Study of the Singularities in the Kohn Variational Method for Scattering Parameters*

KENNETH R. BROWNSTEIN† AND WILLIAM A. MCKINLEY
Rensselaer Polytechnic Institute, Troy, New York
(Received 18 December 1967)

The Kohn variational method for calculating scattering parameters is formulated in an abstract vector space in terms of the extremization of a quadratic form whose coefficients represent matrix elements of $H-E$. This manifestly possesses certain symmetry and invariance properties not apparent in the usual formulation. Under certain conditions the method provides a bound for the exact answer if the number of variational parameters, N, is large enough. The extraneous singularities in the variationally determined value of $\tan\delta$ noticed by Schwartz are analyzed. These are a direct consequence of the Kohn method and the particular normalization employed. It is shown that the pole strength (residue) of each such singularity is the product of two terms, each of which is small; these factors may be expected to decrease with increasing N.

I. INTRODUCTION

A. Kohn Method

FOR purposes of introduction, consider the problem of the S-wave scattering of a spinless particle by a spherically symmetric potential. That is, we seek to solve the Schrödinger equation

$$H\psi=[-d^2/dr^2+V(r)]\psi=E\psi \tag{1}$$

for the reduced radial wave function $\psi(r)$, subject to the single boundary condition $\psi(0)=0$. If the potential $V(r)$ goes to zero fast enough for large r (as will be assumed from now on), then the asymptotic form of ψ for large r is

$$\psi(r)\propto\sin(kr+\delta),\quad k=\sqrt{E}. \tag{2}$$

The phase shift δ so defined is all that is usually required to predict experimental results (scattering cross sections).

The scattering wave function is not square integrable. The successful use of a variational principle for the scattering problem then raises the question of just what type of constraint(s) or normalization should be used, different choices yielding different formulations. A frequent starting point for the discussion of such variational principles is the "Kato identity,"[1,2] whose derivation for the S-wave case is sketched below.

Let $\psi(r)$ denote a trial wave function normalized such that

$$\psi(0)=0, \tag{3a}$$

$$\psi(r)=\sin(kr)+\tan\Delta\cos(kr)\quad\text{for large } r. \tag{3b}$$

Here Δ is a trial phase shift. Similarly, let $\Psi(r)$ denote the exact wave function [satisfying $(H-E)\Psi=0$] normalized such that

$$\Psi(0)=0, \tag{4a}$$

$$\Psi(r)=\sin(kr)+\tan\delta\cos(kr)\quad\text{for large } r, \tag{4b}$$

where δ is the exact phase shift. Finally, let $y(r)=\psi(r)-\Psi(r)$ be the error in the trial wave function. Using the fact that $(H-E)\psi=(H-E)y$, integration of the quadratic functional $L=\int_0^\infty\psi(H-E)\psi dr$ twice by parts yields

$$k\tan\delta=k\tan\Delta-L+\int_0^\infty y(H-E)y dr, \tag{5}$$

one form of the Kato identity. There are several other forms, e.g.,

$$k\cot\delta=k\cot\Delta+L-\int_0^\infty y(H-E)y dr, \tag{6}$$

where the normalization is now such that for large r

$$\psi(r)=\cos(kr)+\cot\Delta\sin(kr), \tag{7a}$$

$$\Psi(r)=\cos(kr)+\cot\delta\sin(kr). \tag{7b}$$

The various forms of the Kato identity are characterized by the fact that the left-hand side of the identity is the exact quantity of interest, while the right-hand side consists of three terms: (a) the zeroth-order estimate of the quantity of interest, i.e., the term corresponding to the left-hand side but appearing in the trial wave function; (b) the quadratic functional $L=\int\psi(H-E)\psi$, which serves as a first-order correction to the above; and (c) the (unknown) error term $\int y(H-E)y$. Using (5) as an example, we obtain the "Kato estimate" denoted by $[k\tan\delta]$ by dropping the last term, which is of second order in the error y. Thus

$$[k\tan\delta]=k\tan\Delta-\int_0^\infty\psi(H-E)\psi dr. \tag{8}$$

This yields a good answer if ψ is "close" to Ψ. In order to achieve this, we may take a flexible function ψ containing parameters [consistent with (3)] and attempt to evaluate these parameters in such a manner as to force ψ to approximate Ψ. There are many ways of doing this, for $(H-E)\Psi=0$ implies that Ψ must satisfy an infinite number of conditions, whereas in practice, we have only a finite number of parameters. This matter

* Based on the first author's (KRB) doctoral dissertation, Rensselaer Polytechnic Institute, 1966.
† Present address: University of Maine, Orono, Me.
[1] L. Spruch, in *Lectures in Theoretical Physics, Boulder, 1961* (Interscience Publishers, Inc., New York, 1962), p. 180.
[2] T. Kato, Progr. Theoret. Phys. (Kyoto) 6, 394 (1951).

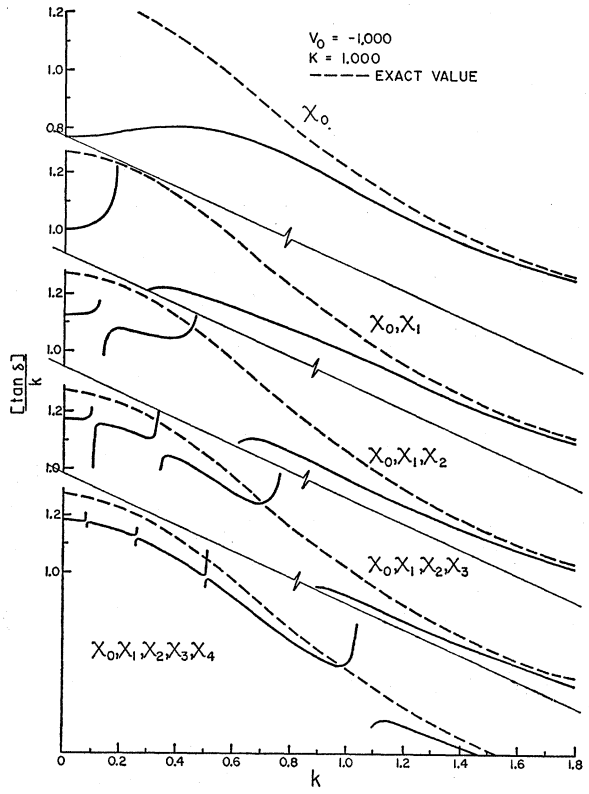

FIG. 1. Variationally determined value of $[\tan\delta]/k$ versus k for the potential $V(r)=V_0$ $(0\leqslant r\leqslant 1)$. See Eqs. (10) for the form of trial wave function. The sequence shown represents an increasing number of variational parameters from one to five. The dashed curve indicates the exact value.

has caused considerable discussion.[3,4] The Kohn[5] method uses the Kato identity not only to obtain $[k\tan\delta]$ from a *given* ψ, but also as a method of *determining* values for the parameters in ψ. Although the parameter Δ seems to play a preferential role in the Kato identity, the Kohn method treats it on an equal footing with all other parameters.

Since $[k\tan\delta]$ and $k\tan\delta$ differ by a term quadratic in the error y, the right-hand side of (8) is stationary with respect to *arbitrary* variations in ψ [consistent with (3)]. The Kohn method asks for stationarity with respect to whatever parameters appear in ψ, including Δ as one of these parameters, i.e.,

$$\frac{\partial}{\partial\lambda_i}\left(k\tan\Delta-\int_0^\infty\psi(H-E)\psi dr\right)=0, \tag{9}$$

giving N equations for the N parameters λ_i. If these equations can be solved, use of these λ_i in (8) gives the value of $[k\tan\delta]$ according to the Kohn method.

[3] T. Wu and T. Ohmura, *Quantum Theory of Scattering* (Prentice-Hall, Inc., Englewood Cliffs, N. J., 1962), pp. 57–68.
[4] F. B. Malik, Ann. Phys. (N. Y.) **20**, 464 (1962).
[5] W. Kohn, Phys. Rev. **74**, 1763 (1948).

B. Characteristics of Kohn Method

Calculations by the Kohn method exhibit certain characteristic features shown in Fig. 1. There is given the S-wave phase shift due to a square well of sufficient depth to give an appreciable phase shift but no bound state. The trial wave function, normalized according to (3), was

$$\psi(r)=\sin(kr)+\sum_{i=0}^{N}C_i\chi_i(r), \tag{10a}$$

where

$$\chi_0(r)=(1-e^{-Kr/2})\cos(kr), \tag{10b}$$

$$\chi_n(r)=r^n e^{-Kr/2},\quad n\geqslant 1. \tag{10c}$$

The term C_0 has the significance of $\tan\Delta$, and K is a nonlinear parameter which was held fixed for each calculation. Typical singularities occur, increasing in number as the number of variational parameters N is increased. Further, as N increases, the widths of the singularities decrease and the variational answer in between them converges on the correct value from below. As K is varied, the location of the singularities in Fig. 1 moves. Figure 2 shows how the phase shift varies with K in the case of a deeper square-well potential, with the energy fixed near a true resonance in the exact result. Calculations of this sort show that these singularities are an artifact of the Kohn method and the chosen normalization, and have nothing to do with any true resonances. All of the features shown here were first noticed in calculations for a more realistic problem, electron-hydrogen scattering, by Schwartz.[6]

Figure 3 shows a Kohn-method calculation for the S-wave phase shift due to a square-well potential, this time using a different normalization. The quantity considered was the logarithmic derivative of the wave function at a fixed value $r=R$. The form of the Kato identity which was used is

$$\Gamma=\gamma+\int_0^R\psi(H-E)\psi dr-\int_0^R y(H-E)y dr, \tag{11}$$

where

$$\Gamma\equiv\left.\frac{d\Psi}{dr}\right|_{r=R},\quad \gamma\equiv\left.\frac{d\psi}{dr}\right|_{r=R},$$

and where the normalization is

$$\psi(0)=0,\quad \psi(R)=1; \tag{12a}$$

$$\Psi(0)=0,\quad \Psi(R)=1. \tag{12b}$$

As before, dropping the error term gives the Kato estimate $[\Gamma]$. If R is greater than the range of the potential, the corresponding estimate of the phase shift is given by

$$[\delta]=\cot^{-1}([\Gamma]/k)-kR. \tag{13}$$

[6] C. Schwartz, Ann. Phys. (N. Y.) **16**, 36 (1961).

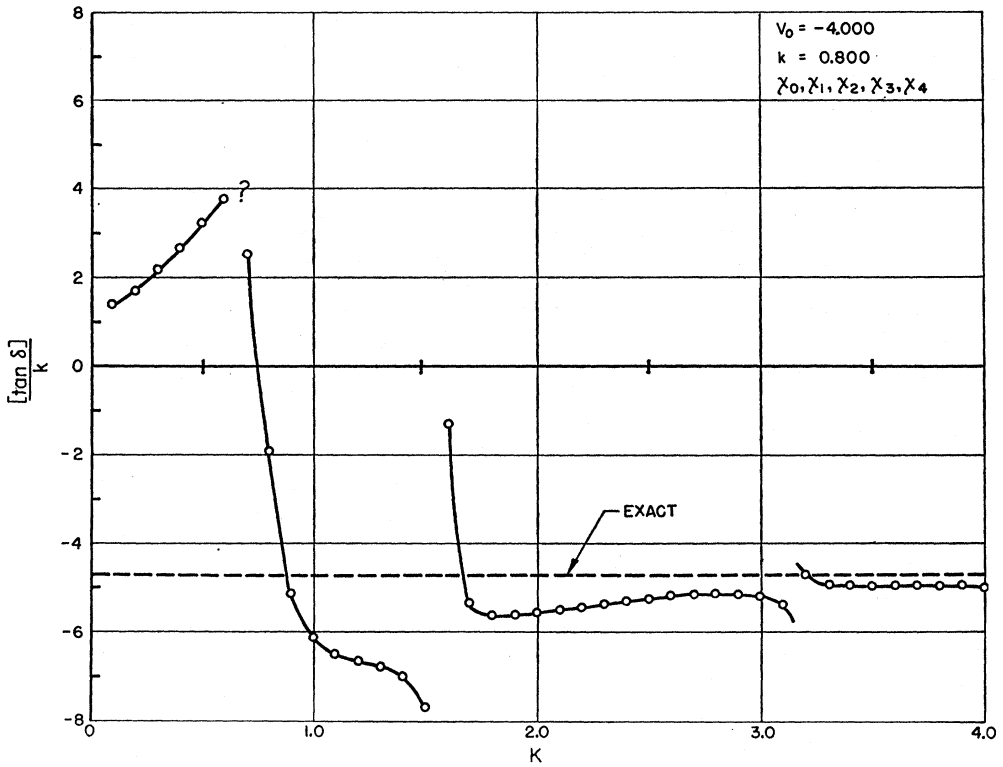

FIG. 2. Variationally determined value of $[\tan\delta]/k$ versus nonlinear parameter K. The trial function is given in Eqs. (10). For each K, $[\tan\delta]/k$ is stationary with respect to C_i $(i=0, \cdots, 4)$. The potential is $V(r)=V_0$ $(0\leqslant r\leqslant 1)$.

The trial wave function was taken as

$$\psi(r)=r/R+\sum_{i=1}^{N} C_i\chi_i(r), \tag{14a}$$

where

$$\chi_n(r)=r^n(r-R). \tag{14b}$$

Figure 3 shows the approximate phase shift $[\delta]$ (mod π). With this normalization there are *no* extraneous singularities. As N is increased, the approximate absolute phase shift approaches the exact value monotonically from below for all k.

Our aim is to analyze the Kohn method so as to understand the characteristic behavior of Figs. 1–3 and to show when the method produces a bound on the scattering parameter of interest.

II. ABSTRACT FORMULATION OF KOHN METHOD

Rather than start with the concept of a trial wave function and a method of obtaining an answer from this trial function (e.g., the Kato identity), we shall start by considering the extremization of a quadratic form whose coefficients represent "matrix elements" of $H-E$. It will be shown that the stationary value of this quadratic form has the desired significance (e.g., tangent of the phase shift); its exact significance will depend upon a concept to be introduced called the *class* of a set of functions.

Thus the emphasis shifts from that of a trial wave function to one of probing the Hamiltonian via its

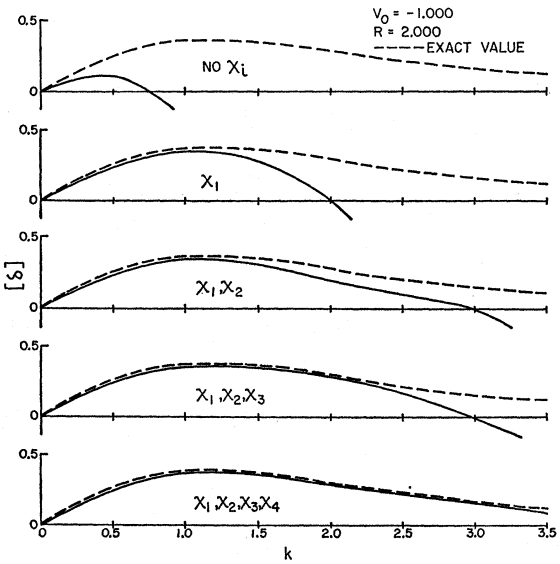

FIG. 3. Variationally determined value of $[\delta]$ versus k. The trial function is given in Eqs. (14). The sequence represents an increasing number of variational parameters from zero to four. The potential is $V(r)=V_0$ $(0\leqslant r\leqslant 1)$.

matrix elements. Although the results of this formalism will be equivalent to the Kohn method, it is hoped that the following presentation will lead to a clearer understanding of its properties.

A. Definition of $[q]$

In what follows, (f,g) will be used to denote the real inner product of two vectors. This satisfies $(f,g)=(g,f)$ as well as the usual linear relationships over the field of real numbers. In terms of this inner product, two vectors f and g will be said to belong to the same *class* if $(f,(H-E)g)=(g,(H-E)f)$, where the operator H is the "Hamiltonian" and the real number E is the "energy." A set of vectors $\{v_1,v_2,\cdots\}$ will be said to belong to one class if all pairs of its elements belong to the same class, i.e., if

$$(v_i,(H-E)v_j)=(v_j,(H-E)v_i) \quad \text{for all } i,\ j. \tag{15}$$

This notion of class is reflexive, symmetric, but not transitive. It is, however, "linear" in the sense that if f and g are any two linear combinations of vectors which belong to one class, then f and g belong to the same class. If the vectors under consideration represent functions of position, the question of "boundary conditions," "asymptotic form at large distances," etc., of these functions, as well as the actual nature of the Hamiltonian (local or nonlocal operator, number of independent variables, etc.), need not be explicitly mentioned here. They are of course implicit in this concept of class.

We start by considering the quadratic form

$$q(C_1,C_2,\cdots,C_N)=L_{ij}C_iC_j+2R_iC_i+T \tag{16}$$

(summation implied on all repeated indices). Here the coefficients L_{ij}, R_i, and T are generated from

the operator $H-E$, (17a)

a set of N (linearly independent) vectors $\{\chi_1,\chi_2,\cdots,\chi_N\}$ belonging to one class, (17b)

and

an additional vector ϕ *not* of the same class as the $\{\chi_i\}$, i.e., the set $\{\chi_1,\cdots,\chi_N,\phi\}$ does not belong to one class, (17c)

as follows:

$$L_{ij}=(\chi_i,(H-E)\chi_j)=L_{ji}, \tag{18a}$$

$$R_i=(\chi_i,(H-E)\phi), \tag{18b}$$

$$T=(\phi,(H-E)\phi). \tag{18c}$$

Although, in general,

$$(\chi_i,(H-E)\phi)-(\phi,(H-E)\chi_i)\neq 0, \tag{19}$$

the R_i are all that will be required.

Let $[q]$ denote the stationary value of q. It then is the value of q subject to the N conditions

$$\partial q/\partial C_i=0, \quad i=1,2,\cdots,N. \tag{20}$$

These variational equations lead to N simultaneous inhomogeneous equations for the N coefficients C_i:

$$L_{ij}C_j+R_i=0, \quad i=1,2,\cdots,N. \tag{21}$$

Substituting into (16), one has

$$[q]=C_iR_i+T, \tag{22}$$

provided that the C_i are determined by (21).

With the variational principle expressed in this manner, all the C_i are treated on an equal footing. The χ_i appear *symmetrically* in all the formulas; only the vector ϕ is treated differently from the others. Although the Kohn method as it is usually applied to the various forms of the Kato identity is equivalent to this formulation (when it employs linear parameters), it does not *manifestly* exhibit the above symmetry. For example, if a trial wave function of the form (10) is used, the term $\int_0^\infty \sin(kr)(H-E)C_0\chi_0 dr$ may be integrated by parts twice to give a system equivalent to (18)–(22), where $\phi=\sin(kr)$ and $C_0=\tan\Delta$.

B. Approximate Wave Function

We may associate with the solution of (21) not only the value $[q]$, but also an approximate "wave function" ψ given by

$$\psi=\phi+C_i\chi_i. \tag{23}$$

It is straightforward to verify that ψ as so defined is invariant under the following transformations applied to the variational recipe: the replacement of the set $\{\chi_i\}$ by another set $\{\chi_i'\}$, whose elements are linear combinations of the χ_i, i.e.,

$$\chi_i'=A_{ij}\chi_j, \quad \text{with} \quad \det(A_{ij})\neq 0; \tag{24a}$$

the addition to ϕ of any linear combination of the χ_i, i.e.,

$$\phi'=\phi+B_i\chi_i. \tag{24b}$$

An alternative way to regard ψ is to seek an approximate solution to $(H-E)\psi=0$, with $\psi=\phi+G_j\chi_j$. A natural thing to do in order to determine the N coefficients G_j is to demand that the N inner products of $(H-E)\psi$ with each of the χ_i vanish: $(\chi_i,(H-E)\psi)=0$, $i=1,2,\cdots,N$. But this leads to $L_{ij}G_j+R_i=0$, that is, back to (21). Thus, in order to determine the C_i, the two sets of conditions

$$\partial q/\partial C_i=0, \quad i=1,2,\cdots,N \tag{25a}$$

$$(\chi_i,(H-E)\psi)=0, \quad i=1,2,\cdots,N \tag{25b}$$

are equivalent.

The stationary value $[q]$ may be expressed compactly in terms of ψ. From (22) we have

$$[q]=C_iR_i+T=C_i(\chi_i,(H-E)\phi)+(\phi,(H-E)\phi),$$

$$[q]=(\psi,(H-E)\phi). \tag{26}$$

This simple relationship results from having determined the C_i variationally as well as from the inclusion of the

term T in q; it thus serves as a justification for this term, since T does not enter directly in the variational equations (20).

C. Definition of Q, Expressions for the Error in $[q]$

Let $\{\eta_i\}$ be a *complete*[7] set of vectors of the same class as the $\{\chi_i\}$, and let

$$\Psi=\phi+G_i\eta_i, \tag{27}$$

where the G_i are determined using the variational principle (20). It is assumed that these equations, generally infinite in number, possess a (unique) solution. Ψ will be called the "exact" wave function. The corresponding value of q, denoted by Q, will be called the "exact" value of q. From (26) we have

$$[q]=(\psi,(H-E)\phi), \quad \text{generated from the } \{\chi_i\}, \tag{28a}$$

$$Q=(\Psi,(H-E)\phi), \quad \text{generated from the } \{\eta_i\}. \tag{28b}$$

Introduce the "difference" wave function

$$y=\psi-\Psi=C_i\chi_i-G_i\eta_i, \tag{29}$$

and note that y is *some* linear combination of the η_i. We have from (28)

$$\begin{aligned}[q]-Q&=(\psi,(H-E)\phi)-(\Psi,(H-E)\phi)\\&=(y,(H-E)\phi)\\&=(y,(H-E)\Psi)-(y,(H-E)D_i\eta_i).\end{aligned}$$

But $(y,(H-E)\Psi)=0$, since y is a linear combination of the η_i and $(\eta_i,(H-E)\Psi)=0$ by (25b). Then

$$\begin{aligned}[q]-Q&=-(y,(H-E)D_i\eta_i)=-(D_i\eta_i,(H-E)y)\\&=(C_i\chi_i-D_i\eta_i-C_i\chi_i,(H-E)y)\\&=(y,(H-E)y)-C_i(\chi_i,(H-E)y).\end{aligned}$$

The last term vanishes, since

$$(\chi_i,(H-E)y)=(\chi_i,(H-E)\psi)-(\chi_i,(H-E)\Psi),$$

and both of these vanish by (25b). Thus

$$[q]-Q=(y,(H-E)y). \tag{30}$$

The error in $[q]$ is the "expectation value" of $H-E$ for the (unnormalized) vector y. See Appendix A for a somewhat more general formula and a comparison with the Kato identity.

It is of interest to compare the value $[q]_N$ associated with the N vectors $\{\chi_i\}$ with the value $[q]_M$ obtained using a larger set of $M>N$ vectors $\{\zeta_i\}$.[8] Here the set $\{\zeta_i\}$ is not necessarily complete, but does contain the $\{\chi_i\}$ in the sense that $\chi_i=B_{ij}\zeta_j$ for some coefficients B_{ij}. Using exactly the same reasoning that led to (30), one can show that

$$[q]_N-[q]_M=(y_{NM},(H-E)y_{NM}), \tag{31}$$

where $y_{NM}=\psi_N-\psi_M$. Thus the variational method not only gives the form (30) for the error in $[q]$ when compared with its *exact* value, but also measures its difference upon enlargement of the set of vectors $\{\chi_i\}$.

It may easily be shown that $[q]$ is invariant under (24a). Transformation (24b), however, may change the numerical value of $[q]$. Since the various wave functions ψ_N, ψ_M, and Ψ are invariant under (24b), it follows that all the difference wave functions are similarly unaffected. Therefore, from (30) and (31), any convergence or one-sided properties of $[q]-Q$ or $[q]_N-[q]_M$ are preserved if ϕ is replaced by $\phi+B_i\chi_i$.

D. Physical Significance of Q

In addition to H and E, the value Q depends upon

$$\text{the definition of inner product } (f,g), \tag{32a}$$

$$\text{the class of the } \{\eta_i\}, \tag{32b}$$

$$\text{the vector } \phi, \tag{32c}$$

and especially upon the relation of these to each other. One method of interpreting Q is to assume that $(H-E)\Psi=0$ (this is not necessarily a consequence of the definition of Ψ) and to then write (28b) as

$$Q=(\Psi,(H-E)\phi)-(\phi,(H-E)\Psi). \tag{33}$$

In the usual case, where inner product means integration and where H involves a second derivative operator, (33) represents "surface terms" only.

We can now see why the prior condition (17c) on ϕ was invoked, for if it were of the same class as the $\{\eta_i\}$, then (33) reduces to the trivial identity $Q=0$. In order to use (33), what is needed are the difference terms (19).

We now present two concrete examples of the determination of the physical significance of Q; these examples are like the ones discussed in the Introduction. In both of these, the Hamiltonian is the usual S-wave operator for the reduced radial wave function: $H=-d^2/dr^2+V(r)$, with $V(r)\to 0$ for $r\to\infty$. The boundary condition at the origin is $\Psi(0)=0$. In each case we must specify the three quantities (32).

Normalization number one:

Let (a) $(f,g)=\int_0^\infty f(r)g(r)dr$, (b) $\chi_i(0)=0$, $\chi_i(r)$ proportional to $\cos(kr)$ for large r, and (c) $\phi=\sin(kr)$, where $k=\sqrt{E}$. It is easily shown that the set $\{\chi_i\}$ does belong to one class. In *this* case, the asymptotic form of Ψ for large r is $\Psi=\sin(kr)+\tan\delta\cos(kr)$, where by definition δ is the phase shift. From (33) integration by parts twice yields

$$Q=-k\tan\delta \tag{34}$$

for the physical significance of Q in this case.

Normalization number two:

Let (a) $(f,g)=\int_0^R f(r)g(r)dr$, (b) $\chi_i(0)=0$, $\chi_i(R)=0$, and (c) $\phi=r/R$, where R is some fixed value of r beyond

[7] Here completeness is used in the following sense: $\{\eta_i\}$ is a maximal linearly independent set of vectors belonging to one class such that each χ_i is some linear combination of the η_j. No claim is made as to the uniqueness of this set, nor is uniqueness necessary in this development.

[8] It is understood, of course, that each ζ_i is some linear combination of the η_j.

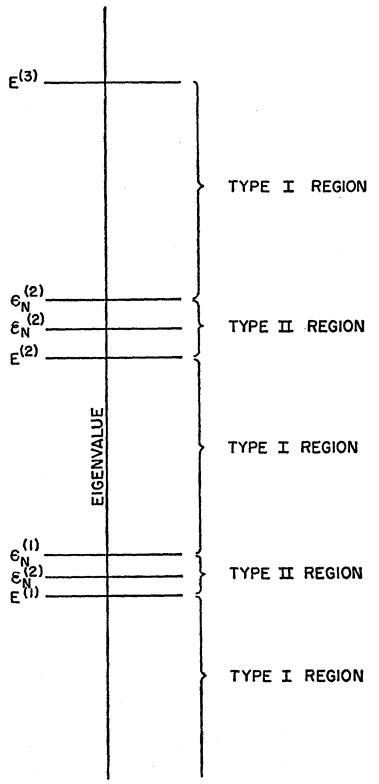

FIG. 4. Illustration of the type-I and type-II regions.

the range of the potential. Once again, it is easy to show that the set $\{\chi_i\}$ does belong to one class. In *this* case, $\Psi(R)=1$. Integration of (33) twice by parts yields

$$Q=\Gamma-1/R \tag{35a}$$

for the physical significance of Q in this case. Here Γ is the logarithmic derivative of Ψ at $r=R$. The phase shift in terms of this Q is then

$$\delta=\cot^{-1}\left(\frac{Q+1/R}{k}\right)-kR \tag{35b}$$

from (13). Of course, this only has meaning mod π.

III. MONOTONICITY RELATIONSHIPS

A. Sign of Error Term $(y,(H-E)y)$

Rosenberg, Spruch, and O'Malley[9] derived a relation that is useful when the number of rows and columns of a real symmetric matrix is increased by one. In our terminology it may be stated as follows: Let (a) $\{v_1,v_2,\cdots,v_K,z\}$ be a set of $K+1$ linearly independent normalizable vectors belonging to one class, (b) $\epsilon_K{}^{(i)}-E$ denote the K eigenvalues (ordered numerically) of $H-E$ when projected onto the K-dimensional space spanned by $\{v_1,v_2,\cdots,v_K\}$, and (c) $\mathcal{E}_K{}^{(i)}-E$ denote the $K+1$ eigenvalues (ordered numerically) of $H-E$ when projected onto the $(K+1)$-dimensional space spanned by $\{v_1,v_2,\cdots,v_K,z\}$. Condition (a) forces the matrix representation of $H-E$ (for this subspace) to be real and symmetric; hence the eigenvalues in (b) and (c) are real. The theorem (based on the fact that the determinant of a matrix is invariant under an orthogonal transformation) is then

$$(z,(H-E)z)=\sum_{i=1}^{K}\Big|\sum_{j=1}^{K}A_{ij}(v_j,(H-E)z)\Big|^2/(\epsilon_K{}^{(i)}-E)$$
$$+\frac{1}{|\lambda|^2}\prod_{i=1}^{K+1}(\mathcal{E}_K{}^{(i)}-E)/\prod_{i=1}^{K}(\epsilon_K{}^{(i)}-E). \tag{36}$$

Here A_{ij} is a matrix which transforms $\{v_i\}$ into an orthonormal set such that the K-by-K matrix of $H-E$ formed using this set is diagonal. λ is defined by

$$\frac{1}{|\lambda|^2}=(z,z)-\sum_{i,j=1}^{K}|(A_{ij}v_j,z)|^2, \tag{37}$$

the significance of λ being that the component of λz orthogonal to the v_i have unit magnitude.

To apply this theorem to the variational principle of Sec. II, choose $K=N$, $v_i=\chi_i$, and $z=y$. Note that y is of the same class as the $\{\chi_i\}$. Provided that the other conditions of the theorem are satisfied, (36) gives a relation for $(y,(H-E)y)$ with the summation term vanishing because $(\chi_j,(H-E)y)=(\chi_j,(H-E)\psi)-(\chi_j,(H-E)\Psi)$, and both of these vanish by (25b). This simplification is a direct consequence of the variational (Kohn) method. Equation (36) then reduces to

$$(y,(H-E)y)=\frac{1}{|\lambda|^2}\prod_{i=1}^{N+1}(\mathcal{E}_N{}^{(i)}-E)/\prod_{i=1}^{N}(\epsilon_N{}^{(i)}-E) \tag{38}$$

for the error term in $[q]-Q$. We therefore have a bound on Q if it can be shown that the right-hand side of (38) has a definite sign. This will be seen to be possible if (a) the eigenvalues of $H-E$, when projected onto the space spanned by the complete set of vectors $\{\eta_i\}$ of the same class as the $\{\chi_i\}$, are discrete and bounded from below, and (b) the set $\{\chi_i\}$ is reasonably complete (to be precisely defined below).

Let $E^{(i)}-E$ denote the (exact) eigenvalues (ordered numerically) of $H-E$ when projected onto the space spanned by the $\{\eta_i\}$. If they are discrete and bounded from below, the Hylleraas-Undheim theorem[10] states that

$$E^{(i)}\leqslant\mathcal{E}_N{}^{(i)}\leqslant\epsilon_N{}^{(i)},\quad i=1,2,\cdots,N \tag{39a}$$

and

$$\epsilon_N{}^{(N)}\leqslant\mathcal{E}_N{}^{(N+1)}. \tag{39b}$$

Further, if the $\{\chi_i\}$ were reasonably complete, we would expect that the $\epsilon_N{}^{(i)}$ give a good approximation to the $E^{(i)}$, at least for the lower eigenvalues. We are thus led to make the following definition. The energy E will be

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

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

said to be in a *type-I region* if

the number of $\epsilon_N^{(i)}$ less than E is equal to the number of $E^{(i)}$ less than E. (40)

Otherwise, E will be said to be in a *type-II region*. These definitions are illustrated in Fig. 4. In view of the inequalities (39), (38) gives

$$(y,(H-E)y)>0, \quad \text{in a type-I region.} \tag{41}$$

Exactly the same reasoning may be applied with $y_{NM}(M>N)$ instead of y in (36) because (a) y_{NM} is of the same class as the $\{\chi_i\}$ and (b) $(\chi_j,(H-E)y_{NM}) = (\chi_j,(H-E)(\psi_N-\psi_M))=0$ $(j=1, 2, \cdots, N)$, so that the summation term in (36) again vanishes. Thus

$$(y_{NM},(H-E)y_{NM})>0, \quad \text{in a type-I region.} \tag{42}$$

Rosenberg's[9] original idea was (a) to take K to be the number of $E^{(i)}$ less than E (this number would have to be determined by other means), and (b) to find a set of K auxiliary vectors $\{v_i\}$ (not identified with the $\{\chi_i\}$) which were "good enough" so that the $\epsilon_N^{(i)}$ would all be less than E. The summation term in (36) would have to be evaluated (it is negative), and an inequality for $(y,(H-E)y)$ would then result since the last term in (36) is again positive. In the method presented here (namely, the identification of $\{v_i\}$ with $\{\chi_i\}$), the summation term in (36) vanishes identically and a stronger inequality is obtained. This simplification was pointed out by Ohmura[11] and later by Rosenberg *et al.*[12] An additional benefit of this is that the computational labor is reduced, because the evaluation of this summation term involves a matrix diagonalization, whereas the Kohn method requires only a matrix inversion (solution of $L_{ij}C_j+R_i=0$). Of course, one still must somehow know that the set $\{\chi_i\}$ is reasonably complete in the sense that the type-I requirement (40) is satisfied.

One further point remains. The matrix theorem (36) can be applied only if y is linearly independent of the χ_i. If y were a linear combination of the χ_i, then $\psi=\Psi$ is *a* solution of the extremization equations (20). Assuming that $\det(L_{ij})\neq 0$, then it is *the* solution. Similar remarks apply to the relative error term y_{NM}. Therefore (41) and (42) may be strengthened to say that if E lies in a type-I region, $[q]$ gives a bound

$$[q]-Q\geqslant 0, \quad \text{with equality } \textit{if and only if } \psi=\Psi, \tag{43}$$

and the monotonicity relation for $M>N$ is

$$[q]_N-[q]_M\geqslant 0, \quad \text{with equality } \textit{if and only if } \psi_N=\psi_M. \tag{44}$$

The case of equality in these formulas may be regarded as the limit of (36) as $\lambda\to\infty$, as follows from the definition (37) of λ.

[11] T. Ohmura, Phys. Rev. **124**, 130 (1961).
[12] L. Rosenberg and L. Spruch, Phys. Rev. **125**, 1407 (1962).

B. Optimum Determination of Additional Parameters

Suppose that for some N the set $\{\chi_1,\cdots,\chi_N\}$ is good enough so that E lies in a type-I region. Consider enlarging this set to $\{\chi_1,\cdots,\chi_M\}$ for $M>N$. Using this set, let $C_1, \cdots, C_N$ be determined variationally, with $C_{N+1}, \cdots, C_M$ left as free parameters; $C_1, \cdots, C_N$ then become functions of $C_{N+1}, \cdots, C_M$. Denote by $[q]_N'$ the corresponding value of q and by ψ' the corresponding wave function ψ; these being functions of $C_{N+1}, \cdots, C_M$. Formula (A3) shows that $[q]_N'-Q=(y',(H-E)y')$, where $y'=\psi'-\Psi$, no matter how $C_{N+1}, \cdots, C_M$ are determined. To apply the matrix theorem (36) in this case, choose $K=N$ (not M), $v_i=\chi_i$ $(i=1, \cdots, N)$, $z=y'$. The summation term in (36) again vanishes because the first N C_i were determined variationally. Hence $[q]_N' -Q\geqslant 0$, with equality if and only if $\psi'=\Psi$. The best way to determine the additional parameters $C_{N+1}, \cdots, C_M$ is therefore such as to minimize $[q]_N'$. This is equivalent to demanding that $\partial q/\partial C_i=0$ $(i=N+1, \cdots, M)$, since q is already stationary with respect to the first N C_i. There can be no extraneous minima since q is only quadratic in the C_i. Thus we are led back to the original variational recipe for *all* the C_i. In addition, if the χ_i themselves contain other parameters [say, $\chi_i=\chi_i(\mu_j)$; the μ_j may be nonlinear parameters], then (43) shows that these too should be determined variationally. One would have to bear in mind that (i) μ_j being nonlinear may result in q having several extrema, and (ii) the condition (40) that E lie in a type-I region might be violated as the μ_j are varied (because the $\epsilon_N^{(i)}$ will be functions of the μ_j). It is understood that the μ_j are not allowed to change the class of the $\{\chi_i\}$.

Another interesting question concerns the determination of any additional parameters that appear in ϕ. That a variational determination is not always the correct procedure may be seen by considering a multiplicative parameter: $\phi\to C\phi$. Extremizing q in this case yields $[q]=Q=0$. The trouble here is that as we vary q (via C), we are also changing Q (for ϕ in effect controls the "normalization" of the wave functions). We must therefore deal with a more restrictive class of parameters, though in practical calculations this restriction is usually satisfied. We want ϕ to contain parameters which do not affect the value of $Q=(\Psi,(H-E)\phi) -(\phi,(H-E)\Psi)$. Since these represent "surface terms," parameters which do not affect ϕ "asymptotically" are permissible. From (43) it is seen that these too should be determined variationally if E lies in a type-I region. The remark (i) above applies again in the case of nonlinear parameters; the remark (ii) does not apply, since ϕ has nothing to do with the $\epsilon_N^{(i)}$.

In summary, provided that E lies in a type-I region the Kohn principle of stationarity becomes a true minimum principle: $[q]>Q$. The addition of more parameters can only *improve* the answer if these parameters are determined variationally; conversely, a variational de-

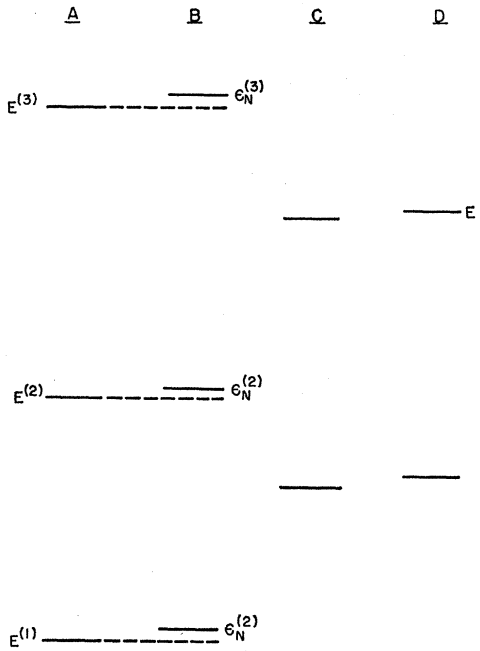

FIG. 5. Energy eigenvalue spectra for the various boundary conditions.

termination of these additional parameters is the optimum determination.

C. Energy Dependence of Monotonicity Relations

Figure 4 can be quite deceptive. It seems to indicate that if the set $\{\chi_i\}$ is good enough, the type-I regions are wide and are interlaced with narrow type-II regions. Thus the "probability" of E being in a type-I region is large. The working procedure would then be to use enough vectors χ_i so that the situation presumably resembled the figure. A statistical argument would then justify that E was (probably) in a type-I region. This would avoid having to check explicitly that the condition (40) were satisfied (the $E^{(i)}$ are usually not known anyway).

What must not be overlooked, however, is the fact that Fig. 4 is drawn for some *fixed* energy E. In many applications the χ_i are themselves functions of the energy (they are usually parametrized by the wave number $k=\sqrt{E}$). In fact, not only the χ_i, but more importantly, even the class of the $\{\chi_i\}$, is a function of E. Therefore the exact eigenvalues $E^{(i)}$ and their approximate values $\epsilon_N{}^{(i)}$ will be functions of E. Thus there is a *correlation* between these spectra and the energy E. It may turn out that although the type-II regions are very narrow, their locations would vary with E in such a manner as to make E lie in a type-II region most of the time. This correlation of the eigenvalues with the energy makes it necessary to apply the previous theorems with care.

The same problem of correlations with energy occurs in the formal solution for $[q]$ as a function of E. For simplicity, take the $\{\chi_i\}$ to correspond to the orthonormal proper vectors of L_{ij}:

$$(\chi_i,(H-E)\chi_j)=(\epsilon_N{}^{(i)}-E)\delta_{ij}. \tag{45}$$

This can always be accomplished by a transformation of the form (24a). The solution of (21) is

$$C_i=-R_i/(\epsilon_N{}^{(i)}-E), \tag{46}$$

and the value $[q]$ as given by (22) is then

$$[q]=-\sum_{i=1}^{N}\frac{R_i^2}{\epsilon_N{}^{(i)}-E}+T. \tag{47}$$

This apparently exhibits the poles and pole strengths of $[q]$. But again, in general, the R_i, $\epsilon_N{}^{(i)}$, and T are themselves functions of E. Therefore, instead of (47), one should write

$$[q]=-\sum_{i=1}^{N}\frac{R_i^2(E)}{\epsilon_N{}^{(i)}(E)-E}+T(E), \tag{48}$$

which should not be taken literally as a Laurent expansion in terms of the energy.

IV. ANALYSIS OF EXAMPLES

A. Application to Example Number One

This example (see Sec. II D) concerns the Hamiltonian $H=-d^2/dr^2+V(r)$, with (a) $(f,g)=\int_0^\infty f(r)g(r)dr$, (b) $\chi_i(0)=0$, $\chi_i(r)$ proportional to $\cos(kr)$ for large r, and (c) $\phi=\sin(kr)$, where $k=\sqrt{E}$. One assumption will be made, so that what follows is really a model of this example, though probably quite a realistic one. This assumption is that there exists an R such that for $r\geqslant R$

$$V(r)=0 \tag{49a}$$

and

all the $\chi_i(r)$ have attained their asymptotic form. (49b)

Under this assumption all the various integrands vanish for $r>R$; the inner product may now be redefined: $(f,g)=\int_0^R f(r)g(r)dr$. The class of the set $\{\chi_i\}$ is then characterized by those functions $f(r)$, $0\leqslant r\leqslant R$, such that

$$f(0)=0,\quad \frac{1}{f}\frac{df}{dr}\bigg|_{r=R}=-k\tan(kR). \tag{50}$$

Using this new inner product, the spectrum $\{E^{(i)}\}$ of H projected onto a complete set of functions of the class specified by (50) is discrete and bounded from below (because R is finite). Since $Q=-k\tan\delta$, the results (43) and (44) become: If E lies in a type-I region,

$$[\tan\delta]\leqslant\tan\delta, \text{ with equality if and only if } \psi=\Psi, \tag{51}$$

and for $M>N$

$$[\tan\delta]_N\leqslant[\tan\delta]_M, \text{ with equality if and only if } \psi_N=\psi_M. \tag{52}$$

Here δ is the exact phase shift and $[\tan\delta]=-[q]/k$.

Of course, the question still remains concerning the probability that E lies in a type-I region. To investigate this, let β_1, β_2, and β_3 denote the logarithmic derivative (evaluated at $r=R$) of χ_i, ϕ, and Ψ, respectively, the β_i being functions of the energy. We have $\beta_1=-k\tan(kR)$, $\beta_2=k\cot(kR)$, and β_3 is unknown. A function $g(r)$ will be said to have "boundary condition β_i" if

$$g(0)=0 \quad \text{and} \quad \frac{1}{g}\frac{dg}{dr}\bigg|_{r=R}=\beta_i.$$

Figure 5, which is drawn for some fixed energy E, gives a qualitative description of the situation. In the figure, column A is the exact spectrum of H when projected onto the space spanned by a complete set of functions $\{\eta_i\}$ having boundary condition β_1 (the $E^{(i)}$); column B is the approximate spectrum obtained by projecting H onto the space spanned by the set $\{\chi_i\}$ (the $\epsilon_N{}^{(i)}$); column C is the exact spectrum of H when projected onto the space spanned by a complete set of functions having boundary condition β_2; and column D is the exact spectrum of H when projected onto the space spanned by a complete set of functions having boundary condition β_3.

If the set $\{\chi_i\}$ is "good," then the levels in column B closely approximate (from above) those in column A, at least for the lower levels. Moreover, if the phase shift δ is not too close to a resonance value $[(M+\frac{1}{2})\pi]$, then the levels in column D approximate those in column C (because then β_2 is close to β_3). In summary, so far, we may expect that

$$\text{B resembles A (from above)}, \tag{53}$$

$$\text{D resembles C}. \tag{54}$$

Now, β_1 and β_2 are "complementary" kinds of boundary conditions (this is somewhat analogous to fixed versus free boundary condition for a vibrating string). We may therefore expect that

$$\text{the levels in C are roughly midway between the levels in A}. \tag{55}$$

Now the question is: Where is the energy E in the figure? By definition of β_3, E must be precisely one of the levels in column D as shown. Thus the correlation previously mentioned between the $E^{(i)}$ and E is such as to automatically make E tend to lie roughly *midway* between some adjacent pair of $E^{(i)}$. This is even a much *better* situation than was qualitatively indicated in Fig. 4; for the $\{\chi_i\}$ has only to be good enough to approximate those $E^{(i)}$ less than E to within roughly one-half an energy-level difference in order to force E to lie in a type-I region.

On the basis of this discussion, it would seem that if $\{\chi_i\}$ is any good at all, Fig. 5 should be true for a wide range of energy, and there is then no reason to expect any singularities at all in $[q]$. We would now like to show that although $\{\chi_i\}$ may be *arbitrarily* good, there must be small regions of energy in which (53) is violated. This will be caused by the fact that A is derived from the *infinite* set $\{\eta_i\}$, while B is derived from the (perhaps very large) *finite* set $\{\chi_i\}$. To do this, we must investigate how the levels in A and B vary with the energy E.

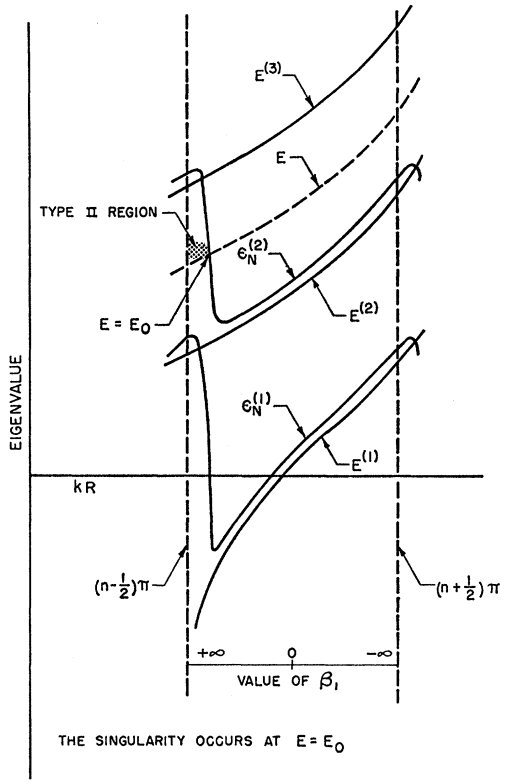

FIG. 6. Expected energy eigenvalue spectra versus the boundary condition $\beta_1=-k\tan(kR)$.

Refer to Fig. 6. As k decreases in the interval

$$(n-\tfrac{1}{2})\pi<kR<(n+\tfrac{1}{2})\pi \tag{56}$$

(with n an integer), $\beta_1=-k\tan(kR)$ increases monotonically from $-\infty$ to $+\infty$. This is a *relaxation* of the boundary condition at $r=R$, thus causing the eigenvalues $E^{(i)}$ to monotonically decrease (see Appendix B). Since $\beta_1=\pm\infty$ correspond to the same physical boundary condition at $r=R$, we must have

$$\lim_{\beta_1\to+\infty} E^{(i)}=\lim_{\beta_1\to-\infty} E^{(i-1)} \quad (i=2,3,\cdots) \tag{57a}$$

and

$$\lim_{\beta_1\to+\infty} E^{(1)}=-\infty. \tag{57b}$$

As k decreases through the interval (56), the spectrum $\{E^{(i)}\}$ thus repeats itself, each level moving down one notch, the lowest level disappearing to negative infinity. This is possible, of course, only because there are an infinite number of levels $E^{(i)}$. Bearing in mind the significance of the superscript, it follows that each $E^{(i)}$ must be a *discontinuous* function of k.

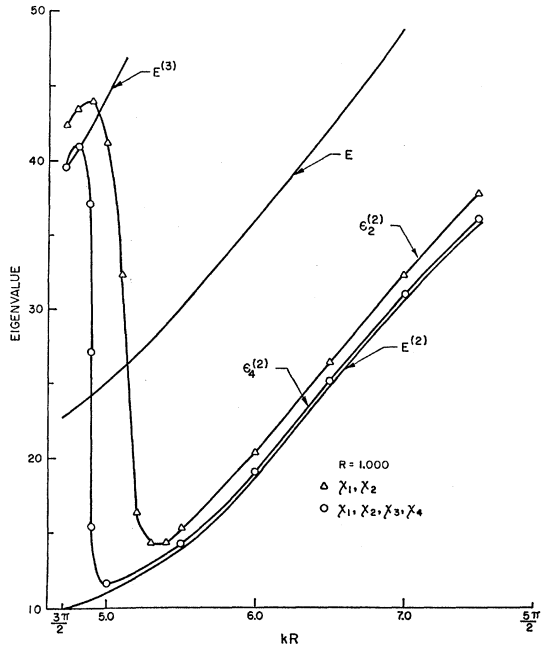

FIG. 7. Numerical illustration of the behavior sketched in Fig. 6. The plotted points are the approximate eigenvalues $\epsilon_N^{(i)}$; the solid curves are the exact eigenvalues $E^{(i)}$. Note the steep slope of the $\epsilon_N^{(i)}$ as they cross E.

However, the finite number of levels $\epsilon_N^{(i)}$ associated with $\{\chi_i\}$ must (a) obey the inequalities $\epsilon_N^{(i)} \geqslant E^{(i)}$ and (b) be *continuous* functions of k (the χ_i and their matrix elements are assumed to be continuous in k; the $\epsilon_N^{(i)}$ as roots of the finite-degree characteristic polynomial are then continuous in k). The approach of the $\epsilon_N^{(i)}$ is then one of nonuniform convergence in k: For any fixed k, the $\epsilon_N^{(i)}$ approach the $E^{(i)}$ from above monotonically with N; however, for any fixed N (no matter how large), there exists a neighborhood (near $\beta_1=+\infty$) in the interval (56) for which the $\epsilon_N^{(i)}$ fail to approximate the $E^{(i)}$ by at least a full energy-level difference $E^{(i+1)}-E^{(i)}$. The reason why this occurs near $\beta_1 \to +\infty$ can be understood by noting that the lowest exact eigenfunction [e.g., for a free particle this is proportional to $\sinh(\sqrt{-E^{(1)}}r)$, $E^{(1)} \to -\infty$] possesses very high curvature and hence is difficult to synthesize using "smooth" functions $\chi_i(r)$.

If the phase shift δ is not too large (mod π), then E lies roughly midway between two adjacent $E^{(i)}$ as shown. The criterion (40) is violated in the small shaded region in the figure, and hence type-II regions *must* occur; they are an inherent feature of the normalization (3) and the variational principle. At the right-hand edge of the shaded region, $E=E^{(j)}$ for some j, and from (48) this corresponds to a singularity in $[q]$. It is easily seen from Fig. 6 and (48) that this singularity is a simple pole with positive pole strength (residue), in agreement with the example of Fig. 1.

If in the interval (56), E happens to cross one of the $E^{(i)}$, there will be an additional singularity in $[q]$ nearby (the type-I criterion can never be met with $E=E^{(i)}$). This additional singularity corresponds to an actual resonance in the problem. This can be seen by noting that if $E=E^{(i)}$, then at that energy $\beta_1=\beta_3$, and the asymptotic phase of Ψ is then that of the χ_i, i.e., $\delta=90°$ (mod π).

A simple numerical example illustrates the behavior of the $\epsilon_N^{(i)}$ discussed above. Take the Hamiltonian as $H=-d^2/dr^2$ and the set $\{\chi_i\}$ as polynomials in r which satisfy the boundary condition $\beta_1=-k\tan(kR)$:

$$\chi_1=Ar+Br^2,$$
$$\chi_n=r^{n-1}(r-R)^2, \quad \text{for} \quad n\geqslant 2,$$

the ratio A/B being adjusted to meet the above boundary condition. R was taken to be unity, and the interval $\frac{3}{2}\pi<kR<\frac{5}{2}\pi$ was investigated, with the results shown in Fig. 7. The plotted points refer to the approximate eigenvalue $\epsilon_N^{(2)}$, with $N=2,4$, and the solid curve is the exact eigenvalue $E^{(2)}$. Even using only two functions, $\epsilon_2^{(2)}$ approximates $E^{(2)}$ quite closely, except near $k=\frac{3}{2}\pi$, as expected. Using four functions, $\epsilon_4^{(2)}$ gives a better approximation; the narrow type-II region near $k=\frac{3}{2}\pi$ has a width of less than 0.2 unit of k (about 5% of the interval investigated). Also shown is the energy ($E=k^2$) as a function of k. The over-all agreement of this example with the previous sketch (Fig. 6) is excellent.

B. Pole Strengths of Singularities of $[q]$

It is of interest to obtain expressions for and estimates of the pole strengths of $[q]$ at the "extraneous" singularities. For simplicity, we may take $\{\chi_i\}$ to satisfy

$$(\chi_i,(H-E)\chi_j)=(\epsilon_N^{(i)}-E)\delta_{ij}. \tag{45}$$

Suppose that a singularity in $[q]$ occurs at $E=E_0$. From (48) the singular part of $[q]$ is

$$+R_j^2(E)/(E-\epsilon_N^{(j)}(E)), \tag{58}$$

with $\epsilon_N^{(j)}(E_0)=E_0$. For the moment, let us neglect the dependence of $\epsilon_N^{(j)}$ upon E. Then the pole strength (residue) p of $[q]$ at this singularity is

$$p=R_j^2=|(\chi_j,(H-E)\phi)|^2, \tag{59}$$

evaluated at $E=E_0$. Because of (45), any linear combination of the χ_i may be added to ϕ without affecting this value of p. One is tempted to add $C_i\chi_i$ to obtain $\psi=\phi+C_i\chi_i$ and then argue that $(H-E)\psi$ is small because ψ is close to Ψ. But this is precisely *not* the case; at the singularity of $[q]$, ψ is infinite, whereas Ψ is finite (Ψ is infinite at an actual resonance, but these are not the singularities being discussed here). However, since we are free to add *any* linear combination $D_i\chi_i$ to ϕ, let s^2 be the smallest norm of all vectors of the form $(H-E)(\phi+D_i\chi_i)$, i.e.,

$$s^2=\min_{Di}[(H-E)(\phi+\sum_i D_i\chi_i),(H-E)(\phi+\sum_i D_i\chi_i)], \tag{60}$$

and use these minimizing D_i. Then (59) becomes

$$p=|(\chi_j,(H-E)(\phi+D_i\chi_i))|^2. \tag{61}$$

Since $(\chi_j,\chi_j)=1$,

$$p\leqslant \text{norm of } [(H-E)(\phi+D_i\chi_i)]=s^2. \tag{62}$$

Now s^2 will be small if there *exists* a linear combination $\phi+D_i\chi_i$ which is close to Ψ, even though the Kohn method does not yield this combination (i.e., the D_i are not the same as the C_i). If we reasonably assume that $\{\chi_i\}$ is rich enough to make s^2 small, we may conclude that p will be small.

Referring to Figs. 6 and 7, it is seen that near $E=E_0$, $\epsilon_N^{(j)}(E)$ is a very rapidly changing function. The above analysis must be modified to include this effect. Expanding $\epsilon_N^{(j)}(E)$ about $E=E_0$ yields

$$p\leqslant s^2(1-d\epsilon_N^{(j)}/dE|_{E=E_0})^{-1} \tag{63}$$

instead of (62). Since, at $E=E_0$, $d\epsilon_N^{(j)}/dE$ is large and negative, the factor multiplying s^2 in (63) is much smaller than unity, making the pole strength much smaller than one would otherwise expect. In the example shown in Fig. 7 with $N=4$, it turns out that $p\leqslant s^2/100$.

Thus we may expect p to be small for two reasons: (i) s^2 is small, and (ii) the correlation of the $\epsilon_N^{(j)}$ with E produces a further dimunition $[1-d\epsilon_N^{(j)}/dE|_{E=E_0}]^{-1}$. Both (i) and (ii) may be expected to become more important as N is increased.

C. Application to Example Number Two

This example (see Sec. II D) concerns the Hamiltonian $H=d^2/dr^2+V(r)$, with (a) $(f,g)=\int_0^R f(r)g(r)dr$, (b) $\chi_i(0)=\chi_i(R)=0$, $\chi_i(r)$ independent of E, and (c) $\phi=r/R$. In this case, the spectrum $\{E^{(i)}\}$ is discrete and bounded from below. Since the χ_i are not parametrized by the energy, Fig. 4 can be taken quite literally (i.e., the $E^{(i)}$ and the $\epsilon_N^{(i)}$ are not functions of E). Thus the probability of E being in a type-I region is large if $\{\chi_i\}$ is reasonably good. Both $[q]$ and Q now possess singularities with positive pole strengths, from (47).

The expression for the phase shift (mod π) in terms of Q is given by (35b). To define an *absolute* phase shift that is continuous in the energy, let

$$[\delta]=\cot^{-1}\left(\frac{[q]+1/R}{k}\right)-kR+l\pi, \tag{64a}$$

$$\delta=\cot^{-1}\left(\frac{Q+1/R}{k}\right)-kR+L\pi. \tag{64b}$$

Here L is the number of $E^{(i)}$ less than E, l is the number of $\epsilon_N^{(i)}$ less than E, and $0<\cot^{-1}<\pi$. Since (i) $[q]\geqslant Q$ and $l=L$ in a type-I region and (ii) $l<L$ in a type-II region, the absolute phase shifts (64) satisfy

$$[\delta]\leqslant\delta, \tag{65a}$$

and for $M>N$,

$$[\delta]_N\leqslant[\delta]_M. \tag{65b}$$

Equations (65) are true for *all* energies. If E lies in a type-I region, they may be strengthened to say that there is equality if and only if $\psi=\Psi$ or $\psi_N=\psi_M$, respectively, and that the amount of inequality is less than π.

V. SUMMARY AND CONCLUSIONS

The Kohn method of calculating scattering quantities (when it employs linear parameters) was shown to be derivable from a more abstract formulation based on the extremization of a quadratic form

$$q(C_1,\cdots,C_N)=\sum_{i,j=1}^{N}L_{ij}C_iC_j+2\sum_{i=1}^{N}R_iC_i+T,$$

whose coefficients L_{ij}, R_i, T represent certain matrix elements of $H-E$, formed from a set of vectors $\{\chi_1,\cdots,\chi_N,\phi\}$. The extremizing C_i define an approximate wave function $\psi=\phi+C_i\chi_i$, and the stationary value of q is then given by $[q]=(\psi,(H-E)\phi)$. If certain conditions are satisfied, $[q]$ approximates an exact value Q in the following manner. The energy axis may be divided into interlacing type-I and type-II regions such that in the type-I regions $[q]$ approaches Q monotonically from above: $[q]_N\geqslant[q]_M\geqslant Q$ for $M>N$. In addition, as the number of variational parameters N is increased, the type-II regions become more narrow. With the exception of these narrow type-II regions, the Kohn principle of stationarity then becomes a true minimum principle: $[q]\geqslant Q$. The implications of these results were investigated for two specific normalizations.

Using a trial function of the form (3) to evaluate the stationary value (8) results in an approximate tangent of the phase shift $[\tan\delta]$. If a certain assumption (49) concerning the range of the trial function and the potential is satisfied, then in the type-I regions: $[\tan\delta]_N\leqslant[\tan\delta]_M\leqslant\tan\delta$ for $M>N$. At the left edge of each type-I region there will be an extraneous singularity in $[\tan\delta]$. These have nothing to do with any actual resonances, but rather are a consequence of the Kohn method and this normalization. The pole strength of each such singularity is less than the product of two terms, each of which is small and may be expected to get smaller as N is increased. In practice, these singularities may be so narrow for large N as to be ignorable (see Fig. 1, with $N=4$).

Using an energy-independent trial function of the form (12) to evaluate the stationary value (11) results in an approximate logarithmic derivative $[\Gamma]$ of the wave function at $r=R$. In the type-I regions we have $[\Gamma]_N\geqslant[\Gamma]_M\geqslant\Gamma$ for $M>N$. It is possible to define an absolute phase shift (64) from this logarithmic derivative. This absolute phase shift obeys $[\delta]_N\leqslant[\delta]_M\leqslant\delta$ for $M>N$, this being valid for all energies. However, in the type-II regions the amount of inequality may be large.

In the type-I regions the amount of such inequality is less then π, so that the result becomes more useful here.

In conclusion, although the Kohn method involves the extremization of a nondefinite quadratic form, evidence was presented to show why the variationally determined scattering parameter $[q]$ will approximate the corresponding exact value Q. In general, the expected behavior is the following. As the number of linear variational parameters N is increased for a fixed value of $k=\sqrt{E}$, eventually $[q]$ provides a bound: $[q]\geqslant Q$. As N is further increased, $[q]$ will monotonically approach Q from above. For a fixed value of N, depending upon the particular normalization employed, there may or may not be extraneous singularities in $[q]$ as a function of k. If present, these singularities will be narrow, their widths decreasing as N is increased.

APPENDIX A: KATO IDENTITY

In Eq. (30), $[q]-Q=(y,(H-E)y)$, $[q]$ was the stationary value of q and $y=\psi-\Psi$ was the difference between the two variationally determined wave functions. This expression is true quite generally even if ψ is *not* determined variationally. To show this, let $\psi=\phi+C_i\chi_i$, where the C_i are *arbitrary*. The value of the (unextremized) quadratic form q is

$$\begin{aligned}q&=L_{ij}C_iC_j+2R_iC_i+T\\&=(C_i\chi_i,(H-E)C_j\chi_j)+2(C_i\chi_i,(H-E)\phi)\\&\qquad+(\phi,(H-E)\phi)\\&=(\psi,(H-E)\psi)+(\psi,(H-E)\phi)-(\phi,(H-E)\psi).\end{aligned}\tag{A1}$$

Similarly,

$$\begin{aligned}Q&=(\Psi,(H-E)\Psi)+(\Psi,(H-E)\phi)\\&\qquad-(\phi,(H-E)\Psi).\end{aligned}\tag{A2}$$

Substituting $y=\psi-\Psi$, where Ψ is the (variationally determined) exact wave function, into (A1),

$$\begin{aligned}q&=(\Psi+y,(H-E)(\Psi+y))\\&\qquad+(\Psi+y,(H-E)\phi)-(\phi,(H-E)(\Psi+y))\\&=(\Psi,(H-E)\Psi)+(\Psi,(H-E)\phi)\\&\quad-(\phi,(H-E)\Psi)+(y,(H-E)y)\\&\quad+2(y,(H-E)\Psi)-(y,(H-E)\Psi)+(\Psi,(H-E)y)\\&\qquad+(y,(H-E)\phi)-(\phi,(H-E)y).\end{aligned}$$

The last four terms combine to give $((\Psi-\phi),(H-E)y)-(y,(H-E)(\Psi-\phi))$, which vanishes, since y and $(\Psi-\phi)$ belong to the same class. The term $2(y,(H-E)\Psi)$ vanishes because Ψ was determined variationally. Thus, using (A2),

$$q=Q+(y,(H-E)y),\tag{A3}$$

which was to be shown. Note that $(H-E)\Psi=0$ was not assumed.

Substitution of (A3) into (A1) yields

$$\begin{aligned}Q&=(\psi,(H-E)\phi)-(\phi,(H-E)\psi)\\&\qquad+(\psi,(H-E)\psi)-(y,(H-E)y).\end{aligned}\tag{A4}$$

This is the analog of the Kato identity. The first two terms, which combine to form a surface term, correspond to $-k\tan\Delta$ in Eq. (5); the next term corresponds to L. $(y,(H-E)y)$ is the usual quadratic error term.

APPENDIX B: PROOF THAT $dE/d\beta$ IS NEGATIVE

Let $f(r)$, $0\leqslant r\leqslant R$, be an eigenfunction (with eigenvalue E) of the operator $H=-d^2/dr^2+V(r)$. The boundary conditions for f are

$$f(0)=0 \quad\text{and}\quad \frac{1}{f}\frac{df}{dr}\bigg|_{r=R}=\beta.$$

We wish to show that E is a monotonically decreasing function of β. In the region $-\infty<\beta<+\infty$, $f(R)\neq 0$. Thus f may be normalized such that $f(R)=1$. Treating f as a function of both r and E, we have, using integration by parts twice,

$$\begin{aligned}0&=\frac{d}{dE}\int_0^R f(H-E)f\,dr\\&=\int_0^R f(H-E)\frac{\partial f}{\partial E}dr-\int_0^R f^2dr\\&=\left(\frac{\partial f}{\partial r}\frac{\partial f}{\partial E}-f\frac{\partial^2 f}{\partial E\partial r}\right)\bigg|_0^R-\int_0^R f^2dr\\&=-\frac{d\beta}{dE}-\int_0^R f^2dr,\end{aligned}$$

$$\frac{dE}{d\beta}=-\left(1\Big/\int_0^R f^2dr\right)<0.$$