ANNALS OF PHYSICS: **42,** 367–374 (1967)

Generalized Eigenvalue Problems*

CHARLES SCHWARTZ

Department of Physics, University of California, Berkeley, California 94720

We examine the equation $L(\lambda)\psi = 0$, where L is a linear operator of the usual sort, except that it is not necessarily self-adjoint, and its dependence on the eigenvalue parameter λ is not necessarily linear. Variational principles, normalization of eigenfunctions, resolution of the identity and operator inversion are some of the aspects discussed.

We are most familiar with the eigenvalue problem in the form

$$(H - \lambda)\psi = 0 \tag{1}$$

with the self-adjoint operator $H = H^\dagger$. Among the well-known results are the following.

(i) Equivalent to (1) is the variational principle

$$\delta J[\tilde{\psi}] = 0,$$

where

$$J[\tilde{\psi}] = \langle \tilde{\psi} \,|\, H - \lambda \,|\, \tilde{\psi} \rangle \tag{2}$$

and $\tilde{\psi}$ is the trial function to be varied. [We use the inner-product notation,

$$\langle \psi_1 \,|\, \mathcal{O} \,|\, \psi_2 \rangle \equiv \int dx \psi_1^\dagger(x) \mathcal{O}(x) \psi_2(x),$$

where x represents whatever independent variables the problem has.]

(ii) Solutions to (1) exist only for some set of eigenvalues λ_n,

$$(H - \lambda_n)\psi_n = 0. \tag{3}$$

(iii) These λ_n may be discrete or continuous, but are real numbers; and for problems with all symmetry removed, we can assume the eigenvalues are nondegenerate.

(iv) The eigenfunctions ψ_n form a complete orthogonal set which can be normalized as

* This research was supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under Grant No. AF-AFOSR-130-63.

$$\langle \psi_n \mid \psi_m \rangle = \delta_{nm} . \tag{4}$$

(v) The identity operator may be expanded as

$$\delta(x - x') = \sum_n \psi_n(x)\psi_n^\dagger(x'). \tag{5}$$

(vi) We can resolve the Green's function as

$$\left\langle G \left| \frac{1}{H - \lambda} \right| F \right\rangle = \sum_n \frac{\langle G \mid \psi_n \rangle \langle \psi_n \mid F \rangle}{\lambda_n - \lambda} . \tag{6}$$

The generalized problem we wish to study is that of the equation

$$L(\lambda)\psi = 0, \tag{7}$$

where the linear operator L may depend on the parameter λ in some complicated manner; and moreover we allow $L^\dagger$ to differ from L. Consequently, we can consider also the adjoint equation

$$L^\dagger(\lambda)\phi = 0,$$

or equivalently,

$$\phi^\dagger L(\lambda) = 0. \tag{8}$$

Our objective is to see which of the familiar properties listed earlier for the simple case (1) carry over to this more general case. Probably none of what we have to say will be new, but some of the results were sufficiently nonobvious to the present writer that it seemed worthwhile to collect and present the following conclusions. No attempt at mathematical rigor will be made.

Think about the familiar problem of finding the energy eigenvalues for the Schrödinger equation. We construct solutions of the differential equation, and then find that the boundary conditions will be satisfied only for some particular values of the energy E which enters as a parameter in the equation. The fact that E occurs linearly [as λ in Eq. (1)] is quite irrelevant to this process. We therefore start our analysis with the following assumption.

ASSUMPTION. Solutions ψ_n of (7) exist only for some (nondegenerate) set of eigenvalues λ_n ; the totality of the eigenfunctions ψ_n form a complete, linearly independent set. The same holds, independently, for the solutions $\phi_m^\dagger$ of (8). (9)

(i′) For a variational principle we construct the following (*1*):

$$J[\tilde{\phi}^\dagger, \tilde{\psi}] = \langle \tilde{\phi} \mid L(\lambda) \mid \tilde{\psi} \rangle. \tag{10}$$

It is obvious that requiring J to be stationary under arbitrary variation of $\tilde{\phi}^\dagger$ will lead to Eq. (7) for $\tilde{\psi}$, and requiring J to be stationary under arbitrary variation of $\tilde{\psi}$ will lead to Eq. (8) for $\tilde{\phi}^\dagger$.

The process of making J stationary will also lead to a value $\tilde{\lambda}$ for the eigen-

value λ. This may be illustrated by considering the case where we choose our trial functions with linear variational parameters, as follows.

$$\tilde{\psi}(x) = \sum_{i=1}^{N} C_i u_i(x), \tag{11a}$$

$$\tilde{\phi}^{\dagger}(x) = \sum_{j=1}^{N} D_j v_j(x), \tag{11b}$$

where the u_i and the v_j are functions chosen from two arbitrary complete sets. The expression for J is then

$$J[D, C] = \sum_{i,j=1}^{N} D_j L_{ji}(\lambda) C_i, \tag{12}$$

where

$$L_{ji}(\lambda) = \int v_j L(\lambda) u_i; \tag{13}$$

and the variation of C_i gives

$$\sum_{j=1}^{N} D_j L_{ji}(\lambda) = 0, \tag{14}$$

while the variation of D_j gives

$$\sum_{j=1}^{N} L_{ji}(\lambda) C_i = 0. \tag{15}$$

Equations (14) and (15) each have nontrivial solutions only if

$$\det \| L_{ji}(\lambda) \| = 0 \tag{16}$$

and this algebraic equation determines the approximate eigenvalues $\tilde{\lambda}_n$. As we consider N tending to infinity, this matrix representation should yield the exact solutions for λ_n, ψ_n, and $\phi_n{}^{\dagger}$. For this special class of variational approximations (linear parameters only) we also see that the stationary value of J, call it $\tilde{J}$, is zero.

An important property of many variational principles is that they give second-order accuracy for some quantity, when the trial function has a first-order error. We shall now see how this comes about for our problem (10). Suppose we have approximate functions $\tilde{\psi}$, $\tilde{\phi}^{\dagger}$ and eigenvalue $\tilde{\lambda}$ close, in first order, to some exact solution ψ, $\phi^{\dagger}$, λ:

$$\tilde{\psi} = \psi + \delta\psi, \qquad \tilde{\phi}^{\dagger} = \phi^{\dagger} + \delta\phi^{\dagger}, \qquad \tilde{\lambda} = \lambda + \delta\lambda. \tag{17}$$

Now we calculate the resulting value of J.

$$\begin{aligned} \tilde{J} = \langle \tilde{\phi} \mid L(\tilde{\lambda}) \mid \tilde{\psi} \rangle &= \langle \phi \mid L(\lambda) \mid \psi \rangle + \langle \phi \mid L(\lambda) \mid \delta\psi \rangle + \langle \delta\phi \mid L(\lambda) \mid \psi \rangle \\ &\quad + \langle \tilde{\phi} \mid (\partial L/\partial\lambda)_{\lambda=\tilde{\lambda}} \mid \tilde{\psi} \rangle (\tilde{\lambda} - \lambda) + \text{second-order small terms.} \end{aligned} \tag{18}$$

The first three terms vanish, and so we can solve for $\delta\lambda$:

$$\lambda = \tilde{\lambda} + \tilde{J}/\tilde{N} + \text{second-order terms.} \tag{19}$$

Here we have defined the generalized normalization integral as

$$N_\lambda = \langle \phi_\lambda \mid -\partial L(\lambda)/\partial\lambda \mid \psi_\lambda \rangle. \tag{20}$$

Formulas (19) and (20) for the eigenvalue λ of (7) may be read as the operator generalization of Newton's method (*2*) for finding a root of the function $L(\lambda)$.

Returning to the special case $L(\lambda) = H - \lambda I$, $\phi^\dagger = \psi^*$, we find

$$N = \langle \psi \mid \psi \rangle$$

and (to second order)

$$\lambda = \tilde{\lambda} + \frac{\langle \tilde{\psi} \mid H - \tilde{\lambda} \mid \tilde{\psi} \rangle}{\langle \tilde{\psi} \mid \tilde{\psi} \rangle} = \frac{\langle \tilde{\psi} \mid H \mid \tilde{\psi} \rangle}{\langle \tilde{\psi} \mid \tilde{\psi} \rangle}. \tag{21}$$

This is the famous Rayleigh–Ritz formula.

Also for the special case of linear variational parameters we saw that $\tilde{J} = 0$ so that $\lambda = \tilde{\lambda}$ to second order.

Thus it appears that none of the great power of variational methods of approximation has been lost in going from (1) to (7). [An exception to this statement is the upper bound property of (21) when applied to the lowest eigenvalue; but the writer considers this to be a matter of lesser importance.]

We now go on to consider the other aspects of the problem.

(ii′) The above matrix representation of (7) tells as much as we can say about the spectrum of eigenvalues λ_n ; for the most part we shall appeal to our Assumption (9).

(iii′) The eigenvalues λ_n may be discrete or continuous, real or complex. However there exist interesting problems, nonlinear in λ and even non self-adjoint, for which the spectrum of λ is still purely real; we know examples of this but no general criteria.

(iv′) We have assumed the ψ_n to be a complete set (and similarly the $\phi_m{}^\dagger$); but in general there is *no orthogonality theorem.*

One exception concerns eigenfunctions which are very close to one another, as in a continuum. For eigenvalues λ_α , λ_β , etc., which are very close to some λ_0 we can write Eq. (7) as

$$[L(\lambda_0) + (\partial L/\partial\lambda)_{\lambda=\lambda_0}(\lambda_\alpha - \lambda_0)]\psi_\alpha = O[(\lambda_\alpha - \lambda_0)^2]. \tag{22}$$

If the right-hand side of this equation becomes negligible, we have a *linear* eigenvalue problem (for λ_α) with $\partial L/\partial\lambda$ playing the role of a weight function. Then, in the usual manner, we can deduce the orthogonality property

$$\langle \phi_\beta \mid (\partial L/\partial\lambda)_{\lambda=\lambda_0} \mid \psi_\alpha \rangle = 0 \quad \text{for} \quad \alpha \neq \beta, \qquad \text{but } \lambda_\alpha \text{ and } \lambda_\beta \text{ very close to } \lambda_0. \tag{23}$$

Some understanding of this special formula (23) may be gained by imagining that the continuous spectrum is something like that of free waves in an infinite box. Then if we further imagine that the operator $\partial L/\partial\lambda$ does not confine the region of integration in (23), we can conclude that the integral will be essentially zero if the wavelengths, labeled by α and β, are not equal.

The above discussion of the variational principle leads us to define a generalized normalization integral for each eigenfunction pair. We offer, as a general criterion, the statement that N should not be infinite. (It is however conceivable that N might be zero for some state, and this special case deserves further study.) This normalizability criterion serves, at least partially, to provide the boundary conditions necessary to make the original eigenvalue problem well defined. Perhaps the complete specification of these boundary conditions is already contained in the statement of the variational principle itself: requiring that J and its variations be bounded, and furthermore that partial integrations should not leave extra surface terms.

(v′) A generalized expansion of the delta-function is

$$\delta(x - y) = \sum_{m,n} P(x)\psi_m(x)W_{mn}\phi_n^{\dagger}(y)Q(y), \tag{24}$$

where P and Q are any nonsingular operators and the numbers W_{mn} will now be determined. Multiply (24) from the left by $\phi_l^{\dagger}(x)Q(x)$ and then integrate over x; there results

$$\phi_l^{\dagger}(y)Q(y) = \sum_n \Big\{ \sum_m \langle\phi_l \mid QP \mid \psi_m\rangle W_{mn} \Big\} \phi_n^{\dagger}(y)Q(y). \tag{25}$$

From the linear independence of the $\phi_n^{\dagger}(y)$ we can then conclude that we require

$$\sum_m \langle\phi_l| QP |\psi_m\rangle W_{mn} = \delta_{ln}. \tag{26}$$

A similar operation with $P(y)\psi_l(y)$ on the right of (24) shows that W is both the right and the left inverse of the matrix

$$\langle\phi_m| QP |\psi_n\rangle. \tag{27}$$

(vi′) Finally we consider an integral involving the Green's function:

$$D(\lambda) = \int dx G^{\dagger}(x) \frac{1}{L_x(\lambda)} F(x) = \left\langle G \middle| \frac{1}{L(\lambda)} \middle| F \right\rangle. \tag{28}$$

This function D may be an analytic function of λ, assuming suitably nice properties for F and G, except when λ approaches some eigenvalue λ_0 of the equation

$$L(\lambda_0)\psi_0 = 0.$$

We can find no explicit resolution of (28) to compare with (6); but we can investigate $D(\lambda)$ in the immediate neighborhood of the point $\lambda = \lambda_0$.

First we insert into (28), just before the function F the unit operator as represented by (24) with the choice $P = L(\lambda)$, $Q = I$:

$$D(\lambda) = \sum_{m,n} \langle G \mid \psi_m \rangle W_{mn} \langle \phi_n \mid F \rangle. \tag{29}$$

The dependence on λ is now buried in W, which is the inverse of the matrix

$$T_{nl} = \langle \phi_n | \, L(\lambda) \, | \psi_l \rangle. \tag{30}$$

Now we set $\lambda = \lambda_0 + \epsilon$ and expand for small ϵ, keeping only the leading term in each case.[1]

$$T_{00} = \epsilon \langle \phi_0 \mid (\partial L/\partial \lambda)_{\lambda=\lambda_0} \mid \psi_0 \rangle = -\epsilon N_0 , \tag{31a}$$

$$T_{0j} = -\epsilon N_{0j} , \tag{31b}$$

$$T_{i0} = -\epsilon N_{i0} , \tag{31c}$$

$$T_{ij} = \langle \phi_i | \, L(\lambda_0) \, | \psi_j \rangle, \tag{31d}$$

where $i, j \neq 0$. If we now write out the equations determining W,

$$\sum_n W_{mn} T_{nl} = \sum_n T_{mn} W_{nl} = \delta_{ml} , \tag{32}$$

separating zero and nonzero components and then keeping only leading terms for $\epsilon \to 0$, we find

$$\sum_k W_{ik} T_{kj} = \sum_k T_{ik} W_{kj} = \delta_{ij} , \qquad i, j, k \neq 0, \tag{33a}$$

$$W_{0j} = -\frac{1}{N_{00}} \sum_k N_{0k} W_{kj} , \tag{33b}$$

$$W_{i0} = -\frac{1}{N_{00}} \sum_k W_{ik} N_{k0} , \tag{33c}$$

$$W_{00} = -\frac{1}{\epsilon N_0} + \text{finite terms.} \tag{33d}$$

Thus we see that W is dominated by the first term in (33d), and our final formula is

$$D(\lambda) \xrightarrow[\lambda \to \lambda_0]{} \frac{\langle G \mid \psi_0 \rangle \langle \phi_0 \mid F \rangle}{\langle \phi_0 \mid - (\partial L/\partial \lambda)_{\lambda=\lambda_0} \mid \psi_0 \rangle} \left(\frac{1}{\lambda_0 - \lambda} \right). \tag{34}$$

The result (34) is really correct only if λ_0 is a discrete eigenvalue; if λ_0 lies in some dense set of eigenvalues λ_α (becoming continuous in some limit), then we

[1] Note: We require that ϵ can be chosen so that λ is not an eigenvalue; this assumes that the spectrum is at most dense along a line, but not filling any area in the complex plane.

proceed as follows. Let the labels α, β stand for the set of states with eigenvalues very close to λ; these labels replace the single label 0 used above. The dominant terms in the matrix W_{mn} are seen to be the elements $W_{\alpha\beta}$, which represent the inverse of the matrix

$$T_{\alpha\beta} = \langle\phi_\alpha| L(\lambda) |\psi_\beta\rangle \cong \langle\phi_\alpha| -\partial L/\partial\lambda |\psi_\beta\rangle(\lambda_\beta - \lambda). \tag{35}$$

Now we can use the special orthogonality relation, Eq. (23), to conclude that

$$T_{\alpha\beta} = N_0(\lambda_\alpha - \lambda)\delta_{\alpha\beta}\,. \tag{36}$$

This leads to

$$W_{\alpha\beta} = \delta_{\alpha\beta}/N_0(\lambda_\alpha - \lambda), \tag{37}$$

and the final formula is

$$D(\lambda) \to \sum_\alpha \frac{\langle G|\psi_\alpha\rangle\langle\phi_\alpha|F\rangle}{N_0(\lambda_\alpha - \lambda)}\,. \tag{38}$$

This is the best we can do to generalize (6), and we see again the appearance of the generalized normalization integral N_0 defined in Eq. (20). (If N_0 should vanish, we would have to carry our analysis further, expecting to find that D has a higher order pole at $\lambda = \lambda_0$.)

As an example of some things we have been discussing, consider the Bethe–Salpeter equation, after rotation to an imaginary time variable (3). The differential operator is

$$L = \{[\mathbf{P}^2 + (P_4 - i\mu_1 E)^2 + m_1^2][\mathbf{P}^2 + (P_4 + i\mu_2 E)^2 + m_2^2] - V(x)\}, \tag{39}$$

$$\mathbf{P}^2 = \sum_{i=1}^{3} P_i^2.$$

This is in a 4-dimensional space and each P is $(-i)$ times a spatial derivative operator. The eigenvalue parameter is the energy E, which obviously occurs in a nonlinear manner. The identification of

$$N = -\partial L/\partial E$$

as an appropriate normalization operator has been made for this equation by several authors (4), but the generality of this form has perhaps not been appreciated.

It is also apparent that, except in the special case $m_1 = m_2$ and $\mu_1 = \mu_2$, the operator (39) is not self-adjoint in the usual sense. However, an extended definition of the adjoint operation to

$$\psi^\dagger(\mathbf{x}, x_4) = \psi^*(\mathbf{x}, -x_4)$$

will allow us to say that

$$L^{\dagger} = L.$$

In such cases as this, where some simple operation transforms the right eigenvector ψ into the left eigenvector $\phi^{\dagger}$, we can say that the operator is effectively self-adjoint.

ACKNOWLEDGMENT

The author has benefited much from several discussions with A. Goldhaber.

RECEIVED: May 26, 1966

REFERENCES

1. Such "two-sided" variational principles have been discussed in: M. BECKER, "The Principles and Applications of Variational Methods." M.I.T. Press Cambridge, Massachusetts, 1964.
2. L. V. KANTOROVICH, "Functional Analysis and Applied Mathematics," Chapter IV. [Translated by C. D. Benster from *Usp. Mat. Nauk* **3** (1948).] National Bureau of Standards, Los Angeles, 1952.
3. G. C. WICK, *Phys. Rev.* **96,** 1124 (1954).
4. M. CIAFALONI AND P. MENOTTI, *Phys. Rev.* **140,** B929 (1965), and earlier papers referred to therein.