations, we obtain, after writing

$$\delta\phi(\mathbf{r})=\int d\mathbf{r}'\delta\phi(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')$$

and putting the coefficient of $\delta\phi(\mathbf{r}')$ equal to zero,

$$[H(\mathbf{r}')-E]\Lambda_a(\mathbf{r},\mathbf{r}')+2\phi(\mathbf{r}')L_n^\dagger(\mathbf{r})=-\delta(\mathbf{r}-\mathbf{r}')\ , \tag{4.3a}$$

and, as follows on multiplying the above from the left by $\phi^\dagger(\mathbf{r}')$,

$$L_n(\mathbf{r})=-\tfrac{1}{2}\phi(\mathbf{r})\ . \tag{4.3b}$$

With Eq. (4.3b) in Eq. (4.3a), we have

$$(H-E)\Lambda_a=-Q\ , \tag{4.3c}$$

where $Q\equiv 1-P$, and $P=\phi\phi^\dagger$ is the projection operator for the function of interest. Λ_a is therefore a Green's function in the generalized sense (Courant and Hilbert, 1937, pp. 354–358) and is perfectly well-defined by (4.3c), even though E is an eigenvalue of H, because $\phi^\dagger Q=0$. However, we still have not taken care of the remaining term $-(\Lambda_a^\dagger\phi)\delta E$ in the variation of (4.1). If δE is of first order, this term requires that $\Lambda_a^\dagger\phi=0$, i.e., that the solution of (4.3c) have no projection on ϕ; this boundary condition on Λ_a can be satisfied, again because $\phi^\dagger Q=0$. If δE is of second order—as, for instance, would be the case if E_t were estimated from the variational principle in Sec. I.B.5—the boundary condition $\Lambda_a^\dagger\phi=0$ can be dropped without destroying the VP for ϕ. With these considerations, the VP for ϕ is complete.

When we consider a bound-state ϕ which is no longer real, $\delta\phi$ and $\delta\phi^\dagger$ would have to be considered linearly independent as discussed in Sec. III.A. A crucial difficulty then would be in the handling of the $\delta\phi^\dagger$ terms resulting from Eq. (4.1) during the analysis of first-order terms, since a relation such as Eq. (4.2) now would no longer be immediate. However, so long as the system is invariant under time reversal and rotation, the powerful theorem due to Wigner (1932) mentioned earlier permits such a relation according to the following scheme. The theorem states that $\phi(\mathbf{r})$ can be expanded in a specified orthonormal basis $\Omega_\alpha^{JM}(\hat{\mathbf{r}})$ of complex functions containing all the angular and spin dependence, with expansion coefficients $R_\alpha(r)$ which contain the radial coordinates and which can be chosen to be real. (Here, J and M specify the angular momentum and α the remaining quantum numbers of the system.) This particular choice or convention of fixing the phase of the complex ϕ, which permits one to write

$$\phi(\mathbf{r})=\Sigma_\alpha R_\alpha(r)\Omega_\alpha^{JM}(\hat{\mathbf{r}})\ , \tag{4.4a}$$

suggests that we pick trial functions of similar form, that is,

$$\phi_t(\mathbf{r})=\Sigma_\alpha R_{\alpha t}(r)\Omega_\alpha^{JM}(\hat{\mathbf{r}})\ , \tag{4.4b}$$

where the $R_{\alpha t}$ are real; it follows immediately that Eq. (4.2) remains valid, and Eqs. (4.3) again follow.

Proceeding now to the most general situation of a complex ϕ of a system that may not be invariant under time reversal and/or rotation, it is clear at the outset that to define ϕ properly we must fix its phase in some fashion. Various alternatives have been discussed in Sec. III.D and, as stated there, a convenient one is to fix the phase of ϕ relative to some arbitrary known function—χ, say—by

demanding that $\phi^\dagger\chi$ be either purely real or purely imaginary:

$$\phi^\dagger\chi \mp \chi^\dagger\phi = 0 \ . \tag{4.5}$$

To construct ϕ_v now requires a more general expression than Eq. (4.1), one which incorporates all the defining equations. We write

$$\phi_v(\mathbf{r}) = \phi_t(\mathbf{r}) + \Lambda_{at}^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}')-E_t]\phi_t(\mathbf{r}')\} + \{[H(\mathbf{r}')-E_t]\phi_t(\mathbf{r}')\}^\dagger\Lambda_{bt}(\mathbf{r},\mathbf{r}') + L_{nt}(\mathbf{r})(\phi_t^\dagger\phi_t - 1) + L_{pt}(\mathbf{r})(\phi_t^\dagger\chi \mp \chi^\dagger\phi_t) \ . \tag{4.6}$$

The subscript p on L_{pt} helps us to keep in mind that it arose from the phase condition. The term in Λ_{bt} is now necessary, because ϕ and $\phi^\dagger$ and, likewise, $\delta\phi$ and $\delta\phi^\dagger$, are completely independent and totally decoupled [no relation such as Eq. (4.2) now obtains] so that the Schrödinger equations for both ϕ and $\phi^\dagger$ have to be incorporated as constraints. So long as Eq. (4.2) was valid, inclusion of this term was unnecessary, as can be easily verified, because it would have led to $(H-E)\Lambda_b = 0$, making this term in Eq. (4.6) explicitly of second order and therefore unnecessary in a variational expression.

Examination of first-order terms in $\delta\phi$ and $\delta\phi^\dagger$ arising from the variation of Eq. (4.6) leads to the defining equations for the Lagrange multipliers [boundary conditions involving the projections of these Lagrange multipliers on ϕ can be avoided by estimating E_t variationally, as discussed above in connection with the VP (4.1)]. We omit the details and record the final result:

$$\Lambda_b(\mathbf{r},\mathbf{r}') = \tfrac{1}{2}\phi(\mathbf{r})L_b(\mathbf{r}') \ , \quad \Lambda_a = \Lambda_{a1} + \Lambda_{a2} \ , \qquad L_p = \tfrac{1}{2}\phi/(\phi^\dagger\chi) \ , \tag{4.7a}$$

where

$$(H-E)L_b = \phi - [\chi/(\phi^\dagger\chi)] \ , \tag{4.7b}$$

$$(H-E)\Lambda_{a1} = -1 + \phi\phi^\dagger \equiv -Q \ , \tag{4.7c}$$

$$(H-E)\Lambda_{a2} = -\tfrac{1}{2}\phi\phi^\dagger + \tfrac{1}{2}\chi\phi^\dagger/(\phi^\dagger\chi) \ . \tag{4.7d}$$

2. Continuum states

While there are technical differences, largely associated with boundary conditions, between the development of VP's in continuum and bound-state problems, there are no differences in principle; and, at least in retrospect, it is astonishing that it took such a long time to develop VP's for continuum problems once they were known for bound-state problems. Indeed, continuum problems can even be simpler, as we shall see, because there is usually no complication of an ambiguous phase (the asymptotic behavior which is specified serving to fix the phase), and because there is no normalization condition. For these reasons, $\phi^\dagger$ never enters into the equations serving to define ϕ, and therefore no term in Λ_b as in the previous paragraph is necessary. VP's for continuum ϕ can be derived both in a partial wave decomposition and for the full $\phi(\mathbf{r})$; for simplicity, we restrict ourselves to potential scattering. We briefly record here the latter VP. For ϕ defined by

$$(H-E)\phi = 0 \ , \tag{4.8a}$$

$$\phi(\mathbf{r}) \sim \exp(i\mathbf{k}\cdot\mathbf{r}) + f(\theta)\exp(ikr)/r \ , \tag{4.8b}$$

where $f(\theta)$ is the scattering amplitude, the VP takes the form

$$\phi_v = \phi_t + \Lambda_t^\dagger[(H-E)\phi_t] \ , \tag{4.9}$$

where the asymptotic form of ϕ_t is given by Eq. (4.8b) with $f(\theta)$ replaced by $f_t(\theta)$. The analysis of $\delta\phi$ terms in Eq. (4.9) again involves a switching over of the $(H-E)$ operator onto Λ. The surface terms that are now generated no longer vanish almost automatically, as they do in the bound-state case; for the surface terms to vanish, Λ^* must be purely outgoing [$\delta\phi$ is purely outgoing, so that if Λ^* is also, the surface term involving $(\delta\phi\ \mathrm{grad}\Lambda^* - \Lambda^*\ \mathrm{grad}\delta\phi)$ will vanish]. Thus Λ is defined by

$$(H-E)\Lambda = -1 \ , \tag{4.10}$$

with $\Lambda^*(\mathbf{r},\mathbf{r}')$ outgoing.

A more complete discussion of VP's for scattering problems is contained in Sec. V.

B. Variational principles for matrix elements

Once a variationally accurate ϕ_v has been derived by the methods of Sec. IV.A, it is, of course, obvious that any matrix element evaluated with it, such as $\phi_v^\dagger W\phi_v$, will automatically be a variational estimate of the exact $\phi^\dagger W\phi$. However, an alternative path that we could follow would be to build directly by our general procedure a VP for the desired matrix element without first getting a variationally good ϕ_v; often the direct approach will be simpler, since obtaining a ϕ_v can be tedious.

The VP derived by the direct approach coincides with the results of Schwartz (1959a, 1959b), Delves (1963a), and Dalgarno *et al.* (1956). That the two approaches can only lead to expressions which can at most differ in second order is again clear from the start, and an actual demonstration of this is given in Gerjuoy, Rau *et al.* (1975). We give below a direct construction of a VP for $\phi^\dagger W\phi$, with W a Hermitian operator and ϕ a bound-state wave function. We write

$$(\phi^\dagger W\phi)_v = \phi_t^\dagger W\phi_t + L_t^\dagger[(H-E_t)\phi_t] + [(H-E_t)\phi_t]^\dagger L_t + \lambda_t(\phi_t^\dagger\phi_t - 1) \ . \tag{4.11}$$

In Eq. (4.11), we have recognized that $\phi^\dagger W\phi$ will be a real quantity, and therefore the left-occurring and right-

occurring Lagrange multipliers in the second and third terms were taken to be merely adjoints of one another; correspondingly, we assume λ_t is real (as well as the exact λ). We also expect ϕ to be real, but have not used this fact. We have also anticipated that even for a general ϕ the specification of the phase should be irrelevant in dealing with a diagonal matrix element, and we have therefore not incorporated the phase specification as a constraint. It would be easy to verify that inclusion of such a term would be redundant, since the term would be explicitly of second order.

Examining $\delta\phi$ and $\delta\phi^\dagger$ terms in Eq. (4.11), we find that each gives

$$(H-E)L=-W\phi-\lambda\phi\ . \tag{4.12}$$

Premultiplication by $\phi^\dagger$ gives $\lambda=-\phi^\dagger W\phi$. L is arbitrary to within a multiple of ϕ; L is made unique by arbitrarily assigning a value to $\phi^\dagger L$. The VP is now complete. VP's for off-diagonal matrix elements $\phi_1^\dagger W\phi_2$ can be derived in exactly the same manner. The complication is merely one of bookkeeping, with several Lagrange multipliers associated with the Schrödinger equations for ϕ_1, $\phi_1^\dagger$, ϕ_2, and $\phi_2^\dagger$, the two normalization conditions and the two conditions specifying the phases of ϕ_1 and ϕ_2. Such a VP (though with a slightly different choice for specifying the phase) is given in Gerjuoy *et al.* (1972). Similarly, VP's for matrix elements involving one or more continuum wave functions can be constructed. The analysis of the surface terms will set ingoing or outgoing boundary conditions on the L's involved. We do not record here explicitly any of these VP's. Sections V and VI treat in detail VP's in scattering theory where the functions involved are continuum wave functions and the operator W coincides with the Hamiltonian.

C. Procedure for choosing trial Lagrange multipliers

The entire emphasis of the present paper lies in the formal construction of VP's, with questions of utility for the most part studiously avoided. However, as noted in the Introduction, some progress of considerable potential value has recently been made with regard to mechanisms for choosing the various L_t's and Λ_t's that must be obtained if the VP's are to be given meaning, and we would like at least to touch upon this development.

The most important question for the considerations of this section is the practical one of using the result, which reduces to the question of choosing trial approximations to the L's and Λ's, since the equations obeyed by them, such as Eqs. (4.3a), (4.7), and (4.12), are complicated and do not admit simple solutions in closed form. This question of obtaining trial Lagrange multipliers is central to this entire review paper since, otherwise, the results are of no practical significance. We have seen in Sec. I and it will turn out again in Secs. V and VI and in many examples of nonlinear equations [see, for example, Rau (1976)] that the equations for L quite often admit immediate solutions, sometimes in terms of the exact ϕ, so that the choice for L_t is quite clear. However, this is not the case for the problems considered in this section and, in fact, in our entire experience the problem of getting trial Lagrange multipliers seems to be most involved in the case of VP's for wave functions and matrix elements. A procedure for resolving this problem has recently been developed (Gerjuoy *et al.*, 1974; Gerjuoy, Rosenberg, and Spruch, 1975); we sketch below a summary of these results.

Equations (4.3c), (4.7), and (4.12) for the exact L's and Λ's, or more generally the $\mathscr{L}$'s, have the structure

$$(H-E)\mathscr{L}=q(\phi)\ ,\quad \phi^\dagger q(\phi)=0\ , \tag{4.13a}$$

where $q(\phi)$ is a formally known functional; for example, in the case of VP's for diagonal matrix elements,

$$q(\phi)=-W\phi+(\phi^\dagger W\phi)\phi\ . \tag{4.13b}$$

The exact $\mathscr{L}$'s are therefore obtained through the inversion of the operator $(H-E)$. Since ϕ is the homogeneous solution of this operator, the solution of the inhomogeneous equation (4.13a) for $\mathscr{L}$ involves an arbitrary additive amount proportional to ϕ. To make $\mathscr{L}$ unique we can therefore supplement Eq. (4.13a) by a condition such as

$$\mathscr{L}^\dagger\phi=0\ , \tag{4.14}$$

thereby explicitly removing from $\mathscr{L}$ any amount proportional to ϕ. There is thus no ambiguity or singularity difficulty in inverting $(H-E)$. The problem arises, however, because neither the exact ϕ nor E is known, nor do we know the exact $\mathscr{L}$ whose approximation we seek for use in the VP (4.11). What we need is a trial $\mathscr{L}$, and the question is how to obtain it. The naive assumption that $\mathscr{L}_t$ can be obtained by replacing ϕ, E, and $\mathscr{L}$ in Eqs. (4.13a) and (4.13b) by the trial quantities proves to be incorrect and leads to a loss of the VP. Thus, for example, in (4.11) the boundary conditions on L (and on ϕ) imply that the second and the third terms on the right can be rewritten as $I^\dagger$ and I, respectively, where $I\equiv\phi_t^\dagger[(H-E_t)L_t]$. With the replacements just indicated, we have $(H-E_t)L_t=q(\phi_t)$, so that $I=\phi_t^\dagger q(\phi_t)=0$, according to the second of (4.13a), with ϕ_t replacing ϕ. As a result, for ϕ_t normalized, our VP (4.11) would collapse to the ridiculous result $(\phi^\dagger W\phi)_v=\phi_t^\dagger W\phi_t$! The loss of the VP was recognized, though not investigated further along the lines of present interest, by Krieger and Sahni (1972). The source of the difficulty is that the equation so obtained from Eqs. (4.13) with the replacement of exact quantities by trial leads to a unique solution for $\mathscr{L}_t$, since there is, in general, no homogeneous solution of $(H-E_t)$; in particular, $(H-E_t)\phi_t\neq 0$. Demanding an analog of Eq. (4.14), namely, that $\mathscr{L}_t$ be orthogonal to ϕ_t, is of no help and $\mathscr{L}_t$ develops a near-singularity as a consequence of inverting $(H-E_t)$. The problem is serious and, in fact, gets *worse* as ϕ_t is improved, since the near-singularity involves $(E-E_t)^{-1}$. This being an inverse second-order quantity, basic statements, such as the demand that $\mathscr{L}-\mathscr{L}_t$ be in some sense of first order of smallness become meaningless, and, in consequence, the VP is lost.

The above discussion suggests the direction in which one might seek the resolution of the difficulty. In modifying Eqs. (4.13) by replacing ϕ, E, and $\mathscr{L}$ by the corresponding trial quantities, we must also modify the Hamiltonian such that ϕ_t is a homogeneous solution of the operator to be inverted. In that case, near-singularities in $\mathscr{L}_t$ can be eliminated by enforcing the requirement

$$\mathscr{L}_t^\dagger \phi_t = 0 \ . \tag{4.15}$$

The procedure followed in the earlier literature (Dalgarno and Stewart, 1960; Dalgarno *et al.*, 1956; Krieger and Sahni, 1972) for modifying H is the very limited one of choosing ϕ_t simple enough that one can find an H_t which obeys $H_t\phi_t = E_t\phi_t$. This usually entails an uncorrelated Hartree-type product wave function for a many-electron atom, in which case H_t is a sum of independent hydrogenic Hamiltonians for each electron. This stringent limitation to extremely simple ϕ_t restricts the usefulness of the VP and perhaps explains why VP's for matrix elements have not been used regularly. An alternative approach (Schwartz, 1959b) does allow a relatively unrestricted ϕ_t, but the number of situations to which it can be applied is very limited. What is really needed to make the procedure competitive with, or superior to, other machine calculations in modern atomic physics is a procedure that can work with more sophisticated ϕ_t. Given any arbitrary ϕ_t which may have arisen from some self-consistent Rayleigh-Ritz calculation, the question is how the Hamiltonian H is to be modified; needless to say, it will no longer be possible to find an H_t of which ϕ_t is an eigenfunction by any straightforward procedure of inspection. The above question of obtaining a trial modified-Hamiltonian $H_{\text{mod},t}$ for use in a sequence of trial equations

$$(H_{\text{mod},t} - E_t)\mathscr{L}_t = q_t(\phi_t) \ , \tag{4.16a}$$

$$\phi_t^\dagger q_t(\phi_t) = 0 \ , \tag{4.16b}$$

now has been answered and various choices for $H_{\text{mod},t}$ given; typically, although not necessarily, the most useful choice for $q_t(\phi_t)$ will be $q(\phi_t)$. One suitable choice for $H_{\text{mod},t}$ in (4.16a) is

$$H^{(1)}_{\text{mod},t} = H - \frac{HP_tH}{E_t} \ , \quad P_t \equiv \phi_t\phi_t^\dagger \ , \quad E_t = \phi_t^\dagger H \phi_t \ . \tag{4.17}$$

Earlier, this operator had been considered in a somewhat different but not unrelated context (see Monaghan and Rosenberg, 1972) and had been shown to have a number of very interesting properties. Thus note that $\phi_t^\dagger H^{(1)}_{\text{mod},t} = 0$. Note also that $H^{(1)}_{\text{mod},t} \to H - EP$ as $\phi_t \to \phi$. Projecting Eq. (4.16a) onto $\phi_t^\dagger$ and using (4.16b), we see that (4.16) also is consistent with the requirement of Eq. (4.15) that $\mathscr{L}_t$ be orthogonal to ϕ_t. Because $\phi_t^\dagger \mathscr{L}_t = 0$, the quantity $(H^{(1)}_{\text{mod},t} - E_t)\mathscr{L}_t$ in (4.16a) approaches $(H - E_t)\mathscr{L}_t$ as $\phi_t \to \phi$. Therefore, since the equations and boundary conditions that define $\mathscr{L}_t$ approach those that define $\mathscr{L}$ as $\phi_t \to \phi$, it follows that $\mathscr{L}_t \to \mathscr{L}$ as $\phi_t \to \phi$, as desired. At the same time, no singularity differences arise in handling Eqs. (4.16), and we consequently have a procedure for obtaining $\mathscr{L}_t$ starting from any arbitrary given ϕ_t. In fact, the above prescription has further merit, because it turns out that, for sufficiently accurate ϕ_t [see Gerjuoy *et al.* (1974) for details], $H^{(1)}_{\text{mod},t}$ has its lowest eigenvalue above E_t. Thus, returning now to the VP (4.11), the functional

$$M(L_{tt}) = L_{tt}^\dagger (H^{(1)}_{\text{mod},t} - E_t) L_{tt} - L_{tt}^\dagger q(\phi_t) - q^\dagger(\phi_t) L_{tt} \tag{4.18}$$

will be a minimum[5] at $L_{tt} = L_t$. A very practical procedure is then available for obtaining a good L_t without inverting any operators, because a form of L_{tt} with open variational parameters can be inserted into Eq. (4.18) and the functional minimized with respect to these parameters, thereby yielding the best possible trial Lagrange multiplier of that form to go with the given ϕ_t into the variational expression [similar remarks pertain to the Λ_t appearing in the VP (4.6)]. Such a subsidiary extremum principle for obtaining the trial Lagrange multiplier makes the results of Secs. IV.A and IV.B of practical significance. One must of course evaluate matrix elements of H to obtain $M(L_{tt})$ as defined by Eq. (4.18), but it is important to note that because of the presence of the projection operator P_t in Eq. (4.17) one need *not* evaluate the very much more difficult matrix elements of H^2, even though $H^{(1)}_{\text{mod},t}$ contains H twice. Preliminary applications have been made by Shakeshaft *et al.* (1976) to diagonal matrix elements (the diamagnetic susceptibility and form factor of helium) and by Wadehra *et al.* (1978) to an off-diagonal matrix element, an electric dipole transition matrix element in helium.

V. THE KOHN VARIATIONAL PRINCIPLE

In this section we will derive or sketch the derivation of the well-known Kohn VP (Kohn, 1948) in various of its forms by our routine procedure. The Kohn principle, along with variants thereon (see Sec. V.A) has been very widely used to estimate various sorts of quantum-mechanical scattering parameters, as, for example, phase shifts, elastic and inelastic scattering amplitudes, elements of the reaction operator, etc. Initially (Secs. V.A and V.B) we confine our attention to the simplest type of quantum-mechanical collision, namely, the potential scattering of spinless particles. More complicated collisions are considered in Sec. V.C; various generalizations

[5]The essential point is that if $AX = f$, where A is Hermitian and positive definite and f is known, then if we introduce

$$J(X_t) \equiv J(X + \delta X) \equiv -X_t^\dagger f - f^\dagger X_t + X_t^\dagger A X_t \ ,$$

we see that $J(X) = -f^\dagger X$. However, we also see that $J(X_t) = -f^\dagger X + \delta X^\dagger A \delta X$, which is $\geq -f^\dagger X = J(X)$.

of the Kohn principle are examined in Sec. V.D. Because of the close formal analogy between quantum-mechanical potential scattering and the scattering of classical acoustic or electromagnetic waves propagating in inhomogeneous media [see, for instance, Kalikstein and Spruch (1964) for casting Maxwell's equations in Schrödinger form], many of the results in Secs. V.A and V.B can be adapted essentially without change to the estimation of acoustic or electromagnetic scattering parameters (see also Kalikstein, 1981). The VP derived in Sec. V.E is of interest in reactor physics (Stacey, 1972).

A. Phase shifts in potential scattering

1. Tangent of the phase shift

In the scattering of spinless particles by a well-behaved spherically symmetric short-range real potential $V(r)$, the solutions ϕ_l associated with the lth partial wave obey

$$(H_l-E)\phi_l(r)=0\ , \tag{5.1a}$$

where $E=\hbar^2k^2/(2m)$ is real, where $\phi_l(r)$ is the lth partial wave (multiplied by r), where the volume element is dr so that $a^\dagger b$ represents $\int a^*(r)b(r)dr$, and where

$$H_l(r)\equiv\frac{\hbar^2}{2m}\left[-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}\right]+V(r)$$

$$\equiv T_l+V(r)\ . \tag{5.1b}$$

A unique solution to Eq. (5.1a) is specified by the requirements that (i) $\phi_l(r)$ be everywhere continuous and continuously differentiable; (ii) $\phi_l(r)$ be regular ($\sim r^{l+1}$) at the origin; and (iii) $\phi_l(r)$ have the asymptotic form as $r\sim\infty$

$$\phi_l(r)\sim\sin(kr-\tfrac{1}{2}l\pi)+\tan\eta_l\cos(kr-\tfrac{1}{2}l\pi)$$

$$=\sec\eta_l\sin(kr-\tfrac{1}{2}l\pi+\eta_l)\equiv\phi_{l\infty}(r)\ , \tag{5.1c}$$

where η_l is the lth phase shift.

The quantity $\phi_l(r)$ is real, so that an obvious starting point for a variational estimate of $\tan\eta_l$ is

$$F_v=\langle\tan\eta_l\rangle_{\rm var}$$

$$=\tan\eta_{lt}+L_{lt}^\dagger[(H_l-E)\phi_{lt}]\ , \tag{5.2}$$

where the real trial solution $\phi_{lt}(r)$ is continuous and continuously differentiable, is regular at the origin, and has the asymptotic form (5.1c), but with η_{lt} in place of η_l. We need not here distinguish between left-occurring and right-occurring adjoint functions. Of course, the last term in Eq. (5.2) is a single integral over r, from 0 to ∞. No trial E_t is required in Eq. (5.2), since we are dealing with a continuum solution, for which the energy E can be regarded as an assigned parameter. The important and nontrivial complications involved in assigning E when the colliding systems are composite, arising from uncertainties concerning bound-state eigenfunctions and eigenenergies, are examined in Sec. V.C.

Varying Eq. (5.2) gives

$$\delta F_v=\delta(\tan\eta_l)+L_l^\dagger[(H_l-E)\delta\phi_l]=0\ . \tag{5.3}$$

When we integrate by parts, Eq. (5.3) yields

$$\delta(\tan\eta_l)+[(H_l-E)L_l]^\dagger\delta\phi_l-\frac{\hbar^2}{2m}\left[L_l(r)\frac{d}{dr}\delta\phi_l(r)-\left[\frac{d}{dr}L_l(r)\right]\delta\phi_l(r)\right]\Bigg|_0^\infty=0\ . \tag{5.4}$$

To ensure that Eq. (5.4) holds for essentially arbitrary $\delta\phi_l(r)$, we must require

$$(H_l-E)L_l(r)=0\ . \tag{5.5a}$$

In addition, the end-point contribution to Eq. (5.4) [the bracketed terms on the left-hand side of Eq. (5.4)] must vanish at $r=0$, because there are no other $r=0$ contributions to Eq. (5.4). Thus, since $\phi_{lt}(r)$ has been so chosen that $\phi_{lt}(0)=0$ and therefore $\delta\phi_l(0)=0$, we infer

$$L_l(0)=0\ . \tag{5.5b}$$

Equations (5.5) imply that L_l is a multiple of ϕ_l, that is,

$$L_l(r)=B_l\phi_l(r)\ , \tag{5.5c}$$

where the factor B_l may depend on η_l but is independent of r. To satisfy Eq. (5.4), we therefore also must require that

$$\delta(\tan\eta_l)-\frac{\hbar^2}{2m}B_l\left[\phi_l(r)\frac{d}{dr}\delta\phi_l(r)-\delta\phi_l(r)\frac{d}{dr}\phi_l(r)\right]\Bigg|^\infty=0\ .$$

Since, for $r\sim\infty$, we have from Eq. (5.1c),

$$\delta\phi_l(r)\sim\delta(\tan\eta_l)\cos(kr-\tfrac{1}{2}l\pi)\ ,$$

we deduce that $(k\hbar^2/2m)B_l=-1$; we therefore have

$$L_l(r)=-(2m/\hbar^2k)\phi_l(r)\ . \tag{5.6a}$$

Equation (5.6a) suggests the choice

$$L_{lt}(r)=-(2m/\hbar^2k)\phi_{lt}(r)\ . \tag{5.6b}$$

The use of Eq. (5.6b) in Eq. (5.2) yields

$$\langle\tan\eta_l\rangle_{\rm var}=\tan\eta_{lt}-(2m/\hbar^2k)\phi_{lt}^\dagger[(H_l-E)\phi_{lt}]\ , \tag{5.7}$$

the Kohn VP.

Note that we could have made the choice

$$L_{lt}(r)=-(2m/\hbar^2 k)\phi_{ltt}(r)\ ,$$

rather than Eq. (5.6b), with ϕ_{ltt} an approximation to ϕ_l different from ϕ_{lt}. Equation (5.7) would then contain two trial functions. A possible advantage might accrue by choosing ϕ_{lt} to be rather simpler in form than ϕ_{ltt}, with the evaluation of $(H_l-E)\phi_{lt}$ thereby being simplified for a calculation involving a given total effort. The possibility of using two trial functions has been noted by many others and arises in many contexts, including, for example, the Rayleigh-Ritz principle. (The variational bound property would, however, be lost for the Rayleigh-Ritz case.)

In the foregoing derivation of Eq. (5.7), the dependence of $\tan\eta_l$ on $\phi_l(r)$ was implicit, in terms of the asymptotic form of ϕ_l as $r\sim\infty$. However, it is also known and is simple to prove [see, for instance, Schiff (1968), Sec. 37] that

$$\tan\eta_l=-(2m/\hbar^2)[rj_l(kr)V(r)]^\dagger\phi_l(r)\ , \tag{5.8a}$$

which explicitly expresses $\tan\eta_l$ as a function of—or, more precisely, an integral over—$\phi_l(r)$. This expression can serve as an alternative starting point for the derivation of a VP for $\tan\eta_l$. The derivation is straightforward and again leads to Eq. (5.7).

If the interaction contains a Coulomb potential in addition to the short-range $V(r)$, we need merely replace Eq. (5.1c) by an expression containing the asymptotic forms of appropriately normalized regular and irregular Coulomb functions. With η_l now representing the phase shift relative to the Coulomb phase, the derivation of the Kohn principle proceeds along almost identical lines.

Returning, for simplicity only, to non-Coulombic potentials, a particular case of considerable interest is that for which $E=0$. (We are then at the bottom of the continuous spectrum, so that, not surprisingly, one can rather easily obtain a variational bound rather than simply a VP. We will not consider the variational bound here but will comment briefly on it later.) From the Wigner theorem (Wigner, 1948) for short-range potentials, we know that η_l behaves, modulo π, as k^{2l+1} as $k\sim 0$. Restricting ourselves, again for simplicity only, to the $l=0$ case, we can therefore write

$$\tan\eta_0\sim -kA\ ,\quad k\sim 0\ , \tag{5.8b}$$

where the scattering length A determines the cross section at zero incident energy. We then can introduce a new wave function $\bar{\phi}_0$, defined by

$$\bar{\phi}_0(r)=\lim_{k\to 0}\frac{\phi_0(r)}{k}\ ,$$

where, by Eq. (5.1c)

$$\bar{\phi}_0(r)\sim r-A \tag{5.8c}$$

as $k\to 0$ at fixed large r. Since any solution of the zero-energy Schrödinger equation is asymptotically a linear combination of r and a constant, the above definition of $\bar{\phi}_0(r)$ simply serves to fix the normalization of our zero-energy solution. In terms of $\bar{\phi}_0(r)$, Eq. (5.7) reduces to

$$\langle A\rangle_{\rm var}=A_t+(2m/\hbar^2)\bar{\phi}_{0t}^\dagger H_0\bar{\phi}_{0t}\ . \tag{5.8d}$$

2. The phase shift itself

Had we desired a variational estimate of η_l rather than of $\tan\eta_l$, we could have started from

$$F_v=\langle\eta_l\rangle_{\rm var}=\eta_{lt}+\hat{L}_{lt}^\dagger[(H_l-E)\phi_{lt}] \tag{5.9}$$

instead of Eq. (5.2). In Eq. (5.9), η_{lt} is, as in the first derivation of the Kohn principle, defined by the asymptotic form of ϕ_{lt}, and the caret is used to distinguish the present adjoint function $\hat{L}_{lt}$ from the adjoint function L_{lt} in Eq. (5.2). Proceeding as before, we find that

$$\begin{aligned}\hat{L}_l(r)&=L_l(r)/\sec^2\eta_l\\&=-(2m/\hbar^2 k\sec^2\eta_l)\phi_l(r)\ .\end{aligned}$$

We thereby find

$$\langle\eta_l\rangle_{\rm var}=\eta_{lt}-(2m/\hbar^2 k\sec^2\eta_{lt})\phi_{lt}^\dagger[(H_l-E)\phi_{lt}]\ . \tag{5.10}$$

Equation (5.10) is a perfectly satisfactory VP for η_l; to make direct contact with the form derived by Hulthén (1944, 1948), we need merely absorb the $\sec^2\eta_{lt}$ factor by introducing

$$\begin{aligned}\hat{\phi}_l(r)&=\phi_l(r)/\sec\eta_l\ ,\\ \hat{\phi}_{lt}(r)&=\phi_{lt}(r)/\sec\eta_{lt}\ ,\end{aligned} \tag{5.11}$$

where $\hat{\phi}_l(r)\sim\sin(kr-\frac{1}{2}l\pi+\eta_l)$ and $\hat{\phi}_{lt}(r)$ has a similar asymptotic form.

3. Other Kohn-type variational principles for phase shifts

In deriving the VP's for $\tan\eta_l$ and η_l, we fixed the normalization of the exact wave function by the boundary condition at infinity. Thus we presumed $\hat{\phi}_l$ of Eq. (5.11) to be a sine function of unit amplitude at infinity, whereas for ϕ_l we presumed the coefficient of the sine term in its asymptotic form (5.1c) in effect to be $\sec\eta_l$. Correspondingly, we have employed trial wave functions having the same normalizations. However, the requirements on the amplitude at infinity of $\phi_l(r)$ and $\hat{\phi}_l(r)$ are specifications of limited quantitative type, of the sort discussed in Sec. III.C. It follows that these boundary conditions at infinity, on $\phi_l(r)$ and $\hat{\phi}_l(r)$, can be regarded as constraints, in which event they can be included, along with appropriate Lagrange multipliers, in the starting expressions for $\langle\tan\eta_l\rangle_{\rm var}$ and $\langle\eta_l\rangle_{\rm var}$, respectively. Were we to do this, we would find the same VP's as before even though trial functions not normalized precisely as in Eq. (5.1c) have

been used. The reason is that $\tan\eta_l$ can be defined independently of the normalization of ϕ_l, so that the Lagrange multiplier associated with the normalization constraint vanishes, just as $\mathscr{L}$ associated with the (irrelevant) phase constraint would vanish in a VP for $\phi^\dagger W\phi$.

If the normalization condition at infinity is incorporated into ϕ_{lt}, there is no reason whatever for the integral in the VP (5.7) to vanish. If, on the other hand, the normalization is regarded as an arbitrary parameter, that is, if ϕ_{lt} contains an arbitrary multiplicative constant C, then making Eq. (5.7) stationary with respect to variations of C requires the integral to vanish. This result is the basis for a method of employing Eq. (5.7) originally proposed by Hulthén (1948), namely, that the adjustable parameters (including $\tan\eta_{lt}$) in ϕ_{lt} be so chosen that the integral is zero, in which event $\tan\eta_{lt}$, so chosen, represents a variational estimate of $\tan\eta_l$.

Moiseiwitsch (1951) extended Hulthén's method to inelastic amplitudes, still for problems wherein the wave function is expanded in partial waves. The extension is not trivial, however, because for such inelastic problems permitting the normalization of ϕ_l to vary introduces correction terms which do not appear in potential scattering, so that merely setting the integral corresponding to that in Eq. (5.7) equal to zero need not yield a stationary estimate of the desired inelastic amplitude. Returning to elastic amplitudes, the relation between the estimates of $\tan\eta_l$ obtained via the Hulthén and Kohn methods—estimates in which the variational parameters are chosen slightly differently, but estimates which differ at most by second order, of course—and some of the difficulties in using these methods, both in composite collisions and in potential scattering, have been discussed by Demkov (1963), Harris and Michels (1971), Nesbet (1969), and Bardsley *et al.* (1972). This subject has been extensively reviewed in Callaway (1978) and Nesbet (1980).

As a matter of fact, the literature contains a variety of VP's for η_l, $\tan\eta_l$, and other functions $F(\eta_l)$, where by Kohn-type we mean that the VP contains ϕ_{lt} in the characteristic combination $\phi_{lt}^\dagger(H-E)\phi_{lt}$. We have found the alternative VP's we have examined to be derivable from our routine procedure, and we see no reason to doubt that any such VP is so derivable. Huang (1949a, 1949b, 1949c), Kato (1950, 1951), Feshbach and Rubinow (1952), and Rubinow (1954, 1955) are among the authors who have examined Kohn-type VP's not discussed in this section. Tamm (1948, 1949) derived a VP formalism very similar to that of Hulthén. Kohn (1951) reformulated his VP for $\tan\eta_l$ in momentum space. The authors listed in the preceding paragraphs also have examined some of the Kohn-type alternatives mentioned in this paragraph, especially the procedure advocated by Rubinow (1955); other relevant numerical results are quoted in Makinson and Turner (1953), Turner and Makinson (1953), Shimamura (1971a, 1971b) and Nesbet and Oberoi (1972). Bibliographies to the Kohn-type VP literature appear in Demkov (1963), Moiseiwitsch (1966), Callaway (1978), and Nesbet (1980).

B. Transition amplitudes in potential scattering

1. Derivation without explicit formula for the amplitude

We continue to consider only the potential scattering of spinless particles. The wave function ϕ_i satisfying the Schrödinger equation (1.3a) when a plane wave of momentum $\hbar\mathbf{k}_i$ is incident on the scattering center can be written in the form

$$\phi_i(\mathbf{r})=e^{i\mathbf{k}_i\cdot\mathbf{r}}+\chi_i(\mathbf{r}) , \tag{5.12}$$

where the scattered part χ_i is outgoing as $\mathbf{r}$ approaches infinity parallel to any final momentum $\hbar\mathbf{k}_f$, that is,

$$\chi_i(\mathbf{r})\sim A(\mathbf{n}_i\to\mathbf{n}_f)e^{ikr}/r , \tag{5.13}$$

as $\mathbf{r}\sim\infty$ parallel to $\mathbf{k}_f$. In (5.13), the quantities $\mathbf{n}_i$ and $\mathbf{n}_f$ are unit vectors, with $\mathbf{k}_i=k\mathbf{n}_i$ and $\mathbf{k}_f=k\mathbf{n}_f$. The amplitude for scattering from the incident direction $\mathbf{n}_i$ to the final direction $\mathbf{n}_f$, namely, $A(\mathbf{n}_i\to\mathbf{n}_f)$, often is reexpressed in the literature as

$$A(\mathbf{n}_i\to\mathbf{n}_f)=-(2m/4\pi\hbar^2)T_{fi} , \tag{5.14}$$

where T_{fi} is the transition amplitude, or matrix element of the transition operator T between the final and initial plane wave states. When $V(\mathbf{r})$ is spherically symmetric, stationary estimates of $A(\mathbf{n}_i\to\mathbf{n}_f)$ can, of course, be obtained by expanding in partial waves and employing the VP derived in Sec. V.A. Suppose, however, that $V(\mathbf{r})\neq V(r)$, or that $V(\mathbf{r})=V(r)$, but that we are interested in estimating $A(\mathbf{n}_i\to\mathbf{n}_f)$ variationally without expanding in partial waves. Our starting point, analogous to Eq. (5.2), is then

$$\begin{aligned}F_v&\equiv\langle A(\mathbf{n}_i\to\mathbf{n}_f)\rangle_{\rm var}\\&=A_t(\mathbf{n}_i\to\mathbf{n}_f)+L_t^\dagger[(H-E)\phi_{it}] ,\end{aligned} \tag{5.15}$$

where

$$\phi_{it}(\mathbf{r})=e^{i\mathbf{k}_i\cdot\mathbf{r}}+\chi_{it}(\mathbf{r}) , \tag{5.16a}$$

$$\chi_{it}(\mathbf{r})\sim A_t(\mathbf{n}_i\to\mathbf{n}_f)e^{ikr}/r , \tag{5.16b}$$

for $\mathbf{r}\sim\infty$ parallel to $\mathbf{n}_f$. We have not included in Eq. (5.15) a term involving $[(H-E)\phi_{it}]^\dagger$. The reason for this omission is that we are trying to estimate a quantity $A(\mathbf{n}_i\to\mathbf{n}_f)$ depending only on ϕ_i, and not at all on $\phi_i^\dagger$, as Eqs. (5.12) and (5.13) show. (The situation is thereby different from that encountered in Secs. III and IV. There we had bound states, and one of the defining equations, the normalization condition, specifically involved $\phi^\dagger$.) The correctness of this assertion will be explicitly demonstrated by our ability to derive a VP starting from Eq. (5.15); the reader can easily verify that the addition of a term involving $[(H-E)\phi_{it}]^\dagger$ complicates the derivation somewhat but leads to precisely the same VP we give below.

Varying Eq. (5.15), we have

$$\delta A(\mathbf{n}_i\to\mathbf{n}_f)+L^\dagger[(H-E)\delta\phi_i]=0 . \tag{5.17a}$$

Analysis of this equation is somewhat tedious, involving the transfer of $(H-E)$ to act on the left. We rewrite Eq. (5.17a) as

$$0=\delta A(\mathbf{n}_i\to\mathbf{n}_f)+[(H-E)L]^{\dagger}\delta\phi_i - \frac{\hbar^2}{2m}\int [L^*\nabla\delta\phi_i-(\nabla L)^*\delta\phi_i]\cdot\mathbf{n}r^2d\Omega \ , \quad (5.17b)$$

where $d\Omega$ is the differential solid angle, $\mathbf{n}$ is the unit vector along $\mathbf{r}$, and the surface integral is at infinity.

We therefore demand that $(H-E)L=0$, or, if we like,

$$(H-E)L^*=0 \ . \quad (5.18a)$$

Inserting $\delta\phi_i=\delta A(\mathbf{n}_i\to\mathbf{n})e^{ikr}/r$ into (5.17b), we then require

$$\delta A(\mathbf{n}_i\to\mathbf{n}_f)=\frac{\hbar^2}{2m}\int \delta A(\mathbf{n}_i\to\mathbf{n})\left[L^*\frac{\partial}{\partial r}\frac{e^{ikr}}{r} - \frac{\partial L^*}{\partial r}\frac{e^{ikr}}{r}\right]r^2d\Omega \ .$$

We therefore require that at infinite r

$$(\hbar^2/2m)r^2\left[L^*\frac{\partial}{\partial r}\frac{e^{ikr}}{r}-\frac{\partial L^*}{\partial r}\frac{e^{ikr}}{r}\right]=\delta(\mathbf{n}-\mathbf{n}_f) \ . \quad (5.18b)$$

At infinite r, the function L^*—which is a solution of Eq. (5.18a)—must be a linear combination of outgoing spherical waves (proportional to e^{ikr}/r) and incoming spherical waves (proportional to e^{-ikr}/r). However, the outgoing part obviously makes no contribution to the left-hand side of Eq. (5.18b). From the Dirac identity (Dirac, 1958, p. 191) asserting

$$r^2\left[e^{-ik\mathbf{n}_f\cdot\mathbf{r}}\frac{\partial}{\partial r}\frac{e^{ikr}}{r}-\frac{e^{ikr}}{r}\frac{\partial}{\partial r}e^{-ik\mathbf{n}_f\cdot\mathbf{r}}\right]\sim -4\pi\delta(\mathbf{n}_f-\mathbf{n}) \ , \quad (5.19a)$$

as $\mathbf{r}\sim\infty$ along $\mathbf{n}$, we therefore conclude that the incoming spherical part of L^* at infinity coincides with the incoming spherical part of $(-m/2\pi\hbar^2)\exp(-ik\mathbf{n}_f\cdot\mathbf{r})$.

In other words, at infinity we have

$$L^*(\mathbf{r})\equiv L^*(\mathbf{r};\mathbf{n}_f) = -(m/2\pi\hbar^2)[e^{-ik\mathbf{n}_f\cdot\mathbf{r}}+\hat{\chi}(\mathbf{r})] \ , \quad (5.19b)$$

where $\hat{\chi}(\mathbf{r})$ is everywhere outgoing at infinity. But we also known that $L^*(\mathbf{r})$ is a solution of the Schrödinger equation (5.18a), whose solutions are uniquely specified by their plane-wave parts. Thus we have, recalling Eq. (5.12),

$$L^*(\mathbf{r})=-(m/2\pi\hbar^2)\phi_{-f}(\mathbf{r}) \ , \quad (5.20a)$$

where

$$\phi_{-f}(\mathbf{r})\equiv[\phi_f^{(-)}(\mathbf{r})]^*=e^{-ik\mathbf{n}_f\cdot\mathbf{r}}+\chi_{-f}(\mathbf{r}) \quad (5.20b)$$

represents a scattered solution of the same type (outgoing scattered part) as our incident wave, but incident along the direction $-\mathbf{n}_f$ rather than $\mathbf{n}_i$.

In view of Eqs. (5.20), the obvious choice of trial L_t in Eq. (5.15) is

$$L_t(\mathbf{r})\equiv L_t(\mathbf{r};\mathbf{n}_f)=-(m/2\pi\hbar^2)\phi^*_{-ft}(\mathbf{r}) \ , \quad (5.21a)$$

where $\phi_{-ft}(\mathbf{r})$ is a trial estimate of the exact $\phi_{-f}(\mathbf{r})$, and has the form

$$\phi_{-ft}(\mathbf{r})=e^{-ik\mathbf{n}_f\cdot\mathbf{r}}+\chi_{-ft}(\mathbf{r}) \ , \quad (5.21b)$$

with $\chi_{-ft}(\mathbf{r})$ everywhere outgoing at infinity. Note that the "scattered parts" χ^*_{-f} and χ^*_{-ft}, of ϕ^*_{-f} and ϕ^*_{-ft}, respectively, are everywhere incoming at infinity; that is, $L(\mathbf{r})$ is proportional to a "time-reversed" scattering solution $\phi^*_{-f}(\mathbf{r})\equiv\phi_f^{(-)}(\mathbf{r})$ of the Schrödinger equation, time reversed because it can be interpreted as composed of a purely incoming spherical wave part χ^*_{-if} which reforms as a plane-wave $\exp(ik\mathbf{n}_f\cdot\mathbf{r})$ going out to infinity along the direction $\mathbf{n}_f$. Using Eqs. (5.21) in Eq. (5.15) yields precisely the Kohn VP for the scattering amplitude $A(\mathbf{n}_i\to\mathbf{n}_f)$.

We add that time-reversal invariance requires that the asymptotic forms of the exact $\phi_i(\mathbf{r})$ and $\phi_{-f}(\mathbf{r})$ be consistent with

$$A(\mathbf{n}_i\to\mathbf{n}_f)=A(-\mathbf{n}_f\to-\mathbf{n}_i) \ .$$

If the asymptotic forms of the trial functions $\phi_{it}(\mathbf{r})$ and $\phi_{-ft}(\mathbf{r})$ are similarly chosen, i.e., if

$$A_t(\mathbf{n}_i\to\mathbf{n}_f)=A_t(-\mathbf{n}_f\to-\mathbf{n}_i) \ ,$$

then it can be proved that the variational estimates of $A(\mathbf{n}_i\to\mathbf{n}_f)$ and $A(-\mathbf{n}_f\to-\mathbf{n}_i)$, computed from their corresponding Kohn VP's, will be identical.

2. Derivation using explicit formula for the amplitude

The above demonstration that the Kohn VP for $A(\mathbf{n}_i\to\mathbf{n}_f)$ (or, equivalently, for T_{fi}) really is stationary differs from the usual demonstration in the literature in that our approach is based on our general prescription for the construction of VP's and deduces the result straightforwardly. In particular, we deduce the information that the adjoint function L^* is proportional to the time-reversed scattering solution defined by Eqs. (5.20). Our approach suggests that it is not likely to be possible to find a simple nontrivial alternative to Eq. (5.15) with L_t given by Eq. (5.21a).

This last remark is illustrated by the following comment on an alternative derivation of the Kohn VP for T_{fi}, still in the context of potential scattering. It is well known that T_{fi} can be expressed as

$$T_{fi}=\psi_f^{\dagger}V\phi_i \ , \quad (5.22a)$$

where the free plane wave ψ_f is given by

$$\psi_f\equiv e^{ik\mathbf{n}_f\cdot\mathbf{r}} \ . \quad (5.22b)$$

Equation (5.22a) is an analog of Eq. (5.8a). Starting from

Eq. (5.22a) as our defining equation for T_{fi}, we are led relentlessly by the systematic approach to the VP given by Eqs. (5.15) and (5.21).

3. Remarks concerning Kohn-type variational principles

We remark that the adjoint functions we have introduced in the VP of Sec. V.A related to partial wave analyses, and in the Kohn VP for T_{fi} just discussed, always turned out to be closely related to a scattering solution of the Schrödinger equation. As a result, the final VP is quadratic or bilinear in the (unknown) scattering solutions. This point sometimes obscures the linearity of the starting point of the Kohn VP. This linearity makes the derivation of the Kohn VP much simpler in a formal sense than would otherwise have been the case. This point has not always been appreciated in the literature.

The reason the adjoint functions in VP's for scattering parameters turn out to be intimately related to continuum solutions of the Schrödinger equation under consideration must be associated with the fact that the Hamiltonian H is the only consequential operator involved in scattering problems. As noted in Sec. IV.B, adjoint functions would be more complicated, depending also on W, if we were to seek VP's for continuum matrix elements of some arbitrary operator W. The situation is analogous to that which obtains for bound-state matrix elements. There, the adjoint function is proportional to the bound-state wave function when considering a VP for the energy of the bound state, the matrix element of H, but not for a VP for any other matrix element.

We also note that in our derivations of the Kohn VP's we could have taken advantage of the known time-reversed properties of these quantities to express $A(\mathbf{n}_i \rightarrow \mathbf{n}_f)$ in terms of the asymptotic behavior of $\phi_f^{(-)*}(\mathbf{r}) \equiv \phi_{-f}$ given by Eq. (5.20b). The starting expressions would then have involved $(H-E)\phi_{-ft}$ rather than $(H-E)\phi_{it}$, and the associated L would have turned out to be proportional to ϕ_i rather than to ϕ_{-f}. The final VP would have been identical to that derived above.

Further, we observe that the result in Eq. (5.7) can be regarded as a VP for an element of the reaction operator, namely $\tan\eta_l$, after decomposition into partial waves. We could also derive a Kohn VP for the matrix elements of the entire reaction operator, before decomposition into partial waves, just as we were able to derive a Kohn VP for matrix elements T_{fi} of the transition operator. The reader is reminded that the reaction operator matrix elements are the quantities characteristically involved in standing-wave descriptions of scattering processes, just as the transition operator matrix elements characteristically are involved in descriptions invoking incident waves plus outgoing scattered waves. In the case of potential scattering, the diagonal elements of the reaction operator in a partial wave representation are $\tan\eta_l$; the corresponding diagonal elements of the transition operator are $\exp(i\eta_l)\sin\eta_l$.

We would now like to comment on a practical problem associated with the use of the Kohn VP for an incident kinetic energy greater than zero, namely, the inversion of the singular operator $H-E$. More precisely, having chosen a trial function containing N linear variational parameters c_n, we choose c_n so as to make $\langle \tan\eta \rangle_v$ stationary. This requires the inversion of $H-E$ in an N-by-N space. Since E is in the continuous spectrum of H, one or more of the eigenvalues of H in the N-by-N space might be very close to E. For N small this will not normally pose a problem, but as we increase N in an attempt to increase the accuracy of the calculation, an eigenvalue may well be very close to E, and the estimate of $\langle \tan\eta \rangle_v$ can be far off. The usual procedure under these circumstances, indicated by numerical instability of the results, is then to consider $\langle \cot\eta \rangle_v$, for one will not have numerical problems for this case too. More generally, Kato (1950), as a purely formal device, introduced an arbitrary parameter θ in the asymptotic form of the wave function. Thus, for $l=0$ potential scattering, Kato wrote in place of Eq. (5.1c)

$$\phi_0(r) \sim \sin(kr+\theta)+\tan(\eta-\theta)\cos(kr+\theta)$$
$$= \sin(kr+\eta)/\cos(\eta-\theta) \, .$$

One thereby generates a VP for $\tan(\eta-\theta)$, which encompasses the results for $\langle \tan\eta \rangle_v$ and $\langle \cot\eta \rangle_v$ for the choice of $\theta=0$ and $\pi/2$, respectively, but allows a continuous choice of θ. The first use of the freedom thereby allowed to avoid a spurious infinity seems to have been made by Kleinman *et al.* (1965). A more recent work on this is Takatsuka and Fueno (1979). Brownstein and McKinley (1968) studied the nature of the singularities which arise on the inversion of $H-E$ in the N-by-N space, and a more thorough analysis was given by Nesbet, to whom references can be found in Nesbet (1980). An alternative prescription was given by Wladawsky (1973). A numerical procedure which in a sense avoids the singularity in $(H-E)^{-1}$ was recently obtained by Winick and Reinhardt (1978).

The situation is much simpler at zero-incident kinetic energy. Again restricting ourselves to $l=0$ potential scattering for simplicity, we have $\langle A \rangle_{\text{var}}$ given by Eq. (5.8d). The evaluation of the linear parameters c_n now requires the inversion of H, and $H^{-1} \equiv G$ is *not* a singular operator. Thus, if for pedagogic purposes we make the further simplification to the free Green's function, G_0, at an energy $\hbar^2 k^2/2m$, we have

$$G_0(r,r';k) = \text{const} \int \frac{\sin k'r \sin k'r'}{k'^2-k^2} k'^2 dk' \, ,$$

and, for all k *but* $k=0$, we must take a principal value to avoid the singularity in the integrand at $k'=k$. At $k=0$, the k'^2 in the denominator is cancelled by the k'^2 in the volume element, and one can say that there is no singularity at $k=0$ because the density of states at zero momentum vanishes. The H^{-1} is not only nonsingular but nonnegative (apart, normally, from a finite number of negative eigenvalues, which can be accounted for even when the eigenvalues are only imprecisely known), and this en-

ables one to obtain quite simply not only a VP for A but an upper variational bound on A (Spruch and Rosenberg, 1959; Rosenberg *et al.*, 1960). The same is true for scattering by a compound system *if* the target eigenfunctions and eigenvalues are known exactly; we consider this in Sec. V.C. The problems which arise in the analysis of the scattering length when the target eigenfunctions and eigenvalues are only imprecisely known will be discussed in Sec. VI.C. (The scattering wave function need not, of course, be known.)

C. Collisions involving composite systems

In this section, we wish to derive the Kohn VP for collisions between composite systems; to treat such collisions our potential scattering treatment must be generalized. Some of the difficulties in the more general case are simply notational. To spare ourselves from as many of these notational problems as possible, we shall assume that the colliding systems are composed of distinguishable spinless particles; for our purposes, this particular simplification retains the essential features which distinguish the many-particle collisions from the potential scattering case. Furthermore, we consider only collisions in which just two systems are incident, and we restrict our considerations to collision systems and energies for which all allowed reactions have just two emergent systems,

$$A+B\rightarrow C+D\ , \tag{5.23}$$

where A, B, C, and D are aggregates (groupings) of the particles involved in the collisions. In other words, only elastic scattering, inelastic excitation or deexcitation, and rearrangement are permitted to occur; breakup is not. Without this restriction to reactions of type (5.23), Green's theorem in the many-particle configuration space becomes much more awkward. The possibility of breakup will be touched upon later. For simplicity only, we also shall assume that the incoming and outgoing pairs do not both have a net charge, for we wish to avoid the complexities associated with Coulomb functions.

With these assumptions we can say that the center of mass wave function describing the collision resulting in the reaction (5.23) has the form

$$\phi_i=\psi_i+\chi_i\ , \tag{5.24}$$

where the scattered part χ_i is everywhere outgoing, and where the incident wave,

$$\psi_i=u_i^{AB}\exp(ik_i^{AB}\mathbf{n}_i\cdot\mathbf{r}^{AB})\ , \tag{5.25}$$

represents A and B approaching each other with relative momentum $\hbar k_i^{AB}\mathbf{n}_i$, with initial bound-state wave functions u_i^A and u_i^B, whose product $u_i^Au_i^B$ is denoted by u_i^{AB}; we do not restrict ourselves to scattering by the targets in their ground state. In Eq. (5.25), $\mathbf{r}^{AB}$ is the displacement of the center of mass of A relative to that of B; the sum of the initial bound-state energies of A and B is ε_i^{AB}, and the total energy in the center-of-mass frame is

$$E=\varepsilon_i^{AB}+(\hbar k_i^{AB})^2/2\mu_{AB}\ , \tag{5.26a}$$

where μ_{AB} is the reduced mass of A,B. It is convenient to index the various pairs of bound states of A,B by the subscript γ, that is, these bound-state wave functions are $u_\gamma^{AB},\gamma=1,2,\ldots,$ with corresponding energies ε_γ^{AB}. Then, if the pair A,B propagate to infinity in the bound state u_γ^{AB}, the corresponding magnitude $\hbar k_\gamma^{AB}$ of their relative momentum will be determined by

$$E=\varepsilon_\gamma^{AB}+(\hbar k_\gamma^{AB})^2/2\mu_{AB}\ . \tag{5.26b}$$

Similarly, for any energetically allowed pair of aggregates P,Q not necessarily identical with the initial pair A,B or the final pair C,D we will have bound-state pair wave functions u_γ^{PQ}, corresponding energies ε_γ^{PQ}, and corresponding momenta $\hbar\mathbf{k}_\gamma^{PQ}$.

With these introductory remarks—all primarily concerned with specifying the notation—out of the way, we now can assert that the everywhere outgoing scattered part χ_i in Eq. (5.24) has the asymptotic behavior for any possible outgoing pair P,Q

$$\chi_i\sim\sum_\gamma A_\gamma^{PQ}(\mathbf{n}_i\rightarrow\mathbf{n})u_\gamma^{PQ}(1/r^{PQ})\exp(ik_\gamma^{PQ}r^{PQ})\ , \tag{5.27}$$

as $\mathbf{r}^{PQ}\sim\infty$ parallel to $\mathbf{n}$, where the sum is over all γ associated with energetically allowed outgoing pairs. Our objective is to estimate variationally the amplitude $A_\Gamma^{CD}(\mathbf{n}_i\rightarrow\mathbf{n}_f)$, which determines the differential cross section for the reaction (5.23), with the initial state defined by Eq. (5.25) and with C,D in the particular eigenstate u_Γ^{CD} moving outward with relative velocity along $\mathbf{n}_f$ in the center-of-mass system. For the present, let us assume there are no uncertainties concerning the eigenstates u_γ^{AB}, u_γ^{CD}, and u_γ^{PQ} or their energies; we will return to this assumption later. Then, as in Eq. (5.15), our starting expression is

$$\begin{aligned}F_v&\equiv\langle A_\Gamma^{CD}(\mathbf{n}_i\rightarrow\mathbf{n}_f)\rangle_{\rm var}\\&=A_{\Gamma t}^{CD}(\mathbf{n}_i\rightarrow\mathbf{n}_f)+L_t^\dagger[(H-E)\phi_{it}]\ ,\end{aligned} \tag{5.28}$$

where E is known. In Eq. (5.28),

$$\phi_{it}=\psi_i+\chi_{it}\ , \tag{5.29}$$

where the asymptotic form of χ_{it} differs from that of χ_i in Eq. (5.27) only in the presence of a subscript t on the scattering amplitude.

The derivation of the VP now proceeds in quite close analogy with Eqs. (5.17)—(5.21), and finally we have

$$L^*(\mathbf{r})=-(\mu_{CD}/2\pi\hbar^2)\phi_{-f}(\mathbf{r})\ , \tag{5.30a}$$

$$\begin{aligned}\phi_{-f}(\mathbf{r})&=(u_\Gamma^{CD})^*\exp(-ik_\Gamma^{CD}\mathbf{n}_f\cdot\mathbf{r}^{CD})+\chi_{-f}(\mathbf{r})\\&\equiv\psi_f^*+\chi_{-f}\equiv\phi_f^{(-)*}(\mathbf{r})\ ,\end{aligned} \tag{5.30b}$$

where χ_{-f} is everywhere outgoing. From Eqs. (5.30) and (5.28) we infer that the desired VP is

$$\begin{aligned}\langle A_\Gamma^{CD}(\mathbf{n}_i\rightarrow\mathbf{n}_f)\rangle_{\rm var}&=A_{\Gamma t}^{CD}(\mathbf{n}_i\rightarrow\mathbf{n}_f)\\&\quad-\frac{\mu_{\rm CD}}{2\pi\hbar^2}\phi_{ft}^{(-)\dagger}[(H-E)\phi_{it}]\ ,\end{aligned} \tag{5.31a}$$

where

$$\phi_{ft}^{(-)}=\psi_f+\chi^*_{-ft}(\mathbf{r}) , \tag{5.31b}$$

with $\chi^*_{-ft}(\mathbf{r})$ everywhere incoming at infinity.

The foregoing analysis, culminating in Eqs. (5.31), provides the desired Kohn VP for the amplitude $A_\Gamma^{CD}(\mathbf{n}_i\to\mathbf{n}_f)$ describing the reaction (5.23), when the various assumptions stated at the beginning of this section hold, and when the bound-state wave functions u_γ^{PQ} and associated energies ε_γ^{PQ} all are exactly known. The assumption that breakup processes do not occur can be relaxed—that is, one can variationally estimate the amplitude for three-body breakup

$$A+B\to C+D+E , \tag{5.32}$$

in which no two of the bodies simultaneously going out to infinity are charged; as an example, $n+d\to n+n+p$ is a possible process we can study but not the counterpart atomic problem $e+A\to e+e+A^+$. It also is permissible to treat reactions (5.23) or (5.32) when three other outgoing bodies $P,Q,R\neq C,D,E$ of (5.32) are energetically able to reach infinity. In such cases, with proper definitions of the scattering amplitudes or associated matrix elements T_{fi} of the transition operator, Eq. (5.15) continues to give the desired VP, though Green's theorem now is much more complicated in form and though there are some subtle mathematical questions connected with the proper definition of the time-reversed solution $\phi_f^{(-)}$ for the reaction (5.32) (Gerjuoy, 1971). A generalization of the Dirac identity (5.19a) is useful. For the $3b$-dimensional vector $\mathbf{r}=(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_b)$, where b is the number of bodies reaching infinity, the required generalization is (Rosenberg, 1964; Lieber *et al.*, 1972a)

$$r^{n-1}\left[e^{-ik\mathbf{n}_f\cdot\mathbf{r}}\frac{\partial}{\partial r}(e^{\pm ikr}/r^d)-(e^{\pm ikr}/r^d)\frac{\partial}{\partial r}e^{-ik\mathbf{n}_f\cdot\mathbf{r}}\right]\sim 2^{d+1}\pi^d k^{1-d}e^{\pm i\pi(d+1)/2}\delta(\mathbf{n}\mp\mathbf{n}_f) , \tag{5.33}$$

where $n=3b$, $d=(n-1)/2$, and $\mathbf{n}_f$ denotes a unit vector in the same $3b$-dimensional space. Equation (5.33) is usable as it stands when $\mathbf{r}_i$ denotes the position of the ith particle in the laboratory system. Equation (5.33) also is usable in the center-of-mass system, provided b is replaced by $b-1$ and the now $(3b-3)$-dimensional vector $\mathbf{r}$ is suitably defined. For two particles in the center-of-mass system, $\mathbf{r}$ is the interparticle distance in ordinary space and Eq. (5.33) quite obviously reduces to Eq. (5.19a).

Detailed "ingenious" derivations [not always taking advantage of Eq. (5.33)] of the VP to which Eq. (5.15) leads in collisions permitting breakup have been given. Nuttall (1967) developed a Kohn-type expression for the elastic scattering amplitude (which was not, however, stationary) when one particle is incident on a bound pair at energies above the breakup threshold. Pieper, Schlessinger, and Wright (1970), Pieper, Wright, and Schlessinger (1970), and Lieber *et al.* (1972a) gave convincing but not rigorous arguments for the validity of the VP resulting from Eq. (5.15) for breakup scattering, and Lieber *et al.* (1972b) made these arguments rigorous through the use of Faddeev equations.

D. Variational identities and generalizations of the Kohn principle

The Kohn VP can be derived via an approach which makes it transparent that the adjoint function L always is itself a scattering function, as we discovered in the various derivations successfully carried through in Secs. V.A–V.C above. This alternative approach involves some ingenuity, as will be seen, and therefore, strictly speaking, falls outside our declared objective of illustrating our routine procedure for deriving VP without exercising ingenuity. Nevertheless, we will discuss this alternative approach because it is simple, has been widely employed in the literature, and has some ramifications which are of general interest and worth noting.

We start with consideration of the expression

$$\phi_f^{(-)\dagger}[(H-E)\phi_{it}]-[(H-E)\phi_f^{(-)}]^\dagger\phi_{it} ,$$

where $\phi_f^{(-)}$ has been defined in Eq. (5.30b). On the one hand, we can use $(H-E)\phi_f^{(-)}=0$. On the other hand, we can use Green's theorem in our multidimensional space to reduce the expression to a surface term; this is quite simple if we restrict ourselves to two-body reactions (5.23). [Three-body breakup reactions can also be handled by this approach, as Lieber *et al.* (1972a) have shown.] For the two-body case, the surface term, which involves the asymptotic forms of ϕ_{it} and $\phi_f^{(-)}$, introduces the amplitudes T_{fit} and T_{fi}; in fact, the surface term is seen to be $T_{fi}-T_{fit}$, and we arrive at the variational identity

$$T_{fi}=T_{fit}+\phi_f^{(-)\dagger}[(H-E)\phi_{it}] . \tag{5.34}$$

However, because ϕ_i satisfies the Schrödinger equation, the last term in Eq. (5.34) is already of first order, so that replacing $\phi_f^{(-)}$ by $\phi_{ft}^{(-)}$ makes only a second-order error. Consequently, we have, as a variational estimate of T_{fi},

$$\langle T_{fi}\rangle_{\text{var}}=T_{fit}+\phi_{ft}^{(-)\dagger}[(H-E)\phi_{it}] . \tag{5.35}$$

We see that we have given an alternative derivation of the Kohn VP in its form (5.15); simultaneously we have made it clear that in our routine starting expression (5.15) the choice L_t proportional to $\phi_{ft}^{(-)}$ [consistent with Eqs. (5.30) and (5.31)] is correct, that is, will yield a variational estimate of T_{fi}.

The variational identity in Eq. (5.34) was first given (Kato, 1951) for potential scattering of a given partial wave. The literature contains the details for potential scattering without a partial-wave decomposition (Spruch, 1962) and for the elastic and inelastic scattering of sys-

tems that can contain identical particles (Rosenberg and Spruch, 1960). Gerjuoy *et al.* (1972) and Gerjuoy, Rau *et al.*, (1975) have pointed out that this type of identity, associated with Kohn-type VP's, can be generalized to all VP's; if we start with our standard formulation, e.g., Eq. (5.15), the step of equating to zero the sum of all first-order terms gives the identity quite generally. Correspondingly, the Kohn VP (as first pointed out in Sec. V.B.3) itself can be considered as a special case of a wider class of VP's, wherein the objective is to estimate variationally quantities $F(\phi)$ which are linear in the unknown ϕ, of the form

$$F=g^{\dagger}\phi \ , \tag{5.36}$$

where g is a known function, and where ϕ is an unknown function, defined by

$$M\phi-\omega=0 \ , \tag{5.37}$$

where ω is a known function, M is a known linear operator, and both ω and M are ϕ independent. Evidently, the Schrödinger equation (1.3a) is of the form (5.37), while appropriate choices of g in Eq. (5.36) generate functionals F which reproduce $\tan\eta_l$ of Eq. (5.8a), T_{fi} of Eq. (5.22a), and ϕ itself. A VP for F will, therefore, produce VP's for $\tan\eta_l$, T_{fi}, and ϕ. We note too, restricting ourselves to potential scattering, that the Schrödinger equation in integral form—the Lippmann-Schwinger (1950) equation—

$$(1+G_{\text{free}}V)\phi-\phi_{\text{free}}=0 \ ,$$

where $G_{\text{free}}\equiv(H_{\text{free}}-E)^{-1}$ and ϕ_{free} are the free Green's function and the incident free wave function, respectively, is also of the form (5.37). The use of the integral equation as the defining equation for ϕ leads in straightforward fashion to a new set of VP's for $\tan\eta_l$ and T_{fi}, the Schwinger VP's, or for ϕ itself, but we will not spell out the details.

It is trivial to obtain a VP for the F defined by Eqs. (5.36) and (5.37). We find, proceeding as usual, that

$$F_v\equiv(g^{\dagger}\phi)_{\text{var}}=g^{\dagger}\phi_t+L_t^{\dagger}(M\phi_t-\omega) \ ,$$

where L_t is an approximation to L, defined by

$$M^{\dagger}L+g=0 \ ,$$

and where we have assumed the validity of the boundary condition

$$L^{\dagger}(M\delta\phi)-(M^{\dagger}L)^{\dagger}\delta\phi=0 \ .$$

We would also like to note that the replacement of L_t by L in the VP for $g^{\dagger}\phi$ is the identity

$$g^{\dagger}\phi=g^{\dagger}\phi_t+L^{\dagger}(M\phi_t-\omega) \ ,$$

which can be regarded as a generalization of the Kato-type identity (5.34).

E. Variational principle for the ratio of linear functionals

We will conclude our discussion of Kohn-type VP's and their generalizations by deriving, via our routine procedure, a VP for

$$F(\phi)=(g_1^{\dagger}\phi)/(g_2^{\dagger}\phi) \ , \tag{5.38}$$

where ϕ again satisfies Eq. (5.37) and g_1 and g_2 are known functions. In so doing, we will provide another illustration of the applicability of our procedures to VP's derived less systematically in the literature. Our derivation for Eq. (5.38) is very similar to that given in Gerjuoy *et al.* (1972). We have

$$F_v=(g_1^{\dagger}\phi_t)/(g_2^{\dagger}\phi_t)+L_t^{\dagger}[M\phi_t-\omega] \ . \tag{5.39}$$

We must then have

$$\delta F_v=(g_1^{\dagger}\delta\phi)/(g_2^{\dagger}\phi)-(g_1^{\dagger}\phi)(g_2^{\dagger}\delta\phi)/(g_2^{\dagger}\phi)^2 +L^{\dagger}M\delta\phi=0 \ .$$

L must therefore satisfy

$$[M^{\dagger}L]^{\dagger}+(g_1^{\dagger})/(g_2^{\dagger}\phi)-[(g_1^{\dagger}\phi)/(g_2^{\dagger}\phi)^2]g_2^{\dagger}=0 \ , \tag{5.40}$$

subject to the boundary condition

$$L^{\dagger}[M\delta\phi]=[M^{\dagger}L]^{\dagger}\delta\phi \ .$$

When we introduce $\hat{L}$ via

$$\hat{L}=-[(g_2^{\dagger}\phi)^{\dagger}/(g_1^{\dagger}\phi)^{\dagger}]L \ ,$$

Eq. (5.39) becomes

$$F_v=[(g_1^{\dagger}\phi_t)/(g_2^{\dagger}\phi_t)]\{1-\hat{L}_t^{\dagger}[M\phi_t-\omega]\} \ , \tag{5.41a}$$

with $\hat{L}_t$ a first-order estimate of the exact $\hat{L}$, which satisfies

$$[M^{\dagger}\hat{L}]^{\dagger}=(g_1^{\dagger})/(g_1^{\dagger}\phi)-(g_2^{\dagger})/(g_2^{\dagger}\phi) \ , \tag{5.41b}$$

a result due to Stacey (1972). [We could of course have obtained a VP for $F(\phi)$ defined by Eq. (5.38) by using the results of the preceding subsection to construct separate VP's for the numerator and denominator, but the present procedure is simpler.]

VI. SCATTERING VARIATIONAL PRINCIPLES, FOR TARGET WAVE FUNCTIONS IMPRECISELY KNOWN

A. Some background

It first became possible to make reasonably accurate estimates of scattering parameters for scattering by compound systems about 1960, because the advent of high-speed computers made it feasible to begin to use the full power of VP's. Indeed, as of 1960, the best estimate of the scattering length A for positron-hydrogen atom (e^+-H) scattering—a relatively simple system by present standards, though the distortion of the H atom by the incident e^+ is considerable—gave the wrong sign for A, the

sign associated with an effectively repulsive potential. At that time, the variational estimates, based on trial functions which contained only two or three variational parameters, varied widely, and there were no objective criteria for judging among VP calculations with roughly the same number of parameters—nor are there such criteria now for VP's. The problem was also analyzed through a variational *bound* approach by Spruch and Rosenberg (1960). For variational upper bounds, whether for scattering lengths or energies or whatever, the rules of the game are objective and trivial—the winning estimate is the lowest estimate, independent of the number of parameters. As the number of parameters one could reasonably use increased to being of the order of 30, variational estimates became much more reliable, and variational bounds were no longer so necessary, at least for simple, well-understood systems.

In the 1960s an enormous effort went into the study of low-energy e^--H scattering, with essentially all calculations based on VP's, apart from a few based on the more powerful variational bounds. (Some calculational techniques which proved fruitful were not originally known to be variational in character, and were only later shown to be so.) One calculated elastic and inelastic total and differential cross sections, the positions and widths of the lower of the infinite number of resonances, etc. There are still problems with the analysis of e^--H scattering at intermediate energies from perhaps 13 up to 100 eV or so, but the main emphasis in e^--atom scattering during the 1970s and now is on the scattering by atoms other than H. For that case, and indeed for almost any but comparatively simple collisions, such as e^+ or e^- or H atom scattering by H atoms, one is faced with a new difficulty: not knowing the target bound-state eigenfunctions (nor the energy eigenvalues). Thus not only does one not know the scattering wave function ϕ—the usual case for a scattering problem—but, even apart from a lack of knowledge of the scattering parameters involved, one does not know ϕ in the asymptotic region, since ϕ asymptotically approaches a sum of products, with each product containing an open-channel target bound-state wave function which is unknown. This can introduce a host of formal *and* practical difficulties; integrals can diverge, and contributions to scattering parameters can be difficult to classify as zeroth, first, or second order. These difficulties have not always been fully appreciated, and there are a number of incorrect statements in the literature in this area.

B. The Demkov variational principle

In discussing scattering by compound systems, we revert to our earlier assumption that breakup cannot take place, that is, that only two-body reactions like (5.23) can occur, but we suppose now that the bound-state eigenfunctions u_γ^{PQ} and their energies ε_γ^{PQ} are only imprecisely known. The energy E of Eq. (5.26a) can thus no longer be regarded as known. The initial relative momentum $\hbar k_i^{AB}$ is still known exactly, having been chosen by the experimentalist, but all other k_γ^{PQ} are not known. Finally, as noted above, we will be unable to choose trial scattering functions which satisfy the appropriate asymptotic boundary conditions, since the latter require a knowledge of all bound-state wave functions associated with open channels. The VP (5.31a), if we are concerned with a scattering amplitude, or the extension of Eq. (5.7) to scattering by a compound target, if we have made a partial wave analysis, is now a formal one, and the question naturally arises as to whether one can develop a Kohn-type VP when some of the relevant target properties are only imprecisely known.

In order to avoid being buried in a morass of subscripts, superscripts, and simple, but lengthy algebra, we begin not with the general case but rather with a relatively simple case which nevertheless contains the essence of the problem.

A prototype problem for the simple case is the scattering of e^+ by a helium atom, in its ground state, at energies below the excitation and pickup thresholds. More generally, we assume that each of the aggregates A and B is incident in its ground state—the subscript i can then be replaced by the subscript 1—and we further assume that the nature of A and B, and the incident kinetic energy

$$\hbar^2(k_1^{AB})^2/2\mu_{AB} \equiv \varepsilon' \tag{6.1}$$

are such that elastic scattering (without exchange) is the only possible process. It is no longer necessary to retain our subscripts and superscripts; we can replace the relative momentum k_1^{AB} by k, the product of ground-state wave functions $u_1^{AB} = u_1^A u_1^B$ by u, the sum of ground-state energies $\varepsilon_1^{AB} = \varepsilon_1^A + \varepsilon_1^B$ by ε, the relative coordinate $\mathbf{r}^{AB}$ by $\mathbf{r}$, the reduced mass μ_{AB} by μ, and the interaction V_{AB} by V. The total energy in the center-of-mass frame is given by Eq. (5.26a), which can now be written as

$$E = \varepsilon + (\hbar k)^2/2\mu \equiv \varepsilon + \varepsilon' \ . \tag{6.2}$$

We perform a partial-wave analysis, with the wave function ϕ_l (which again is the lth partial wave multiplied by r) satisfying

$$(H-E)\phi_l = (H_A + H_B + V + T_{AB} - \varepsilon - \varepsilon')\phi_l = 0 \ , \tag{6.3}$$

where $\phi_l \sim u\theta_{l\infty}$ as $r \sim \infty$, with $\theta_{l\infty} = \theta_{l\infty}(r)$ of the same form as the $\phi_{l\infty}$ defined by Eq. (5.1c); at infinite r the relative kinetic energy operator T_{AB} in (6.3) reduces to T_l of Eq. (5.1b). If we introduce a trial function ϕ_{lt} in the usual way, and assume for the moment that u is known, the Kohn VP is

$$\langle \tan\eta \rangle_v = \tan\eta_t - (2\mu/\hbar^2 k) I_t \ , \tag{6.4}$$

where

$$I_t \equiv \phi_{lt}^\dagger (H - \varepsilon - \varepsilon')\phi_{lt} \ . \tag{6.5}$$

For u only imprecisely known, a natural procedure retaining the VP property would be to replace u by a variational estimate u_v (differing from u by quantities of second order) and simultaneously to replace ε by

$\varepsilon_v \equiv u_v^\dagger [H_A + H_B] u_v$. (If one were to replace u by u_v but not change ε, the new I_t thereby generated would diverge.) While such a procedure would be a perfectly valid one, the evaluation of matrix elements containing u_v would be relatively difficult. It would be far simpler, indeed a blessing, if we could replace u in I_t by a first-order estimate u_t, and could correspondingly replace ε in I_t by $\varepsilon_t \equiv u_t^\dagger [H_A + H_B] u_t$ (implying u_t is normalized), thereby replacing I_t by

$$I_{tt} \equiv \phi_{ltt}^\dagger (H - \varepsilon_t - \varepsilon') \phi_{ltt} , \tag{6.6}$$

where $\phi_{ltt} \sim u_t \theta_{l\infty t}$ as $r \sim \infty$, and the double t subscripts on ϕ_{ltt} and on I_{tt} denote the fact that the full scattering wave function *and* the target bound-state wave functions are represented by trial functions. It would *seem* that the replacement of I_t by I_{tt}, Eq. (6.6), would introduce a first-order error and would thereby violate the VP, but Demkov (1963) has shown that

$$\langle \tan\eta \rangle_v \equiv \tan\eta_t - (2\mu/\hbar^2 k) I_{tt} , \tag{6.7}$$

with I_{tt} defined by Eq. (6.6), is indeed variational.

Demkov's proof will not be repeated here, for although the result is of great importance, the proof is straightforward (though care must be exercised in the treatment of surface terms), tedious, and unilluminating. At this point we *should* present a derivation of the Demkov VP following our routine procedure. We believe a presentation somewhat along those lines may well be possible, and we intend to look into that possibility. At this moment, however, we must state that the Demkov VP, for scattering when target wave functions associated with open channels are not precisely known, may not follow from our routine procedure. Indeed, this is the only problem we have examined for which our procedure does not almost automatically lead to a VP. The difficulties arise because, in attempting a derivation of the VP following our routine procedure, divergent integrals are encountered. The possibility that our routine procedure will lead to formally divergent expressions was not taken into account in our description of our procedure; when such divergences occur, they may have to be dealt with—if at all—by exercises of ingenuity. Fortunately problems involving such divergences have been encountered only rarely—in our experience, at least; otherwise, the whole point of the approach we have developed and have presented in this review would be largely negated.

To illustrate the preceding remarks, let us return to Eq. (5.28), the starting point for the Kohn VP pertinent to collisions involving composite systems. In deriving the VP, it is necessary as we have many times explained, to express $L^\dagger[(H-E)\delta\phi_i]$ in terms of $[(H-E)L]^\dagger \delta\phi_i$. The difference between these two expressions involves a surface integral at infinity which, with proper boundary conditions at infinity, will vanish when the target wave functions and corresponding bound-state energies are precisely known. When these energies are not precisely known, however, assigning the incident kinetic energy ε' of Eq. (6.1) does not specify the outgoing energy in any channel other than the incident, not even in channels corresponding to simple excitation without rearrangement. As a result, for any actual choice of $\delta\phi_i$, the surface integral generally will contain terms which are oscillatory as $r \to \infty$, i.e., which are not well defined and whose vanishing cannot strictly be presumed.

Correspondingly, for any actual choice of $\delta\phi_i$, each of the integrals $L^\dagger[(H-E)\delta\phi_i]$ and $[(H-E)L]^\dagger\delta\phi_i$ generally will diverge for any given total energy E, whether this energy E is the exact energy of the system defined by Eq. (6.2), or is taken to be some other trial value E_t. Even when we ignore the relation $(H-E)\phi_i = 0$, there is no problem with the convergence of $L^\dagger[(H-E)\phi_i]$, of course, because this integral involves only exact (and therefore asymptotically properly behaved) quantities; also, as the remarks beneath Eq. (6.5) have illustrated, one can choose an E_t so that the integral $L_t^\dagger[(H-E_t)\phi_{it}]$ converges for reasonable L_t and ϕ_{it}, even in the present case that the target bound-state wave functions and energies are not precisely known. Thus the difference between the pair of now individually convergent integrals $L_t^\dagger[(H-E_t)\phi_{it}]$ and $L^\dagger[(H-E)\phi_i]$ can be made to converge. However, if we write the difference between them as

$$L^\dagger[(H-E)\delta\phi_i] + \delta L^\dagger[(H-E)\delta\phi_i] - \delta E (L+\delta L)^\dagger (\phi_i + \delta\phi_i) , \tag{6.8}$$

the individual terms, which include pieces that are formally of first, second, and higher order, need not converge. If one term diverges in (6.8), then at least one other term diverges. Labeling terms as being of given order is no longer very meaningful; one cannot throw away a divergent term simply because it is formally of second order.

If such convergence difficulties are ignored, our routine procedures can yield VP's of the Demkov type; in other words, when expressions appearing in our routine procedure are formally divergent, our methods nonetheless may have heuristic value, producing formally stationary and convergent VP's by manipulations requiring little or no ingenuity, but still needing accessory rigorous demonstrations of convergence and stationarity. An illustration of this remark is the following "derivation" of the Demkov VP Eq. (6.7) via the variational identity approach described in Sec. V.D.

Let ϕ_l and ϕ_{ltt} have the asymptotic forms introduced in Eqs. (6.3) and (6.6), namely,

$$\begin{aligned} \phi_l(r) &\sim [\sin(kr - \tfrac{1}{2}l\pi) + \tan\eta \cos(kr - \tfrac{1}{2}l\pi)] u , \\ \phi_{ltt}(r) &\sim [\sin(kr - \tfrac{1}{2}l\pi) + \tan\eta_t \cos(kr - \tfrac{1}{2}l\pi)] u_t , \end{aligned} \tag{6.9}$$

when $r \sim \infty$ and where the trial u_t is normalized. Then it is readily verified, using Green's theorem, that

$$\frac{-\hbar^2 k}{2\mu} u^\dagger u_t (\tan\eta - \tan\eta_t) = \phi_l^\dagger [(H-E)\phi_{ltt}] - [(H-E)\phi_l]^\dagger \phi_{ltt} . \tag{6.10}$$

The last term on the right-hand side of Eq. (6.10) van-

ishes, because ϕ_l is an exact solution of the Schrödinger equation. Dropping that term and then adding the resultant Eq. (6.10) to its adjoint yields

$$\frac{-\hbar^2 k}{2\mu}(u^\dagger u_t + u_t^\dagger u)(\tan\eta - \tan\eta_t) = \phi_l^\dagger[(H-E)\phi_{ltt}] + [(H-E)\phi_{ltt}]^\dagger \phi_l \ . \tag{6.11}$$

We are here seeking a suggestive guide to a VP rather than a rigorous derivation. Since $\phi_{ltt} = \phi_l + \delta\phi_l$ differs from ϕ_l by first order, and since ϕ_l satisfies the Schrödinger equation, replacing ϕ_l by ϕ_{ltt} in Eq. (6.11) makes an error of second order, just as explained in going from Eq. (5.34) to Eq. (5.35). Replacing E defined by Eq. (6.2) with E_t defined by

$$E_t = \varepsilon_t + \varepsilon' \equiv u_t^\dagger (H_A + H_B) u_t + \varepsilon'$$

also makes an error of only second order on the right-hand side of (6.11). Moreover,

$$u^\dagger u_t + u_t^\dagger u = u^\dagger(u+\delta u) + (u+\delta u)^\dagger u = 2 + u^\dagger \delta u + \delta u^\dagger u \ ,$$

which, if we recall Eq. (1.37b), differs from 2 by terms of second order. We conclude therefore that to terms of second order

$$\tan\eta = \tan\eta_t - \frac{2\mu}{\hbar^2 k}\frac{1}{2}\{\phi_{ltt}^\dagger[(H-E_t)\phi_{ltt}] + [(H-E_t)\phi_{ltt}]^\dagger \phi_{ltt}\} \ . \tag{6.12a}$$

Each of the integrals on the right-hand side of Eq. (6.12a) is convergent, so that—granting the validity of our manipulations thus far—we have demonstrated that the right-hand side of Eq. (6.12a) is a variational estimate $\langle \tan\eta \rangle_v$ of the exact $\tan\eta$. Furthermore, application of Green's theorem also gives

$$\phi_{ltt}^\dagger[(H-E_t)\phi_{ltt}] - [(H-E_t)\phi_{ltt}]^\dagger \phi_{ltt} = 0 \ . \tag{6.12b}$$

Using Eq. (6.12b) in (6.12a) immediately yields the Demkov VP Eq. (6.7).

C. Zero-incident kinetic energy

The use of the Demkov principle requires the inversion of $H-E$, a problem discussed towards the end of Sec. V.B in connection with the Kohn principle. However, whereas the Kohn principle does not require the inversion of a singular operator at $\varepsilon'=0$, the Demkov principle does. Thus, restricting ourselves henceforth to $l=0$ and to $\varepsilon'=0$, and assuming there is then only one open channel for $l=0$, we have [as in Eqs. (5.8c) and (5.8d)]

$$\langle A \rangle_{\rm var} = A_t + (2m/\hbar^2)\bar{\phi}_{0t}^\dagger (H-\varepsilon)\bar{\phi}_{0t} \ , \tag{6.13a}$$

$$\bar{\phi}_{0t} \sim (r - A_t)u, \quad r \sim \infty \tag{6.13b}$$

for the Kohn principle, with u and ε known; we also have

$$\langle A \rangle_{\rm var} = A_t + (2m/\hbar^2)\bar{\phi}_{0tt}^\dagger (H-\varepsilon_t)\bar{\phi}_{0tt} \ , \tag{6.14a}$$

$$\bar{\phi}_{0tt} \sim (r - A_t)u_t, \quad r \sim \infty \ , \tag{6.14b}$$

for the Demkov principle, with u and ε only imprecisely known, and with $\varepsilon_t \equiv u_t^\dagger H_0 u_t > \varepsilon$. In line with our previous conventions [see Eq. (5.8c)], the bar in $\bar{\phi}_{0t}$ and $\bar{\phi}_{0tt}$ denotes a zero-incident kinetic energy function, the subscript zero on $\bar{\phi}$ denotes zero angular momentum, the subscript t on $\bar{\phi}$ indicates that the scattering wave function is not known, and the subscript tt on $\bar{\phi}$ indicates that both the bound-state wave function and the scattering wave function are not known.

In practice, the $\bar{\phi}_{0tt}$ used in the VP (6.14a) is obtained by making a linear expansion in a suitably selected subset of some complete set of functions (the dependence on r being adjusted to guarantee the correct asymptotic behavior at infinity, of course); the undetermined coefficients in this expansion are chosen so as to make the right side of (6.14a) stationary for variations of these coefficients about their chosen values. In this procedure, the equations determining the chosen "best" values of the undetermined coefficients are linear, and their solution involves inversion of the matrix $(H-\varepsilon_t)$ in the representation defined by the aforementioned subset. However, while $(H-\varepsilon)^{-1}$ is not singular, since ε is at the (bottom) edge of the spectrum of H—see the end of Sec. V.B—$(H-\varepsilon_t)^{-1}$ *is* singular, since ε_t (which is greater than ε) is embedded in the continuum of H.

Though singular operators in the Kohn VP at nonzero kinetic energies are used regularly, their use can be very troublesome. In particular, in a model-problem analysis of e^+-H scattering, with the ground-state wave function of H assumed to be only imprecisely known, some though not all of the numerical estimates of A based on the Demkov principle and on variants thereof (Peterkop and Rabik, 1971; Houston, 1973; Page, 1975 and 1976) when plotted resembled the schematic figures which appear in texts on the ergodic theorem! Clearly, the inversion of $H-\varepsilon_t$ should be avoided if possible. In order to explain how to do so, we first make a distinction between stationary principles and variational principles, a distinction which often is wholly inconsequential and with which we therefore have not previously been concerned in this article. The Demkov principle for A given by Eq. (6.14) is stationary, that is, the use of $\bar{\phi}_{0tt} = \bar{\phi}_0 + \delta\bar{\phi}_0$ leads to an error $A_v - A = O(\delta\bar{\phi}_0)^2$. On the other hand, it is not always obvious that our choice $\bar{\phi}_{0tt}$ will differ from $\bar{\phi}_0$ by an amount $\delta\bar{\phi}_0$ which can be considered "small"; unless $\delta\bar{\phi}_0$ is small, the error $A_v - A$ of order $(\delta\bar{\phi}_0)^2$ will be large, and there will be no particular advantage to using the variational principle to estimate A. The procedure of varying the undetermined expansion coefficients in the "variational" principle, described in the preceding paragraph, is one way of choosing a $\bar{\phi}_{0tt}$ for which $\delta\bar{\phi}_0$ will be small. If another way of choosing a $\bar{\phi}_{0tt}$ for which $\delta\bar{\phi}_0$ is small can be found, however, a way not involving inverting $(H-\varepsilon_t)$, use of this $\bar{\phi}_{0tt}$ in the Demkov principle

(6.14) without varying any parameters still will give a good estimate of A [error of order $(\delta\bar{\phi}_0)^2$] because of the stationarity property possessed by the VP.

One possibility of avoiding the $(H-\varepsilon_t)$ inversion problem has been suggested by Blau *et al.* (1977), and studied numerically by Aronson *et al.* (1979). Setting $l=0$ and $k=0$ in Eq. (5.8a) and using Eqs. (5.8b) and (5.8c), and accounting for the fact that we are here concerned with scattering by a compound system, we find

$$A=(2m/\hbar^2)(ru)^\dagger V\bar{\phi}_0 \ , \tag{6.15a}$$

where

$$\bar{\phi}_0\sim(r-A)u \ , \tag{6.15b}$$

$k\to 0$ at fixed large r. If we write

$$\bar{\phi}_0=ru-G(\varepsilon)Vru \ , \tag{6.16}$$

where

$$G(\varepsilon)=(H-\varepsilon)^{-1} \ , \tag{6.17}$$

and where there is no need to replace ε in (6.17) by $\varepsilon+i\eta$ with $\eta>0$ and $\eta\to 0$, because we are at the edge of the spectrum, Eq. (6.15a) becomes

$$A=(2m/\hbar^2)(ru)^\dagger[V-VG(\varepsilon)V]ru \ . \tag{6.18}$$

We now introduce an approximation $\tilde{A}$ to A, defined by

$$\tilde{A}\equiv(2m/\hbar^2)(ru_t)^\dagger[V-VG(\underline{\varepsilon})V]ru_t \ , \tag{6.19}$$

where $\underline{\varepsilon}$ is a very good *lower* bound on ε which is independent of the choice of u_t, and where

$$G(\underline{\varepsilon})\equiv(H-\underline{\varepsilon})^{-1}$$

($\underline{\varepsilon}$ will normally be put in by hand, using experimental information on the possible range of values of ε, though for simple systems one might calculate a variational lower bound on ε). Since $\underline{\varepsilon}$ does not lie in the spectrum of H, $G(\underline{\varepsilon})$ is not singular, and one can obtain a variational estimate $G_v(\underline{\varepsilon})$ of $G(\underline{\varepsilon})$ by using the VP for the inverse of an operator discussed in Sec. I.B.3. It is in fact trivial to show that the diagonal elements of $G_v(\underline{\varepsilon})$ yield a variational lower *bound* on the diagonal elements of $G(\underline{\varepsilon})$ if there are no composite bound states; this follows from the result at the end of Sec. I.B.3, remembering H is Hermitian, and recognizing that the inequality $\underline{\varepsilon}<\varepsilon$ then guarantees that $G(\underline{\varepsilon})$ is positive definite. Note that whereas the scattered part of the wave function, $\chi\equiv G(\varepsilon)Vru$, is proportional asymptotically to $(r-A)u$, $\chi_t\equiv G(\underline{\varepsilon})Vru_t$ is an exponentially *decaying* function of r; however, the decay constant will be very small, since it is proportional to $(\varepsilon-\underline{\varepsilon})^{1/2}$, and the incorrect asymptotic form imposes no restriction on the accuracy which can be achieved in the calculation. One thereby obtains a variational estimate $\tilde{A}_v$ of $\tilde{A}$,

$$\tilde{A}_v=\langle\tilde{A}\rangle_{\rm var}$$
$$=(2m/\hbar^2)(ru_t)^\dagger[V-VG_v(\underline{\varepsilon})V]ru_t \ . \tag{6.20}$$

In fact $\tilde{A}_v$ yields an upper bound to $\tilde{A}$, for reasons explained at the end of Sec. V.B.3. For u_t, $\underline{\varepsilon}$, and $G_v(\underline{\varepsilon})$, reasonably accurate approximations to u, ε, and $G(\varepsilon)$, respectively, $\tilde{A}_v$ should represent a reasonably accurate estimate not only of $\tilde{A}$ but also of A. Thus far, we have neither needed nor used the Demkov principle, but on the other hand, our estimate $\tilde{A}_v$, although it yields $\tilde{A}$ to second order, differs in *first* order from A. (We replaced u by u_t and ε by $\underline{\varepsilon}$ in going from A and $\tilde{A}$, and it would have taken still another miracle for $A-\tilde{A}$ to be of second order.) However, in the course of computing $\langle\tilde{A}\rangle_{\rm var}$ we have obtained a good first-order estimate of $\bar{\phi}_0$, namely,

$$ru_t-G_v(\underline{\varepsilon})Vru_t \ ,$$

and this estimate can be further improved by dropping the exponentially decaying factor in $G_v(\underline{\varepsilon})Vru_t$. The resultant wave function now can be used as the trial function $\bar{\phi}_{ott}$ in the Demkov stationary principle (6.14a). Since the linear parameters contained in this trial function would not be varied—i.e., since whatever arbitrary parameters originally appearing in u_t and $G_v(\underline{\varepsilon})$ already have been determined in the course of computing the variational estimate $G_v(\underline{\varepsilon})$—we thereby obtain an estimate of A which is stationary without any need to vary parameters, i.e., without any need to invert the operator $H-\varepsilon_t$.

In other words, even without the Demkov principle we can obtain a variational bound $\tilde{A}_v$ on $\tilde{A}$ which simultaneously should give a reasonable estimate of A, not merely of $\tilde{A}$; this calculation of $\tilde{A}_v$ will involve no numerical instabilities, but for a highly accurate estimate of A without the aid of a VP we will need a very good approximate u_t, since $\tilde{A}-A$ (and therefore $\tilde{A}_v-A$) is of order δu. Such a u_t may be available for, say $e^\pm$-He scattering, but not for more complicated collisions; and even for $e^\pm$-He scattering use of such u_t significantly complicates the calculations. Alternatively, we can use a relatively simple but reasonable approximation u_t and obtain a variational bound $\tilde{A}_v$ on $\tilde{A}$, in the course of which we will obtain an approximate scattering wave function. This approximate function can be readily modified—one need merely drop a factor which generated exponential decay—and used as a trial function in the Demkov principle, the resulting estimate of A being in error by an order of $(\delta u)^2$. We have here an approach which combines a nonvariational bound with no numerical instabilities and a stationary principle to give a variational estimate of A.

VII. VARIATIONAL PRINCIPLES FOR TIME-DEPENDENT TRANSITION AMPLITUDES, THE TIME-TRANSLATION OPERATOR, AND THE DENSITY MATRIX

We consider a system described by a Hermitian Hamiltonian H_0, with energy eigenvalues E_n and orthonormal eigenfunctions ϕ_n:

$$(H_0-E_n)\phi_n=0, \quad \phi_n^\dagger\phi_m=\delta_{nm} \ .$$

The system is subject to an external time-dependent Hermitian interaction $V(\tau)$, and the problem is to determine the normalized wave function $\phi(\tau)$, defined by

$$\left[H-i\hbar\frac{\partial}{\partial\tau}\right]\phi(\tau)=0\ ,\qquad H=H(\tau)=H_0+V(\tau)\ , \tag{7.1}$$

given the normalized wave function at the initial time τ_0. We do not display the spatial dependence of ϕ_n, $\phi(\tau_0)$, $\phi(\tau)$, H_0, H, and $V(\tau)$. Equivalently, we want to know the amplitude $a_n(\tau)$ for finding the system in its nth state at the time τ, given all the initial amplitudes $a_m(\tau_0)$.

Equivalent to a knowledge of the $a_n(\tau)$ is a knowledge of the time translation operator $U(\tau,\tau_0)$, defined by

$$\phi(\tau)=U(\tau,\tau_0)\phi(\tau_0)\ . \tag{7.2}$$

It follows from its definition that U is unitary, that $U(\tau,\tau')U(\tau',\tau)=1$, and therefore that

$$U(\tau,\tau')=U^{-1}(\tau',\tau)=U^{\dagger}(\tau',\tau)\ . \tag{7.3a}$$

We clearly have

$$U(\tau_0,\tau_0)=1\ , \tag{7.3b}$$

and, from Eq. (7.1),

$$\left[H(\tau)-i\hbar\frac{\partial}{\partial\tau}\right]U(\tau,\tau_0)=0\ . \tag{7.4}$$

$U(\tau,\tau_0)$ is uniquely defined by the first-order differential equation (7.4) and the boundary conditions of Eqs. (7.3).

VP's for U in the interaction picture, U_I (Lippmann and Schwinger, 1950; Lippmann, 1956), and for the associated amplitudes $C_n(\tau)$ have long been known. We will derive in this section VP's for $U(\tau,\tau_0)$ and $a_n(\tau)$ using our more general approach. [We note that bounds on transition amplitudes for simple two-state systems have been obtained by Aspinall and Percival (1967), using the properties of the Pauli spin matrices. Using an entirely different approach, these results were generalized by Spruch (1969) to transition amplitudes of almost arbitrary systems and were extended to *variational* bounds; the possibility of obtaining variational bounds rests on a property of U that makes it simpler in many respects than the singular Green's function that occurs in the time-dependent scattering formalism, namely, the isometric property that $U^{\dagger}U\leq 1$. The variational bound was used by Shakeshaft and Spruch (1974) to show for certain classes of potentials that the second Born approximation dominates over the first for the transfer of a light particle in processes analogous to charge transfer.]

The fact that we are now concerned with a first-order rather than a second-order differential equation does not alter the basic approach to the determination of a VP. The starting point for a VP for $U(\tau,\tau')$ follows immediately from Eqs. (7.3) and (7.4). Writing ∂_τ for $\partial/\partial\tau$, and inserting a factor of $-i/\hbar$ for later convenience, we have

$$\langle U(\tau,\tau_0)\rangle_{\rm var}=U_t(\tau,\tau_0)-(i/\hbar)\int_{\tau_0}^{\tau}\Lambda_t(\tau,\tau')[H(\tau')-i\hbar\partial_{\tau'}]U_t(\tau',\tau_0)d\tau'\ , \tag{7.5}$$

where the only restriction placed on the trial time translation operator $U_t(\tau,\tau_0)$ is

$$U_t(\tau_0,\tau_0)=1\ .$$

Proceeding as usual to obtain $\Lambda(\tau,\tau')$ we have

$$0=\delta U(\tau,\tau_0)-\int_{\tau_0}^{\tau}\Lambda(\tau,\tau')[(i/\hbar)H(\tau')\delta U(\tau',\tau_0)+\delta\partial_{\tau'}U(\tau',\tau_0)]d\tau'\ ,$$

where the term in $\delta\Lambda$ vanishes because of Eq. (7.4). Writing

$$\delta\partial_{\tau'}U(\tau',\tau_0)=\partial_{\tau'}\delta U(\tau',\tau_0)\ ,$$

integrating by parts, and using $\delta U(\tau_0,\tau_0)=0$, we find

$$-i\hbar\partial_{\tau'}\Lambda(\tau,\tau')=\Lambda(\tau,\tau')H(\tau')\ , \tag{7.6}$$

$$\Lambda(\tau,\tau)=1\ . \tag{7.7}$$

$\Lambda(\tau,\tau')$ can now be identified with $U(\tau,\tau')$; to see this, we take the adjoint of Eq. (7.4) and use Eq. (7.3a). The replacement of $\Lambda_t(\tau,\tau')$by $U_t(\tau,\tau')$ in Eq. (7.5) gives a VP for $U(\tau,\tau_0)$. As in the estimation of energy eigenvalues and scattering parameters, the only relevant operator in the defining equations of the present problem is H. As in those previously examined cases, the adjoint function Λ in the present problem is not some new function; rather, it is a function that appears in the defining equations.

A VP for transition amplitudes instead of one for $U(\tau,\tau')$ could be obtained by inserting

$$\phi(\tau)=\Sigma_n a_n(\tau)\phi_n \tag{7.8}$$

into Eq. (7.1), premultiplying by $\phi_m^{\dagger}$, and introducing the matrix $\mathbf{H}$ with elements

$$H_{mn}(\tau)=\phi_m^{\dagger}H(\tau)\phi_n\ . \tag{7.9}$$

Introducing the column vector $\mathbf{a}(\tau)$ with components $a_m(\tau)$, we have

$$\partial_\tau\mathbf{a}(\tau)=-(i/\hbar)\mathbf{H}(\tau)\mathbf{a}(\tau)\ . \tag{7.10}$$

The quantity $\mathbf{a}(\tau)$ is completely defined by the first-order equation (7.10) and by its initial value $\mathbf{a}(\tau_0)$ (which is presumed to be known). The VP starting from

$$\langle\mathbf{a}(\tau)\rangle_{\rm var}=\mathbf{a}_t(\tau)+\int_{\tau_0}^{\tau}d\tau'\Lambda_t(\tau,\tau')\times[(-i/\hbar)\mathbf{H}(\tau')-\partial_{\tau'}]\mathbf{a}_t(\tau') \tag{7.11}$$

follows immediately. Proceeding as before, we find that $\Lambda(\tau,\tau')$ satisfies the same equations (7.6) and (7.7) and is therefore again equal to $U(\tau,\tau')$. This close relationship of the VP (7.11) for $\mathbf{a}(\tau)$ to the principle (7.5) for $U(\tau,\tau_0)$

is, of course, to be expected. The more familiar VP's for U_I and $\mathbf{a}_I(\tau)$, the analogs in the interaction representation of U and $\mathbf{a}$, can be written down in an exactly similar way.

If we drop the problem of transition amplitudes, and restrict ourselves to a time-independent Hamiltonian H, whose decomposition into H_0+V need not concern us, we have

$$(H-i\hbar\partial_\tau)\phi(\tau)=0 ,$$

$$\phi(\tau)=U(\tau-\tau_0)\phi(\tau_0) ,$$

$U(\tau,\tau_0)$ reducing to a function of the difference of the arguments, since the system is invariant under time translation. We now have

$$(H-i\hbar\partial_\tau)U(\tau)=0, \quad U(0)=1 , \tag{7.12}$$

while the VP (7.5) for $U(\tau)$ reduces, setting $\tau_0=0$ for convenience—time translation invariance allows this—to

$$\langle U(\tau)\rangle_{\rm var}=U_t(\tau)-\int_0^\tau U_t(\tau-\tau')[(i/\hbar)H+\partial_{\tau'}] \times U_t(\tau')d\tau' , \tag{7.13}$$

where $U_t(0)=1$ but $U_t(\tau')$ is otherwise unrestricted.

Consider now the density matrix $\rho(\beta)$, defined by

$$\rho(\beta)=e^{-\beta H} .$$

$\rho(\beta)$ therefore satisfies $(H+\partial_\beta)\rho(\beta)=0$, $\rho(0)=1$, which is equivalent to Eq. (7.12), with τ replaced by $-i\hbar\beta$. A VP for $\rho(\beta)$ follows immediately from Eq. (7.13) and is given by

$$\langle\rho(\beta)\rangle_{\rm var}=\rho_t(\beta)-\int_0^\beta \rho_t(\beta-\beta')[H+\partial_{\beta'}]\rho_t(\beta')d\beta' ,$$

where the only restriction on $\rho_t(\beta)$ is that $\rho_t(0)=1$. For a time-dependent Hamiltonian, a VP for the density matrix $\rho(\beta,\beta')$ follows in an exactly similar way from Eq. (7.5) with the replacement of τ in these equations by $-i\hbar\beta$.

We conclude this section with a comment on the hermiticity of $V(\tau)$ or, alternatively, $H(\tau)$. If these operators are not Hermitian, the mechanics of the construction of the VP's in this section is not affected and equations like (7.6) for the Lagrange multiplier still follow. However, these multipliers are now new functions and can no longer be identified with the original $U(\tau,\tau_0)$ and $U_I(\tau,\tau_0)$. Correspondingly, if $H\neq H^\dagger$ for H the Hamiltonian in an eigenvalue or scattering problem, one obtains equations for L as usual, but L differs from the solution ϕ of the Schrödinger equation. See also Appendix D.

VIII. LITERATURE SURVEY AND GENERAL REMARKS

The derivations which have been given were largely self-contained; when they were not, we cited specific references to the required reading material. We conclude with some references to the literature on VP's in particular areas, intended to serve readers who wish to pursue those areas in greater depth; we also give some references to the literature on the construction of VP's and on applications of VP's, and to the literature on the formulation of the laws of physics as extremum principles; and we make one or two general comments on VP's.

The classical and quantum-mechanical literature is replete with VP's for all sorts of quantities. In classical physics, these include, among others, frequencies of vibration, first considered at least as far back as 1871 (Rayleigh, 1871, 1873, and 1877, see especially pp. 177 and 178 in Vol. I of the 1945 reprinted edition), wave-guide propagation constants (Schwinger and Saxon, 1968; Cairo and Kahan, 1965; Kalikstein, 1981), and parameters in neutron transport theory (Marshak, 1947; Kostin and Brooks, 1964) and in nuclear reactor physics (Usachev, 1964; Gandini, 1967; Pomraning, 1967b and 1968; Stacey, 1972; Conn and Stacey, 1973; Nelson, 1975; Cheng, 1980). In quantum mechanics, there exist the Rayleigh-Ritz principle (Rayleigh, 1871 and 1873; Ritz, 1908 and 1909) for the ground-state energy, the Hylleraas-Undheim (1930) principle extending the Rayleigh-Ritz principle to excited states,[6] and a host of principles in scattering theory (Schwinger and Saxon, 1968; Hulthén, 1944 and 1948; Kohn, 1948; Lippmann and Schwinger, 1950). In fact, there exist a number of studies devoted almost entirely (Spruch, 1962; Demkov, 1963) or partly (Mott and Massey, 1965; Newton, 1982; Moiseiwitsch, 1966) to VP's in scattering theory. Callaway (1978) and Nesbet (1980) are two recent reviews focusing on the *use* of VP's in atomic scattering. Besides the papers referenced in these reviews, recent work on VP's in atomic scattering includes Maleki and Macek (1980), Watson *et al.* (1980), Lucchese and McKoy (1980), and Maleki (1981); these authors use the Schwinger VP, discussed in Sec. VI, which contains a Green's function.

What the quantum-mechanical VP's mentioned above share is that both energies and scattering parameters can be expressed as matrix elements of H. However, the quantum-mechanical literature also contains a few more general VP's for matrix elements of an arbitrary operator W. These include diagonal matrix elements (Dalgarno and Lewis, 1956; Dalgarno and Stewart, 1956 and 1960; Schwartz, 1959a and 1959b; Delves, 1963a) and off-diagonal matrix elements between two discrete (quadratically integrable) states (Delves, 1963a), or two continuum (nonquadratically integrable) states (Delves, 1963b), as well as between one bound state and one continuum state. There also are VP's for integrals over functions ϕ satisfy-

[6] It seems to be traditional among physicists to refer to Hylleraas and Undheim (1930) or MacDonald (1933) for extension of the energy principle to excited states. However, the result was already contained in the work of Rayleigh and Ritz. It is also interesting that Rayleigh himself responded (Rayleigh, 1911) to the work of Ritz, expressing surprise that Ritz should have regarded the method itself as new. For a more extended account of these historical remarks, see a bibliographical note on p. 67 of Gould (1957).

ing not only homogeneous equations, as in Eq. (1.3a), but also inhomogeneous equations, as well as VP's for the solution of integral equations (Pomraning, 1968; Strieder and Prager, 1967; Robinson and Arthurs, 1968). It is stressed that the above references are merely illustrative and do not begin to encompass the entire vast literature on variational methods. Other books on VP's in quantum mechanics and in classical physics include Gould (1957), Funk (1962), Michlin (1964), Mercier (1963), Becker (1964), Lewins (1965), Donnelly *et al.* (1965), Lanczos (1966), Petrov (1968), Yourgrau and Mandelstam (1968), Biot (1970), Finlayson (1972), and Rektorys (1975). There also is considerable material on VP's and variational estimates in Morse and Feshbach (1953), especially Sec. 9.4. In addition, the relationship between perturbation expansions and VP's has received considerable study, especially in quantum mechanics, for time-dependent as well as time-independent quantities (Hirschfelder *et al.*, 1964, Heinrichs, 1968a and 1968b; Langhoff *et al.*, 1972, especially Sec. V and footnotes 7 and 8; Löwdin and Mukherjee, 1972). Other papers of interest in these and related areas, in which theoretical relations such as hypervirial theorems and experimental data are built into the trial function, commented on briefly in Sec. III.C.4, include Kirtman (1971), Kirtman and Mowery (1971), Heaton and Moiseiwitsch (1971), Sims and Rumble (1973), and McWhirter and Moiseiwitsch (1972).

In general, however, the VP's for $F(\phi)$ encountered in the literature have been obtained by manipulations especially suited to the particular $F(\phi)$ whose variational estimate is sought. It does not seem to be generally recognized that variational estimates of a very wide class of functionals $F(\phi)$ can be systematically constructed, as apparently was first pointed out by Borowitz and Gerjuoy (1965). They give a number of applications of the procedure to problems in different areas, including some in scattering theory, a relatively difficult case because of the comparatively difficult boundary conditions. The basic idea is to incorporate the constraints on ϕ into the VP via Lagrange undetermined multipliers, in a fashion very like the method used when obtaining the extremum of a functional subject to constraints on the argument ϕ of the functional. The same basic idea was arrived at independently by Haymaker and Blankenbecler (1968), by Rau and Spruch, and doubtless by other investigators as well. With differing aims and emphasis, the idea of using Lagrange multipliers as adjoint variables seems to have occurred in several places. In considering the so-called "inverse problem in the calculus of variations," namely, to write down a Lagrangian from which will follow as Euler-Lagrange equations a given set of equations of motion, Atherton and Homsy (1975) show how this can be done, building on powerful theorems established by Vainberg (1964). Complementary VP's leading to bounds on both sides, the so-called "bivariational bounds," are developed in Arthurs (1970), Robinson (1971), Barnsley and Robinson (1974 and 1977), Noble and Sewell (1972), and Sewell and Noble (1978). Examples of recent papers using such complementary principles are Collins (1977) for the heat equation and Cole (1979) for free molecular channel flow where integral equations are involved.

Closer in spirit to the question we investigate—namely, the evaluation of an $F(\phi)$ given defining equations $B(\phi)=0$, is the work in nuclear reactor physics. Starting from the work of Roussopoulos (1953), Pomraning (1967a) gave a general formulation, followed by applications to neutron transport in Pomraning (1967b and 1968). Later work along these lines in reactor physics is in Stacey (1972), Conn and Stacey (1973), and Nelson (1975); a recent paper in the field is Cheng (1980). Parallel work, also in connection with neutron transport problems, with adjoint functions for both linear and nonlinear problems (called the "importance" function in these papers), is in Becker (1964) and Lewins (1965).

These works still fall somewhat short of our goal of a systematic examination of the applicability of the basic procedure to wide classes of problems with a particular focus on problems in quantum mechanics. It is here that one sees most readily that subtleties regarding the proper definition of the functions involved can be crucial for success in formulating a VP. All the work mentioned above and the material in Finlayson (1972) have significant overlap with the work of our school which forms the central focus of this review. The main differences we see are in our concentration on quantum mechanics, in our development, side by side, of so-called variational identities, and in our emphasis on proper attention to pitfalls which can derail the formulation of a VP.

In this review, we discussed in some detail the extent to which the function ϕ must be specified, and examined the various ways in which the associated constraints can be incorporated into the VP. The main papers in which our formulation was developed are Gerjuoy, Rau, and Spruch (1972,1973), Gerjuoy, Rosenberg, and Spruch (1975), Gerjuoy (1974), and Gerjuoy, Rau *et al.* (1975). Applications of the procedure in quantum mechanics are in Mueller *et al.*, (1974), Shakeshaft *et al.* (1976) and Wadehra *et al.* (1978). Rau (1974) considers a problem in astrophysics. The development of VP's for nonlinear differential equations is considered in Rau (1976) and Inokuti *et al.* (1978), for nonlinear integral equations in Rosenberg and Tolchin (1973) and Rau and Rajagopal (1975 and 1976), and for integral difference equations, both linear and nonlinear, in Rau *et al.* (1978) and Inokuti *et al.* (1980).

The utility, or lack of utility, of the VP's we discussed for actual computation of any desired $F(\phi)$ was not more than superficially discussed, in Sec. IV.C. In fact, difficulties do arise in actual applications, as is well known, for example, in applications (Harris and Michels, 1971; Nesbet and Oberoi, 1972; Bardsley *et al.*, 1972; Callaway, 1978; Nesbet, 1980) of the Kohn principle (Kohn, 1948) and variants thereon (Hulthén, 1944 and 1948; Rubinow, 1955). Nor did we discuss the possible uses of these VP's for the purpose of obtaining upper and/or lower variational bounds on desired quantities, a subject on which again there is a considerable literature (Spruch and Rosenberg, 1959 and 1960; Spruch, 1969; Delves, 1963c; Weinhold, 1972; Blau *et al.*, 1973, 1974, 1975, and 1977), some

of which is referenced in recent papers (Gerjuoy *et al.*, 1972; Spruch, 1975). In Gerjuoy *et al.* (1972) and Gerjuoy, Rau *et al.* (1975) it is shown that our procedures routinely yield not merely a VP for $F(\phi)$, but also a generalization of the well-known identity (Kato, 1951) for the tangent of the phase shift. Such a variational identity, in which one expresses the quantity of interest as the sum of a closely related VP for that quantity plus a formally known second-order term, has been one of the important starting points for determining error bounds on the variational estimates of scattering parameters (Spruch and Rosenberg, 1959 and 1960; Rosenberg *et al.*, 1960; Hahn *et al.*, 1962, 1963, and 1964; Sugar and Blankenbecler, 1964), of transition amplitudes (Spruch, 1969), and of functions such as temperature which characterize the equalization processes (Kalikstein and Spruch, unpublished).

Returning to the literature on variational methods, we remark that for the most part the canonical literature has not been concerned with the problem of main interest in the present work, namely, the problem of constructing a VP $\langle F(\phi)\rangle_{\rm var}$ for a given desired $F(\phi)$, thus making it possible to estimate $F(\phi)$ with second-order accuracy using trial functions ϕ_t accurate merely to first order. Instead, the literature has largely been concerned with the construction of stationary principles as an alternative and very compact formulation of the laws of physics or of a given set of equations. Such a usage goes back to the very foundations of physics. Hamilton's principle and the principle of least action are examples in classical mechanics.[7] Now, in theory, any stationary expression which offers an alternative formulation of the laws of physics and thereby of the relations defining ϕ provides a way of computing ϕ, but this is often of little use in practice. For instance, the classical mechanics literature does not record systematic attempts to find the position of a particle as a function of time variationally from Hamilton's principle [for some efforts of this sort, see Miller (1957), Luttinger and Thomas (1960), and Lyness and Blatt (1961 and 1962)], probably because Hamilton's principle involves an integration over time from initial instant t_0 to final instant t_1, where t_1 is not known until the dynamical paths carrying the particles from their initial to their final configurations are known. On the other hand, in heat transfer problems various stationary functionals apparently furnish reasonably accurate estimates of the unknown temperature distribution function ϕ determining the heat flow (Biot, 1970), and as a second example, employment of the Rayleigh-Ritz VP is one of the most practical ways to compute the quantum-mechanical ground-state eigenfunctions for complicated atoms, molecules, or nuclei. Nevertheless, the estimates of ϕ obtained in this way are accurate only to first order, because they make some particular functional of ϕ stationary, not normally ϕ itself. Correspondingly, these estimates of ϕ furnish only a first-order estimate of any desired functional $F(\phi)$ other than the particular functional (such as the energy) which was made stationary. Presumably the "best" VP for estimating ϕ is the one which makes ϕ itself stationary, that is, which will yield ϕ to second order using trial estimates ϕ_t accurate only to first order; VP's for ϕ itself were constructed by the methods detailed above in Sec. IV.A. (See also Appendix C.) With such VP's, it is possible to estimate any desired $F(\phi)$ with second-order accuracy by simply using the estimate of ϕ accurate to second order. Further, constructing a VP $\langle\phi\rangle_{\rm var}$ for ϕ itself opens the way to the construction of "supervariational principles" for $F(\phi)$ that are good to higher order than the second (Blankenbecler, 1966; Gerjuoy, Rau *et al.*, 1975).

In passing, we note that there is a third use of VP's and of variational bounds, namely, the determination of the functional *form* of an expression, such as the velocity dependence of a cross section. Applications include effective range theory (Schwinger, 1947), which defines the low energy form of a cross section, and the high-velocity limit of charge exchange and similar cross sections (Shakeshaft and Spruch, 1974).

We make two additional remarks bearing on the practical role of this paper's general procedures for constructing VP's. For particular classes of problems—for example, for the VP's associated with scattering processes discussed in Sec. V, systematic approaches other than the very broadly applicable procedures we describe may exist. An example was given in Sec. V.D. The second remark is that, given a VP for a given quantity, the most economical way of obtaining a second VP for the same quantity or related quantities often may not be the direct approach, but rather the choice of a particular form of trial function in the VP already known. Thus, in the context of Sec. V, the insertion of a trial function ϕ_t into the Schwinger VP, with ϕ_t having a specified dependence on a second function ϕ_{tt}, yields the Kohn VP (see, for example Lieber *et al.*, 1972a, 1972b) in terms of the trial function ϕ_{tt}; similarly, the Kohn VP can be made to yield the Schwinger principle (Kohn, 1948; Carew and Rosenberg, 1973). Another example was given in Sec. V.A.

We also note that it may be possible to derive VP's from the Rayleigh-Ritz principle. A starting point might be papers by Risberg (1956), Percival (1957, 1960), and Rosenberg and Spruch (1960), in which variational bounds for potential scattering are derived using the Rayleigh-Ritz method by converting a difficult portion of the scattering analysis to the analysis of a bound-state problem. VP's in the R-matrix formalism also belong in this class (Chatwin and Purcell, 1971; Fano and Lee, 1973; Lee, 1974; Burke and Robb, 1975; Shimamura, 1978).

[7]The viewpoint dates back at least as far as Hero of Alexandria some 2000 years ago. He recognized that the path of a ray of light in traveling between two points via reflection from a plane mirror is the path of minimum length. The concept that light knew in advance along which path to proceed was at one time ascribed teleological significance. See also the wonderful discussion in Chap. 19 of Feynman *et al.* (1964).

ACKNOWLEDGMENTS

The support required for our efforts over the years was provided largely by the Army Research Office, the Office of Naval Research and the National Science Foundation; it is our pleasure to acknowledge that support, presently supplied by the ONR under Contract No. N00014-76-C-0317 and by the NSF Grant Nos. PHY79-10413 and PHY76-05721.

Present-day academic-year faculty life can be more complicated than the lives of the Kwakiutl Indians as described by Franz Boas and Ruth Benedict, and more time consuming than the life of a medieval serf. The opportunity to think unhurriedly often occurs only during summer months, and we thank the Aspen Center for Physics for the many summers during which we were provided stimulating and lovely surroundings. One of us (L.S.) would also like to thank the Institute for Advanced Study, in Princeton, where he was in residence during the very final stages of his share of the work on this review.

Finally, we come to the *sine qua non,* collaborators, who come in the form of lowly graduate students, medium-low postdoctoral research associates, and distinguished colleagues and visitors. Much of the work which culminated in this review was performed on New York University and University of Pittsburgh premises, and a number of people were involved. It is *pro forma* for students to acknowledge their faculty advisors in their theses. However, faculty members do not always have the opportunity to acknowledge their indebtedness to graduate students for their nearly endless efforts, often (and properly) based on ideas suggested by but only dimly understood by their advisors. A number of these extreme efforts bear fruit only in weeding out the unpromising approaches. Graduate students, some of whom continued on in a postdoctoral capacity, and all of whom contributed in the general area of variational principles, include Irving Aronson, Ralph Bartram, Robert Blau, David Epstein, Paul Henry, Kalman Kalikstein, Martin J. Kelly, Chemia J. Kleinman, Ronald O. Mueller, Thomas F. O'Malley, Brian Thomas, and Jogindra M. Wadehra. Michael J. Lieber came as a postdoctoral associate, as did Yukap Hahn and Robin Shakeshaft, who also returned later as visitors. To all of them our many thanks. Professor Kalikstein deserves a second thank you for a careful reading of the manuscript. One of us (A.R.P.R.) acknowledges several conversations with Ugo Fano and Mitio Inokuti; in trying to explain our formalism and point of view, he found that his own ideas were clarified by the questions they raised. Finally, we accord a bouquet of thanks to Leonard Rosenberg, for countless discussions over the years, as enlightening as they were pleasant.

APPENDIX A: NOTATION FOR GENERALIZED LAGRANGE MULTIPLIERS

There are considerable notational difficulties in a general formulation of the Lagrange multipliers $\mathscr{L}$, though these difficulties are rarely troublesome in facing a concrete problem. Some comments on notation may nevertheless be helpful. The illustrative examples we have considered have shown that the generalized Lagrange multipliers $\mathscr{L}$ occurring in VP's may be of quite different types. Even in the comparatively simple problems treated in Sec. I.B we have seen cases in which the quantities $\mathscr{L}$ were ordinary numerical r-independent multipliers (Secs. I.B.1 and I.B.2), operators (Sec. I.B.3), ordinary functions (Secs. I.B.4 and I.B.5) and a vector function of two variables (Sec. I.B.7). Unfortunately, it appears impossible to devise a simple notation which distinguishes between all the diverse possibilities for $\mathscr{L}$. The notation we have employed exposes the spatial dependence of the particular $\mathscr{L}$ under consideration, while simultaneously continuing to emphasize the identification of the adjoint functions $\mathscr{L}$ with generalized Lagrange multipliers. Keeping the spatial dependence of the various $\mathscr{L}$ clearly in mind usually facilitates the derivations of the relations determining the exact $\mathscr{L}$.

Specifically, when an $\mathscr{L}$ is $\mathbf{r}$ independent, it is denoted by the greek lower case λ. When an $\mathscr{L}$ is an ordinary function of $\mathbf{r}$, it is denoted by $L(\mathbf{r})$. When an $\mathscr{L}$ depends on two coordinates, it is denoted by $\Lambda(\mathbf{r},\mathbf{r}')$; the Lagrange multiplier Λ introduced in Sec. I.B.3 presumably would be so indexed in a coordinate representation. For instance, [recalling Eq. (1.18b)] the operator Λ would be the Green's function

$$\Lambda = -\frac{1}{H-E} \equiv -G(E)$$

or, in coordinate representation,

$$\Lambda(\mathbf{r},\mathbf{r}') \equiv -G(\mathbf{r},\mathbf{r}';E) ,$$

when A of Sec. I.B.3 is the operator $(H-E)$, with H the quantum-mechanical Hamiltonian.

The symbols λ, L, and Λ suffice to describe the dependences on $\mathbf{r}$ of all $\mathscr{L}$ encountered in this work; more complicated dependences of $\mathscr{L}$ on $\mathbf{r}$ can occur, of course, and can be dealt with by the methods we describe. [A formal statement is that $\mathscr{L}$ is a mapping of $B(\phi)$ onto $F(\phi)$ and has the characteristic structure of this mapping.] We will now illustrate circumstances under which these various types of adjoint functions can occur; a few arbitrarily chosen illustrative examples are detailed in Table I below, reference to which may help the reader to follow the discussion in the next two paragraphs.

The adjoint function $\mathscr{L}$ will be an $\mathbf{r}$-independent λ when $F(\phi^\dagger,\phi)$ is an $\mathbf{r}$-independent number (for example, the matrix element $\phi^\dagger W\phi$) and when the corresponding constraint $B(\phi^\dagger,\phi)=0$ imposes an $\mathbf{r}$-independent condition [as in the normalization condition (3.1)]. When the constraint is over a range of values of $\mathbf{r}$—an example being the Schrödinger equation (1.3a)—but F still is an $\mathbf{r}$-independent number, the corresponding adjoint function $\mathscr{L}$ will be an ordinary function of $\mathbf{r}$, that is, $L(\mathbf{r})$; also in this case $\mathscr{L}^\dagger B$ or $B^\dagger\mathscr{L}$ will always involve an integral over $\mathbf{r}$. When the constraint is $\mathbf{r}$ independent but F is a function of $\mathbf{r}$, the associated $\mathscr{L}$ again will be a function

TABLE I. Notation for Lagrange multipliers. F_n and B_n denote single numbers. $F(\mathbf{r})$ and $B(\mathbf{r})$ are functions. F_j and B_j denote multicomponented collections of numbers, and $F_j(\mathbf{r})$ and $B_j(\mathbf{r})$ denote multicomponented collections of functions.

F	B	$\mathscr{L}$	Integration variable in $\mathscr{L}^\dagger B$	Index summed in $\mathscr{L}^\dagger B$
F_n	B_n	λ		
F_n	$B(\mathbf{r})$	$L(\mathbf{r})$	$\mathbf{r}$	
$F(\mathbf{r})$	B_n	$L(\mathbf{r})$		
$F(\mathbf{r})$	$B(\mathbf{r}')$	$\Lambda(\mathbf{r},\mathbf{r}')$	$\mathbf{r}'$	
F_n	B_j	λ_j		j
F_n	$B_j(\mathbf{r})$	$L_j(\mathbf{r})$	$\mathbf{r}$	j
F_j	B_k	λ_{jk}		k
$F_j(\mathbf{r})$	$B_k(\mathbf{r}')$	$\Lambda_{jk}(\mathbf{r},\mathbf{r}')$	$\mathbf{r}'$	k

of $\mathbf{r}$—that is, in this case again $\mathscr{L}=L(\mathbf{r})$; now, however, neither $\mathscr{L}^\dagger B$ nor $B^\dagger\mathscr{L}$ will involve an integral over $\mathbf{r}$. An example occurs in Eq. (4.1). Usually it will be obvious whether or not the symbolic products $\mathscr{L}^\dagger B$ and $B^\dagger\mathscr{L}$ involve integration. The case $\mathscr{L}\equiv\Lambda(\mathbf{r},\mathbf{r}')$ occurs when F is a function of $\mathbf{r}$ [the first variable in $\Lambda(\mathbf{r},\mathbf{r}')$] and the constraint is over a range of values; now both $\Lambda^\dagger B$ and $B^\dagger\Lambda$ will be functions of $\mathbf{r}$ and will involve integration over $\mathbf{r}'$. See Eq. (4.1). Note that in this last case, where there is integration over $\mathbf{r}'$ but not $\mathbf{r}$, the notation implies that the adjoint operation on $\Lambda^\dagger B$ does not involve interchange of $\mathbf{r}$ and $\mathbf{r}'$ in Λ, that is, if sums over components are unnecessary (as in quantum-mechanical problems with spinless particles)

$$\Lambda^\dagger B=\int d\mathbf{r}'\Lambda^*(\mathbf{r},\mathbf{r}')B(\mathbf{r}')\ ,$$
$$(\Lambda^\dagger B)^\dagger=B^\dagger\Lambda=\int d\mathbf{r}'B^*(\mathbf{r}')\Lambda(\mathbf{r},\mathbf{r}')\ . \tag{A1}$$

The point is that for the purposes of multiplication by B the $\mathbf{r}$ in $\Lambda(\mathbf{r},\mathbf{r}')$ is to be thought of as a parameter rather than a variable. In many circumstances, an adjoint function of the type $\Lambda(\mathbf{r},\mathbf{r}')$ will be identifiable as a Green's function relevant to the problem under consideration, or will be closely related to such a Green's function.

In general, all the Lagrange multiplier types λ, L and Λ may be multicomponented, that is, may need to be indexed by discrete subscripts, as, for example, for quantum-mechanical problems involving particles possessing spin. The necessary discrete indexings of the $\mathscr{L}$ are inferred just as the necessary $\mathbf{r}$ dependences were inferred, from the given discrete indexings attached to $F(\phi^\dagger,\phi)$ and the particular $B(\phi^\dagger,\phi)$ of interest.

Generally, each of the λ, L and Λ may appear in the fundamental starting point of the VP (2.3) and its analogs as a formal multiplier to the left or to the right of its associated (possibly adjointed) constraint. In this event, it often is convenient to subscript these generalized Lagrange multipliers by $\underline{a}$ (left occurring in the starting expression) or $\underline{b}$ (right occurring in the starting expression); we have generally followed this practice. The main value of the $\underline{a},\underline{b}$ subscript notation is that the $\mathscr{L}_{ai},\mathscr{L}_{bi}$ associated, respectively, with the original and adjointed versions of the ith constraint frequently turn out to be simply related to each other, as, for example, in Eqs. (3.7) and (3.9). Of course, maintaining the distinction between the terms involving $\mathscr{L}_a$ and $\mathscr{L}_b$ is superfluous in some instances, as, for example, with the normalization constraint (3.1); in Eqs. (3.12) and (4.1) we required only one Lagrange multiplier λ associated with such a normalization constraint. Similarly, the distinction between $\mathscr{L}_{ai}$ and $\mathscr{L}_{bi}$ is an unnecessary complication when it is obvious *a priori* that $\mathscr{L}_{ai}$ and $\mathscr{L}_{bi}$ should be identical [recall the discussion connected with Eqs. (3.7) and (3.9)]. For this reason, in the variational starting expressions of Sec. IV.B and later sections we have employed $\mathscr{L}_i$ for $\mathscr{L}_{ai}=\mathscr{L}_{bi}$ whenever possible, once it had become clear (via the simple illustrative problems in Secs. III.B.1 and IV.A) that the equations would indeed imply $\mathscr{L}_{ai}=\mathscr{L}_{bi}$.

APPENDIX B: VARIATIONAL BOUNDS ON POWER DISSIPATION

This appendix, which supplements Secs. I.B.1 and III.C.5, is a digression from the focus of the paper, for we will be concerned with the relative merits of different VP's, any two of which necessarily differ only to second order. More precisely, we will be concerned with the possibility of obtaining not simply VP's but variational bounds. (The distinction between principles and bounds is our only interest here; the problem of two resistors in parallel is a totally trivial one, and we make no effort to recast our results in a form which is applicable to many resistors.) In Secs. I.B.1 and III.C.5 we saw that depending on whether we do or do not require the two trial currents to satisfy the current conserving constraint (1.6), we are led to different VP's for P_v, namely, Eqs. (1.12) and (3.37b), respectively. We study the second-order errors associated with the two results. The introduction of $I_{1t}=I_1+\delta I_1$ and $I_{2t}+\delta I_2$ into Eq. (1.12), obtained without the current conservation constraint, gives

$$P_v=P-(R_1+R_2)\delta I_1\delta I_2\ . \tag{B1}$$

This result verifies that Eq. (1.12) indeed is a VP for P, but that it is *not* a variational bound, since $\delta I_1\delta I_2$ can be of either sign; δI_1 and δI_2 are regarded as independent in Eqs. (1.12) and (B1). On the other hand, if we insert $I_{1t}=I_1+\delta I_1$ into Eq. (3.37b) obtained with the current conservation constraint, we find

$$P_v=P+(\delta I_1)^2(R_1+R_2)\ . \tag{B2}$$

Therefore, Eq. (3.37b) is both a VP *and* a variational upper bound for P. As stated at the beginning of Sec. I.B.1 this is precisely the well-known extremum principle for resistor networks (Smythe, 1968, p. 252).

It is hardly surprising that the second approach, in which we chose $I_{2t}=I-I_{1t}$, led to a variational bound.

Since we had eliminated one of the trial currents, namely, I_{2t}, the second-order error term had to be of the form of a const$\times(\delta I_1)^2$, and we had to obtain a variational bound. This suggests that if we eliminate I_{2t} by choosing the trial I_{1t}, I_{2t} so that they satisfy Eq. (1.7) rather than Eq. (1.6) (that is, if we choose $I_{2t}=I_{1t}R_1/R_2$), we will again obtain a variational bound. Indeed, starting from

$$P_v = I_{1t}^2 R_1 + (I_{1t}R_1/R_2)^2 R_2 + \lambda_t'[I_{1t}+(I_{1t}R_1/R_2)-I] ,$$

we find

$$P_v = 2R_1 I I_{1t} - R_1\left[1+\frac{R_1}{R_2}\right] I_{1t}^2 , \tag{B3a}$$

which is yet a third VP for P. Moreover, Eq. (B3a) can be rewritten as

$$P_v = P - R_1\left[1+\frac{R_1}{R_2}\right](\delta I_1)^2 , \tag{B3b}$$

so that we now have a *lower* variational bound.

At the end of Sec. III.C.5, we gave a numerical example. The choice $I_{1t}=4.5$, $I_{2t}=3.5A$ was seen to give from Eq. (3.37b) an upper bound of 122 W on the power (exact value 120 W). The same choice in Eq. (B3a) gives 118.8 W, a lower bound.

APPENDIX C: SYSTEM GOVERNED BY NONLINEAR EQUATION NOT DERIVABLE FROM LAGRANGIAN

This appendix illustrates the remarks made in footnote 2. We shall construct a VP for the displacement of a particle whose equation of motion is not derivable from a Lagrangian. To illustrate the utility of our procedures further, we deliberately have chosen to examine a nonlinear equation of motion.

Consider the one-dimensional problem of a particle moving in the x direction, from $\tau=0$ to $\tau=1$, where we denote the time by τ to avoid confusion with t denoting "trial." The equation of motion is postulated to be

$$\left[\frac{d^2x}{d\tau^2}\right]^2 + a^3 x\frac{dx}{d\tau} = 0 , \tag{C1}$$

where a is a positive real constant.

It is readily seen that (C1) is not derivable from a Lagrangian—i.e., that it cannot be deduced from an analog of Hamilton's principle

$$\delta\int_0^1 d\tau\, N(x,\dot{x},\tau) = 0 , \tag{C2}$$

where the "Lagrangian" of this system is denoted by N to avoid confusion with our standard notation for generalized Lagrange multipliers (Appendix A). The Lagrange equation is

$$\frac{d}{d\tau}\frac{\partial N}{\partial \dot{x}} - \frac{\partial N}{\partial x} = 0 ,$$

or

$$\frac{\partial^2 N}{\partial\tau\partial\dot{x}} + \frac{\partial^2 N}{\partial x\partial\dot{x}}\dot{x} + \frac{\partial^2 N}{\partial\dot{x}^2}\ddot{x} - \frac{\partial N}{\partial x} = 0 . \tag{C3}$$

Only the third term in Eq. (C3) involves $\ddot{x}$, and this term is linear in $\ddot{x}$. Therefore, Eq. (C3) cannot possibly represent Eq. (C1), which is quadratic in $\ddot{x}$. Moreover, the deficiency of Eq. (C3) cannot be remedied by letting N depend explicitly on $\ddot{x}$, that is, by generalizing Eq. (C2) to

$$\delta\int_0^1 d\tau\, N(x,\dot{x},\ddot{x},\tau) = 0 .$$

In this event Lagrange's equation becomes

$$\frac{d^2}{d\tau^2}\frac{\partial N}{\partial\ddot{x}} - \frac{d}{d\tau}\frac{\partial N}{\partial\dot{x}} + \frac{\partial N}{\partial x} = 0 . \tag{C4}$$

The $d^2/d\tau^2$ operation in Eq. (C4) yields a term $(\partial^2 N/\partial\ddot{x}^2)(d^4x/d\tau^4)$, and this is the only term in Eq. (C4) involving $d^4x/d\tau^4$. It follows that Eq. (C4) cannot represent Eq. (C1) unless

$$\frac{\partial^2 N}{\partial\ddot{x}^2} = 0 ,$$

or

$$N(x,\dot{x},\ddot{x},\tau) = \ddot{x}\,f(x,\dot{x},\tau) + g(x,\dot{x},\tau) , \tag{C5}$$

where f and g are arbitrary. With the form (C5) for N, Eq. (C4) again contains no terms quadratic in $\ddot{x}^2$, and therefore again cannot represent Eq. (C1).[8]

Nevertheless, deriving a VP for an essentially arbitrary quantity depending on $x(\tau)$'s obeying Eq. (C1) is no problem. For example, suppose we seek a VP for

$$F = \int_0^1 d\tau\, w(x,\tau) , \tag{C6a}$$

where w is a specified function, with $x(\tau)$ obeying Eq. (C1) and obeying the boundary conditions

$$x(0)=\alpha , \quad \dot{x}(0)=\beta , \tag{C6b}$$

with α and β given real constants. Our starting point if we proceed as usual is

$$F_v = \int_0^1 d\tau\, w(x_t,\tau) + L_t^\dagger\left[\left[\frac{d^2x_t}{d\tau^2}\right]^2 + a^3 x_t\frac{dx_t}{d\tau}\right] , \tag{C7}$$

where $x_t(\tau)$ is a trial solution of Eq. (C1) obeying the boundary conditions (C6b), and where the notation in the last term of Eq. (C7) signifies there is an integral over τ from $\tau=0$ to $\tau=1$. The usual manipulations now lead to

[8] We remark that even when the equations of motion $B_i(\mathbf{r},\dot{\mathbf{r}},t)=0$ are not derivable from a Lagrangian, it may be possible to find a set of equations $\hat{B}_i(\mathbf{r},\dot{\mathbf{r}},t)=0$ which are derivable from a Lagrangian and whose solutions $\mathbf{r}(t)$ simultaneously guarantee $B_i(\mathbf{r},\dot{\mathbf{r}},t)=0$, although the set $\hat{B}_i$ is not identical to the original set B_i (Okubo, 1981). However, this possibility in no way diminishes the point of this appendix.

$$\delta F_v = \int_0^1 d\tau \frac{\partial w}{\partial x}\delta x + \int_0^1 d\tau \left[2\frac{d^2}{d\tau^2}(\ddot{x}L) - a^3 x\dot{L} \right]\delta x + \left[a^3 xL\delta x + 2\ddot{x}L\delta\dot{x} - 2\left[\frac{d}{d\tau}(\ddot{x}L) \right]\delta x \right]\Bigg|_0^1 \tag{C8}$$

for the first variation of F_v. To make the right-hand side of (C8) vanish for arbitrary δx, it will be necessary to require

$$2\frac{d^2}{d\tau^2}(\ddot{x}L) - a^3 x\dot{L} + \frac{\partial w}{\partial x} = 0 . \tag{C9}$$

In addition, because we have decided to choose x_t so that δx and $\delta\dot{x}$ vanish at $\tau=0$ but need not vanish at $\tau=1$, guaranteeing there is no contribution from the boundary terms in Eq. (C8) requires that the desired solution to Eq. (C9) satisfy the boundary conditions

$$L(1)=\dot{L}(1)=0 . \tag{C10}$$

Correspondingly, for x_t a first-order estimate of the exact x, Eq. (C7) should yield a variational estimate of F defined by Eq. (C6a) whenever the trial L_t in Eq. (C7) is a first-order estimate of the exact L satisfying Eqs. (C9) and (C10). One way, but by no means necessarily the only or even the best way of choosing L_t (recall the discussion in Sec. IV.C) is to choose the trial L_t as the solution of

$$2\frac{d^2}{d\tau^2}(\ddot{x}_t L_t) - a^3 x_t\dot{L}_t + \frac{\partial w}{\partial x_t} = 0 , \tag{C11}$$

subject to the boundary conditions $L_t(1)=\dot{L}_t(1)=0$. The fact that Eq. (C9) [and therefore its final version (C11)] is linear, even though our starting equation (C1) for x was nonlinear, is characteristic and noteworthy.

In the special case that

$$F = \int_0^1 d\tau\, x(\tau)\delta(\tau - c) = x(c) , \quad 0<c<1$$

Eq. (C7) provides a variational principle for $x(\tau)$ at any point $\tau=c$ in the interval $\tau=0$ to $\tau=1$. It is necessary to replace the generic term $\partial w/\partial x$ in Eq. (C9) by $\delta(\tau-c)$. Equivalently, one now replaces Eq. (C9) by

$$2\frac{d^2}{d\tau^2}(\ddot{x}L) - a^3 x\dot{L} = 0 ,$$

$$0\le\tau<c- \text{ and } c+<\tau\le 1 , \tag{C12a}$$

subject to

$$\left[2\frac{d}{d\tau}(\ddot{x}L) - a^3 xL \right]\Bigg|_{c-}^{c+} = -1 , \tag{C12b}$$

where $c+$ and $c-$ denote, respectively, points immediately to the right and to the left of $\tau=c$. Correspondingly, the term

$$-\left[a^3 xL\delta x + 2\ddot{x}L\delta\dot{x} - 2\left[\frac{d}{d\tau}(\ddot{x}L) \right]\delta x \right]\Bigg|_{c-}^{c+} \tag{C12c}$$

now must be added to the right-hand side of Eq. (C8), because the integral involving L_t in Eq. (C7) now must run from $\tau=0$ to $\tau=c-$ and then from $\tau=c+$ to $\tau=1$.

Equation (C12b) means that L now is a Green's function for the second-order differential equation (C12a). Because Eq. (C12a) is homogeneous in the domain $c+<\tau\le 1$, the boundary conditions (C10) imply $L(\tau)$ is identically zero in that domain. Because δx and $\delta\dot{x}$ need not vanish at $\tau=c$, the coefficients of these quantities each must vanish on the right-hand side of Eq. (C8). Recalling that (C12c) now is included in (C8), and that $\partial w/\partial x$ in (C8) now is $\delta(\tau-c)$, one sees that we now must require

$$\left[a^3 xL - 2\frac{d}{d\tau}(\ddot{x}L) \right]\Bigg|^{c-} + 1 = 0 \tag{C13a}$$

and

$$L(c-)=0 . \tag{C13b}$$

Equation (C13a) is gratifyingly identical with Eq. (C12b), given that $L(c+)=\dot{L}(c+)=0$. Equation (C13b), which guarantees there will be no terms proportional to $\delta\dot{x}(c)$ on the right of Eq. (C8), shows that $L(\tau)$, like conventional Green's functions, is continuous at $\tau=c$. Incidentally, the fact that $L(\tau)=0$ for $c\le\tau\le 1$ means that the variational estimate of $x(c)$ from the VP (C7) will not involve the trial estimate $x_t(\tau)$ for $\tau>c$, as one should expect. When, as in the present problem, the solution $x(c)$ is specified by the initial values $x(0)$ and $\dot{x}(0)$, the actual $x(c)$ can be thought to result from a step-by-step evolution of $x(\tau)$ via the differential equation (C1), from $\tau=0$ to $\tau=c$ but not [when $x(c)$ is sought] through times beyond $\tau=c$. Indeed, the very same VP would be reached more simply if, instead of (C7), the starting point [in the present problem, where $F\equiv x(c)$] were taken to be

$$x_v(c) = x_t(c) + \int_0^c d\tau\, L_t(\tau)\left[\left[\frac{d^2 x_t}{d\tau^2} \right]^2 + a^3 x_t(\tau)\frac{dx_t}{d\tau} \right] . \tag{C13c}$$

The above considerations have shown that a VP for the solution to Eq. (C1) can be constructed, even though Eq. (C1) is not derivable from a Lagrangian. Of course this VP of ours for $x(\tau)$ does not have the form (C2) of Hamilton's principle, nor would our VP have that form if the boundary conditions (C6b) were replaced by the customary (Hamilton's principle) boundary conditions that $x_t(\tau)=x(\tau)$ at the end points, i.e., by the requirement $\delta x=0$ at $\tau=0$ and $\tau=1$. These boundary conditions would require $L=0$ at $\tau=0$ and $\tau=1$, as is readily seen from Eq. (C8), and would produce corresponding modifications of the boundary conditions (C13) at $\tau=c$, but would not alter the differential equation (C9) for the exact L.

The solution of Eq. (C1) can be reduced to quadratures via the substitution

$$x(\tau) = \exp\int^\tau d\tau\, g(\tau) ,$$

but the functional form of $x(\tau)$ obtained in this way is

very complicated for arbitrary $x(0),\dot{x}(0)$, so that Eq. (C1) probably is best solved numerically for arbitrary α and β in Eq. (C6b). In the special case that $\alpha=1$ and $\beta=-a$, however, it is obvious that the solution to Eq. (C1) subject to the boundary conditions (C6b) is

$$x(\tau)=e^{-a\tau}\ . \tag{C14}$$

This result permits a simple test of the power of the VP (C7). Suppose for the case $\alpha=1,\beta=-a$ we wish to estimate $x(c)$ in the interval $0<c<1$, not realizing that there is the exact solution (C14). We will make the simplest possible choices for the trial functions $x_t(\tau)$ and $L_t(\tau)$ subject to the requirement that the trial quantities obey the same boundary conditions (C6b), (C10), and (C13) as do the exact functions. Therefore we will choose

$$x_t(\tau)=1-a\tau+b\tau^2\ , \tag{C15a}$$

with b a variational parameter; this $x_t(\tau)$ satisfies Eq. (C6b) for $\alpha=1,\beta=-a$; the quadratic term in (C15a) is needed (as will be seen immediately below) to determine $\dot{L}_t(c-)$, which is not specified by the trial version of Eqs. (C13) when $\ddot{x}_t(c)=0$. Using Eq. (C15a) in the trial version of Eqs. (C13), we see that the boundary conditions on L_t can be satisfied by a linear approximation to $L_t(\tau)$, namely,

$$\begin{aligned} &L_t(\tau)=\frac{1}{4b}(\tau-c)\ ,\quad 0\le\tau<c-\\ &L_t(\tau)=0\ ,\quad c+<\tau\le 1\ . \end{aligned} \tag{C15b}$$

Substituting these simple approximate forms (C15) into Eq. (C7) or (equivalently) Eq. (C13c), making the resultant expression stationary with respect to b, and using the value of b thereby obtained immediately yields the desired variational estimate $x_v(c)$. For comparison with the exact $x(\tau)$ given by (C14) it is convenient to expand the exact and variationally estimated $x(\tau)$ in powers of ac. We find

$$\langle x(c)\rangle_{\rm var}=1-ac+\frac{a^2c^2}{2}-\frac{a^3c^3}{6}+\frac{a^4c^4}{18}+\cdots\ . \tag{C16}$$

Thus, even with the very simple trial functions (C15), the VP (C7) [in the case $\alpha=1,\beta=-a$ in (C6b)] yields an estimate of $x(c)$ which is exact through terms of order a^3c^3, and whose a^4c^4 term is in error by only $a^4c^4/72$.[9]

APPENDIX D: ENERGY EIGENVALUE FOR A NON-HERMITIAN HAMILTONIAN WITH BOUNDARY CONDITIONS EXPLICITLY INVOLVING THE EIGENVALUE

We seek the eigenvalues E of

$$(H-E)\phi=0\ ,\quad -\infty<x<0-\ ,\quad 0+<x<\infty\ , \tag{D1a}$$

$$H\equiv\frac{d^2}{dx^2}-2x\frac{d}{dx}\ , \tag{D1b}$$

subject to the boundary conditions that ϕ and $d\phi/dx$ are everywhere continuous except at $x=0$, where $d\phi/dx$ (but not ϕ) is discontinuous, with the discontinuity

$$\left.\frac{d\phi}{dx}\right|_{0+}-\left.\frac{d\phi}{dx}\right|_{0-}=-aE\phi(0)\ , \tag{D2}$$

a being a positive real constant; as in Eqs. (C12), the symbols $0+$ and $0-$ denote points immediately to the right and left, respectively, of $x=0$.

Although H of (D1b) is not self-adjoint (see below), its eigenvalues E are real when $a=0$ in Eq. (D2); indeed, when $a=0$ the eigenfunctions are Hermite polynomials (Morse and Feshbach 1953, p. 786). The reality of the eigenvalues in the $a=0$ case can be understood by noting that the function $u=\phi\exp(-\frac{1}{2}x^2)$ satisfies the equation

$$\left[\frac{d^2}{dx^2}-x^2+1\right]u=Eu\ . \tag{D3}$$

When u and du/dx can be supposed everywhere continuous, as is the case when $a=0$ in Eq. (D2), the operator within the parentheses on the left-hand side of Eq. (D3) is obviously self-adjoint. Moreover, the eigenvalues remain real even when $a\neq 0$ in Eq. (D2), as can be seen by multiplying Eq. (D3) [which is valid for $a\neq 0$ as well as for $a=0$] by u^*, integrating over all space and then integrating by parts, noting that u satisfies the same boundary condition (D2) as does ϕ.

The foregoing considerations correctly suggest that finding a VP for the present model problem would be simplified by introducing $u=\phi\exp(-\frac{1}{2}x^2)$, i.e., by

[9] Actually, the VP in this problem is stationary at two values of b, $b=\pm(a^2/2)(1-ac/3)^{1/2}(1-a^3c^3/20)^{-1/2}$. The result (C16) is obtained using the plus sign in this formula; using the negative sign yields an estimate of $x(c)$ which is incorrect as early as the a^2c^2 term. Finding multiple roots for the values of the variational parameters making a VP stationary is commonplace; deciding *a priori* which root will give the best variational estimate demands some physical or mathematical insight into the problem and can be difficult. For the VP for $x(c)$ examined in this appendix, the fact that the positive value of b gives a better variational estimate may be understandable on the basis that the original differential equation (C1) requires that x and $\dot{x}$ have opposite signs, which in turn implies that the particle velocity must *increase* from negative values to zero as the particle moves toward $x=0$ from points $x>0$ to the right of the origin. In other words, the actual solution to the differential equation (C1) subject to the boundary conditions $x(0)=1,\dot{x}(0)=-a$ is expected to have a positive acceleration $\ddot{x}(\tau)$, at least for small τ when $x(\tau)$ and $\dot{x}(\tau)$ surely remain greater than 0 and less than 0, respectively. The acceleration in the trial estimate $x_t(\tau)$ of (C15a) is $2b$; therefore for τ small, $b>0$ in (C15a) corresponds to a much more physically reasonable choice of $x_t(\tau)$ than does $b<0$, and should yield a correspondingly better variational estimate of $x(c)$ at small ac.

working with the Hermitian operator of Eq. (D3) rather than with H of Eq. (D1b). Nevertheless, for the purposes of this appendix we shall continue to work with H, so as to illustrate the construction of the VP for the eigenvalue when the differential equation for the eigenfunction involves an operator which is not immediately self-adjoint, with the additional complication that the boundary conditions on the eigenfunction involve a discontinuity in derivative proportional to the eigenvalue itself. However, in the derivation which follows we shall recognize that the eigenvalues E are real, because this permits us to deal solely with real functions, thereby very much simplifying the analysis without losing the essential features. The derivation of the VP when E is complex is sketched at the end of this appendix.

For real functions v and w, the adjoint $H^\dagger$ of H is defined by the relation

$$(H^\dagger v)^\dagger w = v^\dagger (Hw) , \tag{D4}$$

where (in the present problem) the notation of Eq. (D4) implies integration over x from $x=-\infty$ to ∞, but does not require taking the complex conjugate of any quantities. It is implicit that v and/or w vanish sufficiently rapidly at $\pm\infty$ to ensure convergence of the integrals in Eq. (D4).

Integrating $v^\dagger(Hw)$ by parts, we find that

$$\begin{aligned} v^\dagger(Hw) &= \int_{-\infty}^{\infty} dx\, v(Hw) \\ &= \int_{-\infty}^{\infty} dx (H^\dagger v) w \\ &\quad + \left[2xvw - \left[v\frac{dw}{dx} - w\frac{dv}{dx} \right] \right] \Bigg|_{0-}^{0+} , \end{aligned} \tag{D5a}$$

where

$$H^\dagger \equiv \frac{d^2}{dx^2} + 2x\frac{d}{dx} + 2 . \tag{D5b}$$

The manipulations involved in Eq. (D5a) are much as in Eqs. (C12). Equation (D5b) shows $H^\dagger \neq H$, i.e., H is not self-adjoint. In addition, in order that $H^\dagger$ satisfy the equality (D4), it is necessary to impose boundary conditions on v and w which will guarantee that there is no contribution to (D5a) from the discontinuity at $x=0$. If we assume that w satisfies Eq. (D2), satisfactory boundary conditions on v are seen to be: v continuous at $x=0$, but dv/dx discontinuous, with

$$\left.\frac{dv}{dx}\right|_{0+} - \left.\frac{dv}{dx}\right|_{0-} = -aEv(0) , \tag{D6}$$

i.e., in the present problem the functions v for which Eq. (D4) holds happen to satisfy the same boundary conditions as do the functions w, when the functions w are supposed to obey Eq. (D2).

The eigenvalues E of

$$(H^\dagger - E)\psi = 0 , \tag{D7}$$

subject to the boundary conditions that ψ is everywhere continuous but $d\psi/dx$ satisfies Eq. (D6), coincide with the eigenvalues E specified by Eqs. (D1) and (D2); this result is the immediate consequence of multiplying Eq. (D1a) by ψ and using Eq. (D4). Now let E represent any one of these eigenvalues, e.g., the lowest, and let ϕ and ψ denote their corresponding eigenfunctions satisfying Eqs. (D1) and (D7), respectively. Further, let E_t, ϕ_t, ψ_t denote first-order trial quantities to the exact E, ϕ, ψ, where we shall require that ϕ_t, ψ_t satisfy the same sort of boundary conditions as ϕ, ψ, i.e., that ϕ_t is continuous at $x=0$ but that $d\phi_t/dx$ has the discontinuity

$$\left.\frac{d\phi_t}{dx}\right|_{0+} - \left.\frac{d\phi_t}{dx}\right|_{0-} = -aE_t\phi_t(0) , \tag{D8}$$

and similarly for ψ_t. Then, according to our standard procedure, a starting point for a variational estimate of E should be

$$\begin{aligned} E_v \equiv \langle E \rangle_{\rm var} &= E_t + L_{at}^\dagger [(H-E_t)\phi_t] \\ &\quad + [(H^\dagger - E_t)\psi_t]^\dagger L_{bt} . \end{aligned} \tag{D9}$$

The first-order variation of Eq. (D9) gives

$$\begin{aligned} \delta E_v &= \delta E + L_a^\dagger [(H-E)\delta\phi - \phi\delta E] \\ &\quad + [(H^\dagger - E)\delta\psi - \psi\delta E]^\dagger L_b . \end{aligned} \tag{D10a}$$

But, as in Eq. (D5a), we have

$$\begin{aligned} L_a^\dagger [(H-E)\delta\phi] &= [(H^\dagger - E)L_a]^\dagger \delta\phi \\ &\quad - \left[L_a \frac{d\delta\phi}{dx} - \delta\phi \frac{dL_a}{dx} \right] \Bigg|_{0-}^{0+} , \end{aligned} \tag{D10b}$$

$$\begin{aligned} [(H^\dagger - E)\delta\psi]^\dagger L_b &= \delta\psi^\dagger [(H-E)L_b] \\ &\quad - \left[L_b \frac{d\delta\psi}{dx} - \delta\psi \frac{dL_b}{dx} \right] \Bigg|_{0-}^{0+} . \end{aligned}$$

Also, Eqs. (D2) and (D8) imply

$$\left.\frac{d}{dx}\delta\phi\right|_{0+} - \left.\frac{d}{dx}\delta\phi\right|_{0-} = -a[E\delta\phi(0) + \phi(0)\delta E] , \tag{D10c}$$

and similarly for $\delta\psi$ at $x=0$. Using Eqs. (D10b) and (D10c) in Eq. (D10a), we find that δE_v will vanish to first order if

$$(H-E)L_b = 0 , \tag{D11a}$$

$$(H^\dagger - E)L_a = 0 , \tag{D11b}$$

$$1 - L_a^\dagger \phi - \psi^\dagger L_b + aL_a(0)\phi(0) + aL_b(0)\psi(0) = 0 , \tag{D11c}$$

and L_a, L_b each are continuous at $x=0$ but have a first derivative discontinuity satisfying Eqs. (D2) or (D6).

Equation (D11a) and these boundary conditions on L_b show that L_b is a multiple of ϕ, i.e.,

$$L_b = c_b\phi . \tag{D12a}$$

Similarly,

$$L_a = c_a\psi . \tag{D12b}$$

The magnitudes of the multiplicative factors c_a, c_b remain to be determined. With the simplest choice $c_a = c_b = c$, Eq. (D11c) yields

$$c = \frac{1}{2} \frac{1}{\psi^\dagger \phi - a\psi(0)\phi(0)} . \tag{D13a}$$

Therefore, first-order estimates of L_b and L_a should be $L_{bt} = c_t \phi_t$ and $L_{at} = c_t \psi_t$, with

$$c_t = \frac{1}{2} \frac{1}{\psi_t^\dagger \phi_t - a\psi_t(0)\phi_t(0)} . \tag{D13b}$$

Substituting these L_{bt} and L_{at} into Eq. (D9) yields the desired VP

$$E_v = E_t + \frac{1}{2} \frac{1}{\psi_t^\dagger \phi_t - a\psi_t(0)\phi_t(0)} \{ \psi_t^\dagger [(H - E_t)\phi_t] + [(H^\dagger - E_t)\psi_t]^\dagger \phi_t \} . \tag{D14}$$

Equation (D14), like Eq. (1.41)—to which Eq. (D14) reduces when $a = 0$ and $H^\dagger = H$—is independent of the normalization of the trial functions ϕ_t and ψ_t. That the first variation of (D14) is zero, i.e., that (D14) really is a VP under the stated boundary conditions, can be readily verified directly from (D14).

When the eigenvalue E is not real, the problem of finding a VP is more complicated in several respects. Suppose we had not recognized that the eigenvalues E of Eq. (D3) remain real even when $a \neq 0$ in Eq. (D2). Then v and w will have to be treated as complex functions in Eq. (D4), and it will be necessary to replace Eq. (D6) by

$$\left. \frac{dv}{dx} \right|_{0+} - \left. \frac{dv}{dx} \right|_{0-} = -aE^* v(0) \tag{D15a}$$

in order that Eq. (D4) shall continue to hold; Eq. (D5b) remains unchanged however. Correspondingly, Eq. (D7) is replaced by

$$(H^\dagger - E^*)\psi = 0 , \tag{D15b}$$

that is, the eigenvalues of $H^\dagger$ are the quantities E^* when the eigenfunctions ψ are required to obey Eq. (D15a).

The result (D15b) suggests that with a single simple modification—namely, replacing E_t by E_t^* in the last term of Eq. (D9)—the right-hand side of Eq. (D9) will be a suitable starting point for a variational estimate of E when E is complex. However, this starting point has the deficiency that it contains but a single term involving E_t^*, which will cause difficulties when attempting to set the coefficient of δE^* equal to zero (when E is complex, the first-order variations δE and δE^* should be regarded as independent, just as the variations $\delta\phi$ and $\delta\phi^\dagger$ were taken to be independent in Sec. III.A and later sections). Also, because in practice the values of the parameters (introduced via the trial quantities) making the real part of the VP stationary need not coincide with the values making the imaginary part stationary, it is desirable to construct separate VP's for the real and imaginary parts of E.

In light of the above considerations, a suitable starting point for a VP for the real part of E should be

$$\langle \mathrm{Re}(E) \rangle_{\mathrm{var}} = \tfrac{1}{2}(E_t + E_t^*) + \tfrac{1}{2}\{L_{at}^\dagger [(H - E_t)\phi_t] + (L_{at}^\dagger [(H - E_t)\phi_t])^*\} + \tfrac{1}{2}\{[(H^\dagger - E_t^*)\psi_t]^\dagger L_{bt} + ([(H^\dagger - E_t^*)\psi_t]^\dagger L_{bt})^*\} . \tag{D16}$$

Proceeding as usual, we again deduce Eqs. (D12), and eventually arrive at the VP

$$\langle \mathrm{Re}(E) \rangle_{\mathrm{var}} = \tfrac{1}{2}(E_t + E_t^*) + \frac{1}{2c}\{\psi_t^\dagger [(H - E_t)\phi_t] + (\psi_t^\dagger [(H - E_t)\phi_t])^*\} + \frac{1}{2c}\{[(H^\dagger - E_t^*)\psi_t]^\dagger \phi_t + ([(H^\dagger - E_t^*)\psi_t]^\dagger \phi_t)^*\} , \tag{D17a}$$

where

$$c = \psi_t^\dagger \phi_t + \phi_t^\dagger \psi_t - a[\psi_t^*(0)\phi_t(0) + \psi_t(0)\phi_t^*(0)] . \tag{D17b}$$

Equations (D17) obviously reduce to Eq. (D14) when all the trial quantities in (D17) are purely real.

REFERENCES

Aranoff, S., and J. K. Percus, 1968, Phys. Rev. **166**, 1255.

Aronson, I., R. Blau, C. J. Kleinman, L. Rosenberg, and L. Spruch, 1979, Phys. Rev. A **19**, 1568.

Arthurs, A. M., 1970, *Complementary Variational Principles* (Clarendon, Oxford).

Aspinall, L., and I. C. Percival, 1967, in *Fifth International Conference on the Physics of Electronic and Atomic Collisions, Leningrad,* edited by I. P. Flaks (Nauka, Leningrad), p. 109.

Atherton, R. W., and G. M. Homsy, 1975, Stud. Appl. Math. **54**, 31.

Bardsley, J. N., E. Gerjuoy, and C. V. Sukumar, 1972, Phys. Rev. A **6**, 1813.

Barnsley, M. F., and P. D. Robinson, 1974, Proc. R. Soc. London, Ser. A **338**, 527.

Barnsley, M. F., and P. D. Robinson, 1977, J. Inst. Math. Its Appl. **20**, 485.

Becker, M., 1964, *The Principles and Applications of Variational Methods* (MIT, Cambridge, Mass.).

Bellman, R., and R. Kalaba, 1959, J. Math. Mech. **8**, 683.

Bellman, R., and G. M. Wing, 1975, *An Introduction to Invariant Imbedding* (Wiley-Interscience, New York).

Bethe, H. A., and E. E. Salpeter, 1957, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer, Berlin).

Biot, Maurice A., 1970, *Variational Principles in Heat Transfer* (Oxford, London).

Blankenbecler, R., 1966, Phys. Rev. **152**, 1471.
Blau, R., A.R.P. Rau, and L. Spruch, 1973, Phys. Rev. A **8**, 131.
Blau, R., L. Rosenberg, and L. Spruch, 1974, Phys. Rev. A **10**, 2246.
Blau, R., L. Rosenberg, and L. Spruch, 1975, Phys. Rev. A **11**, 200.
Blau, R., L. Rosenberg, and L. Spruch, 1977, Phys. Rev. A **15**, 1475.
Borowitz, S., and E. Gerjuoy, 1965, University of Colorado, Boulder, Colorado, Joint Institute of Laboratory Astrophysics Report No. 36.
Brownstein, K. R., and W. A. McKinley, 1968, Phys. Rev. **170**, 1255.
Burke, P. G., and W. D. Robb, 1975, Adv. At. Mol. Phys. **11**, 143.
Cairo, L., and T. Kahan, 1965, *Variational Techniques in Electromagnetism* (Gordon and Breach, New York).
Callaway, J., 1978, Phys. Rep. **45**, 89.
Carew, J., and L. Rosenberg, 1973, Phys. Rev. D **7**, 1113.
Chatwin, R. A., and J. E. Purcell, 1971, J. Math. Phys. **12**, 2024.
Cheng, E. T., 1980, Nucl. Sci. Eng. **74**, 147.
Cole, R. J., 1979, Proc. R. Soc. Edinburg, Sec. A **82**, 211.
Collins, W. D., 1977, Proc. R. Soc. Edinburgh, Sec. A **77**, 273.
Conn, R. W., and W. M. Stacey, Jr., 1973, Nucl. Fusion **13**, 185.
Courant, R., 1937, *Differential and Integral Calculus,* translated by E. J. McShane (Interscience, New York), Vol. 1.
Courant, R., and D. Hilbert, 1937, *Methods of Mathematical Physics* (Interscience, New York), Vol. 1.
Dalgarno, A., and J. T. Lewis, 1956, Proc. R. Soc. London, Ser. A **233**, 70.
Dalgarno, A., and A. L. Stewart, 1956, Proc. R. Soc. London, Ser. A **238**, 269.
Dalgarno, A., and A. L. Stewart, 1960, Proc. R. Soc. London, Ser. A **257**, 534.
Dalgarno, A., W. D. Davison, and A. L. Stewart, 1956, Proc. R. Soc. London, Ser. A **257**, 115.
Delves, L. M., 1963a, Nucl. Phys. **41**, 497.
Delves, L. M., 1963b, Nucl. Phys. **45**, 313.
Delves, L. M., 1963c, Nucl. Phys. **45**, 443.
Demkov, Yu. N., 1963, *Variational Principles in the Theory of Collisions* (MacMillan, New York).
Dirac, P. A. M., 1958, *Principles of Quantum Mechanics,* 4th ed. (Oxford University, London).
Donnelly, R. J., R. Herman, and I. Prigogine, 1965, *Non-Equilibrium Thermodynamics, Variational Principles and Stability* (University of Chicago, Chicago).
Epstein, S. T., and J. O. Hirschfelder, 1961, Phys. Rev. **123**, 1495.
Fano, U., and C. M. Lee, 1973, Phys. Rev. Lett. **31**, 1573.
Feshbach, H., and S. Rubinow, 1952, Phys. Rev. **88**, 484.
Feynman, R. P., R. B. Leighton, and M. Sands, 1964, *The Feynman Lectures on Physics* (Addison-Wesley, Reading), Vol. II.
Finlayson, B. A., 1972, *The Method of Weighted Residuals and Variational Principles* (Academic, New York).
Funk, Paul, 1962, *Variationsrechnung und ihre Anwendung in Physik und Technik* (Springer, Berlin).
Gandini, A., 1967, J. Nucl. Energy **21**, 755.
Gerjuoy, E., 1971, Philos. Trans. R. Soc. London, Ser. A **270**, 197 (Number 1203).
Gerjuoy, E., A. R. P. Rau, L. Rosenberg, and L. Spruch, 1974, Phys. Rev. A **9**, 108.
Gerjuoy, E., A. R. P. Rau, L. Rosenberg, and L. Spruch, 1975, J. Math. Phys. **16**, 1104.
Gerjuoy, E., A. R. P. Rau, and L. Spruch, 1972, J. Math. Phys. **13**, 1797.
Gerjuoy, E., A. R. P. Rau, and L. Spruch, 1973, Phys. Rev. A **8**, 662.
Gerjuoy, E., L. Rosenberg, and L. Spruch, 1975, J. Math. Phys. **16**, 455.
Goldstein, H., 1980, *Classical Mechanics,* 2nd ed. (Addison-Wesley, Reading, Mass).
Gould, S. H., 1957, *Variational Methods for Eigenvalue Problems* (University of Toronto, Toronto).
Hahn, Y., T. F. O'Malley, and L. Spruch, 1962, Phys. Rev. **128**, 932.
Hahn, Y., T. F. O'Malley, and L. Spruch, 1963, Phys. Rev. **130**, 381.
Hahn, Y., T. F. O'Malley, and L. Spruch, 1964, Phys. Rev. **134**, B397 and B911.
Harris, F. E., and H. H. Michels, 1971, "Expansion methods for electron-atom scattering," in *Methods in computational physics,* Vol. 10 (Atomic and Molecular Scattering), edited by Berni Alder, Sidney Fernbach, and Manuel Rotenberg (Academic, New York), p. 143.
Haymaker, R. W., and R. Blankenbecler, 1968, Phys. Rev. **171**, 1581.
Heaton, M., and B. L. Moiseiwitsch, 1971, J. Phys. B **4**, 332.
Heinrichs, Jean, 1968a, Phys. Rev. **172**, 1315.
Heinrichs, Jean, 1968b, Phys. Rev. **176**, 2168.
Hirschfelder, J. O., and C. A. Coulson, 1962, J. Chem. Phys. **36**, 941.
Hirschfelder, J. O., W. Byers-Brown, and S. T. Epstein, 1964, Adv. Quantum Chem. **1**, 255.
Houston, S. K., 1973, J. Phys. B **6**, 136.
Huang, S. S., 1949a, Phys. Rev. **76**, 477.
Huang, S. S., 1949b, Phys. Rev. **76**, 866.
Huang, S. S., 1949c, Phys. Rev. **78**, 1878.
Hulthén, L., 1944, K. Fysiogr. Saellsk. Lund, Foerh. **14**, No. 21.
Hulthén, L., 1948, Ark. Mat. Astron. Fys. A **35**, No. 25.
Hylleraas, E. A., and B. Undheim, 1930, Z. Phys. **65**, 759.
Inokuti, M., D. Douthat, and A. R. P. Rau, 1980, Phys. Rev. A **22**, 445.
Inokuti, M., H. Sekine, and T. Mura, 1978, "General use of the Lagrange multiplier in nonlinear mathematical physics," in *Variational Methods in the Mechanics of Solids,* edited by S. Nemat-Nasser (Pergamon, Oxford), p. 156.
Kalaba, R., 1959, J. Math. Mech. **8**, 519.
Kalikstein, K., 1981, J. Math. Phys. **22**, 1433.
Kalikstein, K., and L. Spruch, 1964, J. Math. Phys. **5**, 1261.
Kantorovich, L. V., and G. P. Akilov, 1964, "Newton's method," Chap. IV of *Functional Analysis in Normed Spaces,* translated by D. E. Brown and edited by A. P. Robertson (Macmillan, New York), p. 695.
Kato, Tosio, 1950, Phys. Rev. **80**, 475.
Kato, Tosio, 1951, Prog. Theor. Phys. (Japan) **6**, 394.
Kirtman, Bernard, 1971, J. Chem. Phys. **55**, 1457.
Kirtman, B., and R. L. Mowery, 1971, J. Chem. Phys. **55**, 1447.
Kleinman, C. J., Y. Hahn, and L. Spruch, 1965, Phys. Rev. **140**, A413.
Kohn, W., 1948, Phys. Rev. **74**, 1763.
Kohn, W., 1951, Phys. Rev. **84**, 495.
Kostin, M. D., and H. Brooks, 1964, J. Math. Phys. **5**, 1691.
Krieger, J. B., and V. Sahni, 1972, Phys. Rev. A **6**, 919.
Lanczos, Cornelius, 1970, *The Variational Principles of Mechanics,* 4th ed. (University of Toronto, Toronto).

Langhoff, P. W., S. T. Epstein, and M. Karplus, 1972, Rev. Mod. Phys. **44**, 602.
Lee, C. M., 1974, Phys. Rev. A **10**, 584.
Lewins, J., 1965, *Importance: The Adjoint Function* (Pergamon, London).
Lieber, M., L. Rosenberg, and L. Spruch, 1972a, Phys. Rev. D **5**, 1330.
Lieber, M., L. Rosenberg, and L. Spruch, 1972b, Phys. Rev. D **5**, 1347.
Lippmann, B. A., 1956, Phys. Rev. **102**, 264.
Lippmann, B. A., and J. Schwinger, 1950, Phys. Rev. **79**, 469.
Löwdin, P. O., and P. K. Mukherjee, 1972, Chem. Phys. Lett. **14**, 1.
Lucchese, R. R., and V. McKoy, 1980, Phys. Scr. **21**, 366.
Luttinger, J. M., and R. B. Thomas, 1960, J. Math. Phys. **1**, 121.
Lyness, J. N., and J. M. Blatt, 1961, J. Austr. Math. Soc. **2**, 153.
Lyness, J. N., and J. M. Blatt, 1962, J. Austr. Math. Soc. **2**, 357.
MacDonald, J. K. L., 1933, Phys. Rev. **43**, 830.
McWhirter, J. D. G., and B. L. Moiseiwitsch, 1972, J. Phys. B **5**, 1439.
Makinson, R. E. B., and J. S. Turner, 1953, Proc. Phys. Soc. A **66**, 857.
Maleki, N., 1981, Ph.D. thesis (University of Nebraska), unpublished.
Maleki, N., and J. Macek, 1980, Phys. Rev. A **21**, 1403.
Marshak, R. E., 1947, Phys. Rev. **71**, 688.
Mercier, Andre, 1963, *Variational Principles of Physics* (Dover, New York).
Michlin, Solomon G., 1964, *Variational Methods in Mathematical Physics* (MacMillan, New York).
Miller, P. H., 1957, Am. J. Phys. **25**, 30.
Moiseiwitsch, B. L., 1951, Phys. Rev. **82**, 753.
Moiseiwitsch, B. L., 1966, *Variational Principles* (Interscience, New York).
Monaghan, W., and L. Rosenberg, 1972, Phys. Rev. A **6**, 1076.
Morse, P. M., and H. Feshbach, 1953, Methods of Theoretical Physics (McGraw-Hill, New York), Vols. I and II.
Mott, N. F., and H. S. W. Massey, 1965, *The Theory of Atomic Collisions,* 3rd ed. (Oxford University Press, London).
Mueller, R. O., A. R. P. Rau, and L. Spruch, 1974, Phys. Rev. A **10**, 1511.
Nelson, Paul, 1975, Nucl. Sci. Eng. **56**, 340.
Nesbet, R. K., 1969, Phys. Rev. **179**, 60.
Nesbet, R. K., 1980, *Variational Methods in Electron-Atom Scattering Theory* (Plenum, New York).
Nesbet, R. K., and R. S. Oberoi, 1972, Phys. Rev. A **6**, 1855.
Newton, Roger G., 1982, *Scattering Theory of Waves and Particles,* 2nd ed. (Springer, New York).
Noble, B., and M. J. Sewell, 1972, J. Inst. Math. Its Appl. **9**, 123.
Nuttall, J., 1967, Phys. Rev. Lett. **19**, 473.
Okubo, S., 1981, Phys. Rev. A **23**, 2776.
Page, B. A. P., 1975, J. Phys. B **8**, 2486.
Page, B. A. P., 1976, J. Phys. B **9**, 67.
Percival, I. C., 1957, Proc. Phys. Soc. (London) **70**, 494.
Percival, I. C., 1960, Phys. Rev. **119**, 159.
Peterkop, R., and L. Rabik, 1971, J. Phys. B **4**, 1440.
Petrov, Iu. P., 1968, *Variational Methods in Optimum Control Theory* (Academic, New York).
Pieper, S. C., L. Schlessinger, and Jon Wright, 1970, Phys. Rev. D **1**, 1674.
Pieper, S. C., Jon Wright, and Leonard Schlessinger, 1970, Phys. Rev. D **2**, 1561.
Pomraning, G. C., 1967a, J. Math. Phys. **8**, 149.
Pomraning, G. C., 1967b, J. Math. Phys. **8**, 2096.
Pomraning, G. C., 1968, J. Math. Phys. (Cambridge, Mass.) **47**, 155.
Rasiel, Y., and D. R. Whitman, 1965, J. Chem. Phys. **42**, 2124.
Rau, A. R. P., 1974, Mon. Not. R. Astr. Soc. **168**, 273.
Rau, A. R. P., 1976, J. Math. Phys. **17**, 1338.
Rau, A. R. P., and A. K. Rajagopal, 1975, Phys. Rev. B **11**, 3604.
Rau, A. R. P., and A. K. Rajagopal, 1976, Phys. Rev. B **14**, 3052.
Rau, A. R. P., M. Inokuti, and D. Douthat, 1978, Phys. Rev. A **18**, 971.
Rayleigh, J. W. S., 1871, Phil. Trans. R. Soc. London, Ser. A **161**, 77; reprinted in *Scientific Papers* (Dover, New York, 1964), Vol. I, No. 5.
Rayleigh, J. W. S., 1873, Proc. Lond. Math. Soc. **4**, 357; reprinted in *Scientific Papers* (Dover, New York, 1964), Vol. I, No. 21.
Rayleigh, J. W. S., 1877, *The Theory of Sound;* reprinted (Dover, New York, 1945).
Rayleigh, J. W. S., 1911, Philos. Mag. **22**, 225; reprinted in *Scientific Papers* (Dover, New York, 1964), Vol. VI, No. 357.
Rektorys, K., 1975, *Variational Methods in Mathematics, Science, and Engineering* (Reidel, Holland).
Risberg, V., 1956, Arch. Math. Naturvidensk. **53**, 1.
Ritz, W., 1908, J. Reine Angew. Math. **135**, 1.
Ritz, W., 1909, Ann. Phys. (Leipzig) **28**, 737.
Robinson, P. D., 1971, in *Nonlinear Functional Analysis and Applications,* edited by L. B. Rall (Academic, New York), p. 507.
Robinson, P. D., and A. M. Arthurs, 1968, J. Math. Phys. **9**, 1364.
Rosenberg, L., 1964, Phys. Rev. **134**, B937.
Rosenberg, L., 1974, Phys. Rev. D **9**, 1789.
Rosenberg, L., and L. Spruch, 1960, Phys. Rev. **120**, 474.
Rosenberg, L., L. Spruch, and T. F. O'Malley, 1960, Phys. Rev. **118**, 184.
Rosenberg, L., and S. Tolchin, 1973, Phys. Rev. A **8**, 1702.
Roussopoulos, P., 1953, C. R. Acad. Sci. **236**, 1858.
Rubinow, S. I., 1954, Phys. Rev. **96**, 218.
Rubinow, S. I., 1955, Phys. Rev. **98**, 183.
Schiff, L. I., 1968, *Quantum Mechanics,* 3rd ed. (McGraw-Hill, New York).
Schwartz, Charles, 1959a, Ann. Phys. (N.Y.) **6**, 156.
Schwartz, Charles, 1959b, Ann. Phys. (N.Y.) **6**, 170.
Schwinger, J., 1947, Phys. Rev. **72**, 742.
Schwinger, J., and D. S. Saxon, 1968, *Discontinuities in Waveguides* (Gordon and Breach, New York).
Scott, W. R., and B. Kirtman, 1972, J. Chem. Phys. **56**, 1685.
Sewell, M. J., and B. Noble, 1978, Proc. R. Soc. London, Ser. A **361**, 293.
Shakeshaft, R., L. Rosenberg, and L. Spruch, 1976, Phys. Rev. A **14**, 1989.
Shakeshaft, R., and L. Spruch, 1974, Phys. Rev. A **10**, 92.
Shimamura, I., 1971a, J. Phys. Soc. Jpn. **30**, 1702.
Shimamura, I., 1971b, J. Phys. Soc. Jpn. **31**, 852.
Shimamura, I., 1978, "R-matrix theories," in *The Physics of Electronic and Atomic Collisions,* Invited Papers, X ICPEAC, edited by G. Watel (North-Holland, Amsterdam) p. 213.
Sims, J. S., and J. R. Rumble Jr., 1973, Phys. Rev. A **8**, 2231.
Smythe, W. R., 1968, *Static and Dynamic Electricity,* 3rd ed. (McGraw-Hill, New York).
Spruch, L., 1962, in *Lectures in Theoretical Physics,* edited by W. E. Brittin, W. B. Downs, and J. Downs (Wiley, New York)

Vol. 4, p. 161.
Spruch, L., 1969, in *Lectures in Theoretical Physics,* edited by S. Geltman, K. T. Mahantappa, and W. E. Brittin (Gordon and Breach, New York), Vol. XI-C, p. 77.
Spruch, L., 1975, in *The Physics of Electronic and Atomic Collisions,* Invited Papers IX ICPEAC, edited by J. S. Risley and R. Geballe (University of Washington Press, Seattle), p. 685.
Spruch, L., and L. Rosenberg, 1959, Phys. Rev. **116**, 1034.
Spruch, L., and L. Rosenberg, 1960, Phys. Rev. **117**, 143.
Stacey, W. M., Jr., 1972, J. Math. Phys. **13**, 1119.
Strieder, W., and S. Prager, 1967, J. Math. Phys. **8**, 514.
Sugar, R., and R. Blankenbecler, 1964, Phys. Rev. **136**, B472.
Takatsuka, K., and T. Fueno, 1979, Phys. Rev. A **19**, 1011 and 1018.
Tamm, I. E., 1948, Zh. Eksp. Teor. Fiz. **18**, 337.
Tamm, I. E., 1949, Zh. Eksp. Theor. Fiz. **19**, 74.
Usachev, L. N., 1964, J. Nucl. Energy **18**, 571.
Vainberg, M. M., 1964, *Variational Methods for the Study of Nonlinear Operators* (Holden-Day, San Francisco).
Wadehra, J. M., L. Spruch, and R. Shakeshaft, 1978, Phys. Rev. A **18**, 344.
Watson, D. K., R. R. Lucchese, V. McKoy, and T. N. Rescigno, 1980, Phys. Rev. A **21**, 738.
Weinhold, F., 1972, Adv. Quantum Chem. **6**, 299.
Wigner, E. P., 1932, Nachr. Ges. Wiss. Goettingen, Math.-Phys. Kl. **32**, 35.
Wigner, E. P., 1948, Phys. Rev. **73**, 1002.
Winick, J. R., and W. P. Reinhardt, 1978, Phys. Rev. A **18**, 910 and 925.
Wladawsky, I., 1973, J. Chem. Phys. **58**, 1826.
Yourgrau, W., and S. Mandelstam, 1968, *Variational Principles in Dynamics and Quantum Theory* (Saunders, Philadelphia).