DERIVATIVES WITH RESPECT TO PARAMETERS OF THE EIGENVECTORS OF A HAMILTONIAN AND THEIR APPLICATION

A. Ya. Tsaune

UDC 539.101:182.192

A method is described for finding the first- and higher-order derivatives of the eigenvectors of a Hamiltonian with respect to its parameters. This method is useful even when the explicit dependence of the eigenvectors on the parameters is not known. The method is based on a transfer of the differentiation from the eigenvector to the Hamiltonian and on a separate analysis of the derivatives of the projections of the eigenvector onto the corresponding subspace and onto the orthogonal complement of this subspace. Conditions governing the position of the eigenvector being differentiated in its degenerate subspace are analyzed. This method can be used in certain fundamental problems, and it can be related to steady-state Rayleigh–Schrödinger perturbation theory.

In many physical problems the derivatives of eigenvectors of a Hermitian operator with respect to parameters must be found. In the overwhelming majority of cases, however, the explicit dependences of these vectors on the parameters are unknown, so a direct differentiation cannot be carried out. In the method described below this difficulty is avoided by transferring the differentiation from the vector to the operator. We will use the Hamiltonian as the Hermitian operator.

We start from the stationary Schrödinger equation

$$(H-IE_s)|s\sigma\rangle = 0, \tag{1}$$

where H is the Hamiltonian defined in Hilbert space Λ; I is the unit operator; E_s is an eigenvalue; $|s^\sigma\rangle$ is an eigenvector; s denotes the (generally degenerate) corresponding subspace, and σ denotes a vector in $\Lambda^{(s)}$.

The Hamiltonian is a function of the parameters R_j ($j = 1, 2, \ldots, N$), the set of which we denote by R; we thus have

$$H = H(R). \tag{2}$$

We obviously also have

$$E_s = E_s(R), \quad |s\sigma\rangle = |s\sigma(R)\rangle. \tag{3}$$

We will omit the R below for brevity; to indicate that H, E, etc. correspond to fixed values of the parameters ($R = R_0$) we write

$$\begin{gathered} H(R_0) \equiv H_0, \quad E_s(R_0) \equiv E_{s0}, \quad |s\sigma(R_0)\rangle \equiv |s\sigma\rangle_0, \\ \langle s\sigma'(R_0)|H(R_0)|s\sigma(R_0)\rangle \equiv \langle s\sigma'|H|s\sigma\rangle_0. \end{gathered} \tag{4}$$

The eigenvectors can be orthonormalized for any parameters; i.e., we can write

$$\langle s'\sigma'|s\sigma\rangle = \delta_{ss'}\delta_{\sigma\sigma'}. \tag{5}$$

We introduce the following notation for the derivatives:

$$\frac{\partial H}{\partial R_j} \equiv \partial_j H, \quad \frac{\partial E_s}{\partial R_j} \equiv \partial_j E_s, \quad \frac{\partial|s\sigma\rangle}{\partial R_j} \equiv |\partial_j s\sigma\rangle \tag{6}$$

and we will assume that H is differentiable with respect to its parameters as many times as necessary.

Dnepropetrovsk Chemical Engineering Institute. Translated from Izvestiya Vysshikh Uchebnykh Zavedenii Fizika, No. 3, pp. 74–78, March, 1971. Original article submitted February 11, 1970.

To find the first derivative, we differentiate (1) and transform to the equilibrium values of the parameters, finding

$$(H - IE_s)_0 | \partial_j s\sigma \rangle_0 = - [\partial_j (H - IE_s)]_0 | s\sigma \rangle_0. \tag{7}$$

We cannot find $|\partial_j s\sigma\rangle_0$ directly from this expression, since $(H-IE_S)$ has no inverse. We thus single out from Λ the subspace Λ_0' such that

$$\Lambda = \Lambda_0^{(s)} \oplus \Lambda_0', \tag{8}$$

where $\Lambda_0^{(s)}$ is $\Lambda^{(s)}$ at $R = R_0$, and $\oplus$ indicates the orthogonal sum. Obviously, $\Lambda_0^{(s)}$ is an invariant subspace of operator H_0, and projection onto this subspace does not remove the vectors from the region of definition. We assume that Λ_0' is also invariant; then H_0 and I, respectively, generate [1]

the operators $H_0^{(s)}$ and $I^{(s)}$ in $\Lambda_0^{(s)}$ (9a)

and the operators H_0' and I' in Λ_0'. (9b)

We introduce P_{S0} and $(I-P_S)_0$, the orthogonal projection operators projecting onto $\Lambda_0^{(s)}$ and Λ_0', respectively [1]; we can now show that we have [1]

$$P_{s0} (H - IE_s)_0 = (H - IE_s)_0 P_{s0}, \tag{10a}$$

$$(I - P_s)_0 (H - IE_s)_0 = (H - IE_s)_0 (I - P_s)_0. \tag{10b}$$

Multiplying (7) from the left by P_{S0} and $(I-P_S)_0$; and using (10a), (10b), and (9); we find

$$(H - IE_s)_0^{(s)} P_{s0} | \partial_j s\sigma \rangle_0 = - P_{s0} [\partial_j (H - IE_s)]_0 | s\sigma \rangle_0, \tag{11a}$$

$$(H - IE_s)_0' (I - P_s)_0 | \partial_j s\sigma \rangle_0 = - (I - P_s)_0 [\partial_j (H - IE_s)]_0 | s\sigma \rangle_0. \tag{11b}$$

Since we have

$$| \partial_j s\sigma \rangle_0 = P_{s0} | \partial_j s\sigma \rangle_0 + (I - P_s)_0 | \partial_j s\sigma \rangle_0, \tag{12}$$

Eq. (11b) gives the second part of the derivative, while (11a) holds only if its right side vanishes, since matrix $(H-IE_S)_0^{(s)}$ is of rank zero. Otherwise, we must transform to new eigenvectors in $\Lambda_0^{(s)}$:

$$| s\tau \rangle_0 = \sum_\sigma | s\sigma \rangle_0 \langle s\sigma | s\tau \rangle_0, \tag{13}$$

which satisfy

$$P_{s0} [\partial_j (H - IE_s)]_0 | s\tau \rangle_0 = 0 \quad (j = 1, 2, \ldots, N). \tag{14}$$

The solution of this equation always exists; $(\partial_j E_S)_0$ serves as an eigenvalue. Two cases are possible:

1) The vectors $|s\tau\rangle_0$ ($\tau = 1, 2, \ldots, M$) do not depend on j; i.e., there exists in $\Lambda_0^{(s)}$ at least one set of vectors which satisfy (14) for all j.

2) There is no set of vectors which satisfy (14) for all j.

We assume the first case. Then according to the value of $(\partial_j E_S)_0$ we can resolve $\Lambda_0^{(s)}$ into $\Lambda_0^{(s1)}$, corresponding $(\partial_j E_S)_0^{(1)}$, with $|s\tau^{(1)}\rangle_0$ ($\tau^{(1)} = 1, 2, \ldots, m_1$); $\Lambda_0^{(s2)}$, corresponding to $(\partial_j E_S)_0^{(2)}$, with $|s\tau^{(2)}\rangle_0$ ($\tau^{(2)} = 1, 2, \ldots, M_2$); ...; and $\Lambda_0^{(sn)}$, corresponding to $(\partial_j E_S)_0^{(n)}$, with $|s\tau^{(n)}\rangle_0$ ($\tau^{(n)} = 1, 2, \ldots, M_n$). Here we have

$$\Lambda_0^{(s)} = \sum_{\nu=1}^{n} \oplus \Lambda_0^{(s\nu)}, \quad \sum_{\nu=1}^{n} M_\nu = M. \tag{15}$$

The problem now reduces to that of determining

$$| \partial_j s\tau \rangle_0 = P_{s0} | \partial_j s\tau \rangle_0 + (I - P_s)_0 | \partial_j s\tau \rangle_0, \tag{16}$$

in which the second term is found from (11b) with $|s\sigma\rangle_0$ replaced by $|s\tau\rangle_0$, and the first term satisfies

$$(H - IE_s)_0^{(s)} P_{s0} | \partial_j s\tau \rangle_0 = 0. \tag{17}$$

Since matrix $(H-IE_S)_0^{(s)}$ is of rank zero, the solution of this equation is not restricted in any manner, and we can assume

$$P_{s0} | \partial_j s\tau \rangle_0 = 0. \tag{18}$$

It is not difficult to show that this condition does not violate the condition for orthonormality within $\Lambda_0^{(s)}$:

$$\langle s\tau' | s\tau \rangle = \delta_{\tau\tau'}, \tag{19}$$

which follows from (5). The first derivative can thus always be determined from simply the second term in (16).

To find the second derivative, we differentiate (1) (where $|s\sigma\rangle$ has been replaced by $|s\tau\rangle$) twice; fix the parameters; multiply the resulting equation from the left by P_{s0} and $(I-P_s)_0$; and use (10a), (10b), and (9); finding

$$(H-IE_s)_0^{(s)}P_{s0}|\partial_{\kappa j}s\tau\rangle_0 = -P_{s0}[\partial_{\kappa j}(H-IE_s)]_0|s\tau\rangle_0 - P_{s0}[\partial_j(H-IE_s)]_0|\partial_\kappa s\tau\rangle_0 - P_{s0}[\partial_\kappa(H-IE_s)]_0|\partial_j s\tau\rangle_0, \quad (20a)$$

$$\begin{aligned}(H-IE_s)_0'(I-P_s)_0|\partial_{\kappa j}s\tau\rangle_0 &= -(I-P_s)_0[\partial_{\kappa j}(H-IE_s)]_0|s\tau\rangle_0 \\ &-(I-P_s)_0[\partial_j(H-IE_s)]|\partial_\kappa s\tau\rangle_0 - (I-P_s)_0[\partial_\kappa(H-IE_s)]_0|\partial_j s\tau\rangle_0.\end{aligned} \quad (20b)$$

The second equation gives $(I-P_s)_0|\partial_{\kappa j}s\tau\rangle_0$ unambiguously, while the first requires that the right side of (10a) vanish. Otherwise, we must further refine the vectors within each $\Lambda_0^{(s\nu)} \subset \Lambda_0^{(s)}$ ($\nu = 1, 2, \ldots, n$) in (15); i.e., we must transform to

$$|s\theta^{(\nu)}\rangle_0 = \sum_{\tau^{(\nu)}=1}^{M_\nu} |s\tau^{(\nu)}\rangle_0 \langle s\tau^{(\nu)}|s\theta^{(\nu)}\rangle_0$$
$$(\theta^{(\nu)} = 1, 2, \ldots, M_\nu; \quad \nu = 1, 2, \ldots, n), \quad (21)$$

which satisfy the equation found by equating the right side of (20a) to zero (after the first derivatives are substituted in). The eigenvalues of this system are $(\partial_{\kappa j}E_s)_0$, and different values of these eigenvalues lead to a further resolution of each $\Lambda_0^{(s\nu)}$ such that we have

$$\Lambda_0^{(s\nu)} = \sum_{\varphi=1}^{m} \oplus \Lambda_0^{(s\nu\varphi)}, \quad (22)$$

where $\Lambda_0^{(s\nu\varphi)}$ is formed from $|s\theta^{(\nu\varphi)}\rangle_0$ $(\theta^{(\nu\varphi)} = 1, 2, \ldots, M_{\nu\varphi})$ and where we have

$$\sum_{\varphi=1}^{m} M_{\nu\varphi} = M_\nu. \quad (23)$$

As in the analysis of the first derivative, we assume that for all k, j = 1, 2, . . . , N there exists a single set $|s\theta^{(\nu)}\rangle_0$; then in the expression for the second derivative.

$$|\partial_{\kappa j}s\theta\rangle_0 = P_{s0}|\partial_{\kappa j}s\theta\rangle_0 + (I-P_s)_0|\partial_{\kappa j}s\theta\rangle_0 \quad (24)$$

the second term is given by (20b) (where $|s\tau\rangle_0$ has been replaced by $|s\theta\rangle_0$ in the derivatives), and the first term can be written as

$$P_{s0}|\partial_{\kappa j}s\theta\rangle_0 = \sum_{\theta'=1}^{M} |s\theta'\rangle_0 \langle s\theta'|\partial_{\kappa j}s\theta\rangle_0. \quad (25)$$

The coefficients on the right side of this equation are found by a repeated differentiation of the orthonormality condition analogous to (19):

$$\langle s\theta'|s\theta\rangle = \langle s\theta|s\theta'\rangle = \delta_{\theta\theta'} \quad (26)$$

It is not difficult to show that these coefficients can be written in the symmetric form

$$\mathrm{Re}\langle s\theta|\partial_{\kappa j}s\theta'\rangle_0 = \mathrm{Re}\langle s\theta'|\partial_{\kappa j}s\theta\rangle_0 = -\frac{1}{2}[\mathrm{Re}\langle\partial_j s\theta'|\partial_\kappa s\theta\rangle_0 + \mathrm{Re}\langle\partial_j s\theta|\partial_\kappa s\theta'\rangle_0], \quad (27a)$$

$$\mathrm{Im}\langle s\theta|\partial_{\kappa j}s\theta'\rangle_0 = -\mathrm{Im}\langle s\theta'|\partial_{\kappa j}s\theta\rangle_0 = \frac{1}{2}[\mathrm{Im}\langle\partial_j s\theta'|\partial_\kappa s\theta\rangle_0 - \mathrm{Im}\langle\partial_j s\theta|\partial_\kappa s\theta'\rangle_0]. \quad (27b)$$

Here Re and Im are the signs of the real and imaginary parts.

The procedure for finding the third and higher derivatives is obvious. The eigenvectors to be differentiated are refined, and the resolution of $\Lambda_0^{(s)}$ is continued. It is also obvious that this process is related to a "removal of the degeneracy" as the parameters are varied.

If there does not exist a set of vectors which satisfy (14) for all j [see case 2 below Eq. (14)], no further differentiation can be carried out, since certain vectors from $\Lambda_0^{(s)}$, must be used as original vectors for certain j, while other vectors must be used for other j. The set of parameters R_j of this type cannot be used to simultaneously find the derivatives of a given vector.

Schrödinger equation (1) was differentiated in this derivation; this approach was first used in [2] (see also [3]) to find the first derivative of the eigenvector and the second derivative of the corresponding nondegenerate eigenvalue. All the equations were obtained on the basis of eigenvectors of the Hamiltonian.

The derivatives of the eigenvectors of the Hamiltonian are also of independent interest, since they can be used in the following general problems:

1. To determine the eigenvectors for a set of parameters different from the fixed set, through the use of the series

$$|s\theta(R_0+\Delta R)> = |s\theta>_0 + \sum_j |\partial_j s\theta>_0 \Delta R_j + \frac{1}{2}\sum_{\kappa j}|\partial_{\kappa j} s\theta>_0 \Delta R_j \Delta R_\kappa + \dots \quad (28)$$

2. To find the derivatives of the matrix elements of arbitrary (parameter-dependent) operators A in the basis of eigenvectors of the Hamiltonian:

$$[\partial_j <s'\theta'|A|s\theta>]_0 = <\partial_j s'\theta'|A|s\theta>_0 + <s'\theta'|\partial_j A|s\theta>_0 + <s'\theta'|A|\partial_j s\theta>_0. \quad (29)$$

In particular, this procedure allows us to find the derivatives of the energy,

$$E_s = <s\theta|H|s\theta>. \quad (30)$$

3. To expand a matrix element in a series in terms of increments in the parameters; e.g., for the energy, to write

$$E_s(R_0+\Delta R) = E_{s0} + \sum_j (\partial_j E_s)_0 \Delta R_j + \frac{1}{2}\sum_{\kappa j}(\partial_{\kappa j} E_s)_0 \Delta R_j \Delta R_\kappa + \dots \quad (31)$$

Equations (28) and (31) can evidently be thought of as the equations of a "multiparameter perturbation theory."

In conclusion we turn to two particular cases. Let us assume that the Hamiltonian depends only on a single parameter, in the usual form for steady-state Rayleigh–Schrödinger perturbation theory [4]:

$$H = H^{(0)} + RH', \quad (32)$$

and E_s is nondegenerate. Then the second index in the expression for an eigenvector can be omitted. We write all the operators in the basis of eigenvectors of $H^{(0)}$; then we have

$$P_{s0} = |s>_{0\,0}<s|, \quad (I-P_s)_0 = \sum_{r\neq s}|r>_{0\,0}<r|, \quad (33)$$

and

$$<r|(H-IE_s)'^{-1}|t> = \frac{\delta_{rt}}{E_r^{(0)}-E_s^{(0)}} \quad (r,\, t\neq s). \quad (34)$$

Using (11b), (18), (20b), (25), (27a), and (27b), and using $(\partial E_s)_0 = <s|\partial H|s>_0$, we find

$$P_{s0}|\partial s>_0 = 0, \ (I-P_s)_0|\partial s>_0 = -\sum_{r\neq s}|r>_0 \frac{<r|\partial H|s>_0}{E_r^{(0)}-E_s^{(0)}}, \quad (35a)$$

$$P_{s0}|\partial^2 s>_0 = -|s>_0 \sum_{r\neq s}\frac{|<r|\partial H|s>_0|^2}{(E_r^{(0)}-E_s^{(0)})^2}, \quad (35b)$$

$$(I-P_s)_0|\partial^2 s>_0 = 2\sum_{r\neq s}\sum_{t\neq s}|r>_0 \frac{<r|\partial H|t>_0<t|\partial H|s>_0}{(E_r^{(0)}-E_s^{(0)})(E_t^{(0)}-E_s^{(0)})} - 2\sum_{r\neq s}|r>_0\frac{<r|\partial H|s>_0<s|\partial H|s>_0}{(E_r^{(0)}-E_s^{(0)})^2}. \quad (35c)$$

Also using (12) and (24), substituting everything into (28), and setting $\Delta R = 1$, we find Eq. (25.14) of [4], which gives the eigenvector according to second-order perturbation theory.

Substituting (35a) into (29), using (12), and setting s' = s, we find

$$[\partial <s|A|s>]_0 = <s|\partial A|s>_0 - \sum_{r\neq s}\frac{<s|\partial H|r>_0<r|A|s>_0}{E_r^{(0)}-E_s^{(0)}} - \sum_{r\neq s}\frac{<s|A|r>_0<r|\partial H|s>_0}{E_r^{(0)}-E_s^{(0)}},$$

which is the sum-rule equation of [5].

LITERATURE CITED

1. N. I. Akhiezer and I. M. Glazman, Theory of Linear Operators in Hilbert Space [in Russian], Nauka, Moscow (1966).
2. B. W. Brown, Proc. Camb. Phil. Soc., 58, 251 (1958).
3. J. C. Y. Chen, J. Chem. Phys., 38, No. 1, 283 (1963).
4. L. Schiff, Quantum Mechanics, McGraw–Hill, New York (1968).
5. L. Salem, Phys. Rev., 125, 1788 (1962).