Proc. Indian Acad. Sci. (Chem. Sci.), Vol. 96, Nos 3 & 4, February 1986, pp. 121–126.

A generalization of Brillouin's theorem and the stability conditions in the quantum-mechanical variation principle in the case of general trial wave functions

PER-OLOV LÖWDIN*

Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Fl 32611, USA

* NFCR Senior Investigator, Professor Emeritus at Uppsala University, Uppsala, Sweden

Abstract. In connection with the quantum-mechanical variation principle, it is shown that the Brillouin theorem and the stability conditions usually associated with the Hartree-Fock scheme in the many-electron theory may be generalized to the case of arbitrary trial functions depending on a set of linear or non-linear complex parameters.

1. The eigenvalue problem and the variation principle

The purpose of this paper is to show that the Brillouin theorem (Brillouin 1933, 1934) and the stability conditions (Thouless 1961; Adams 1962) usually associated with the Hartree-Fock scheme in many-electron theory are general features of the variation principle even in the case when the trial wave function is of a more general nature and depends on a set of linear or non-linear complex parameters.

Many of the fundamental problems in quantum mechanics may be formulated as eigenvalue problems with proper boundary conditions of the type

$$H\Psi = E\Psi, \tag{1}$$

or more generally

$$\Omega_1\Psi = \lambda\Omega_2\Psi, \tag{2}$$

where H, Ω_1, and Ω_2 are self-adjoint operators; here we will also assume that Ω_2 is positive definite. If one introduces the variational expression

$$I = \langle\Phi|\Omega_1|\Phi\rangle/\langle\Phi|\Omega_2|\Phi\rangle = A/B, \tag{3}$$

where Φ is an arbitrary trial wave function, the eigenvalue problem (2) is equivalent with the variation principle:

$$\delta I = 0, \tag{4}$$

for the first-order variation. In a previous paper (Löwdin *et al* 1981), we considered the

Dedicated to Professor Sadhan Basu on the occasion of his 65th birth anniversary.

change of I associated with a finite variation

$$\Delta\Phi = \Phi' - \Phi \tag{5}$$

of the trial wave function Φ; one has

$$\Delta I = I' - I = (A'/B') - (A/B) = (1/B')\,(A' - IB') = \\ = (1/B')\,(\Delta A - I\Delta B). \tag{6}$$

Introducing the combined operator

$$Q = \Omega_1 - I\Omega_2, \tag{7}$$

we obtained:

$$B'I = \Delta A - I\Delta B = \langle \Delta\Phi | Q | \Phi \rangle + \langle \Phi | Q | \Delta\Phi \rangle + \\ + \langle \Delta\Phi | Q | \Delta\Phi \rangle, \tag{8}$$

where the quantity $B' = \langle \Phi' | \Omega_2 | \Phi' \rangle$ is always positive.

In the treatment of a many-particle system, let us now introduce a general trial wave function Φ which depends not only on the particle coordinates but also on a set of linear or non-linear complex parameters

$$\mathbf{c} = \{c_1, c_2 \ldots c_p\}:$$
$$\Phi = \Phi(\mathbf{c}) = \Phi(c_1, c_2, \ldots c_p). \tag{9}$$

Giving these parameters a finite variation, so that $c_k' = c_k + \Delta c_k$, one obtains the Taylor expansion

$$\Delta\Phi = \Phi(\mathbf{c} + \Delta\mathbf{c}) - \Phi(\mathbf{c}) = \sum_k \Delta c_k \cdot \Phi_k + \frac{1}{2!} \sum_{k,l} \Delta c_k\, \Delta c_l\, \Phi_{kl} + \ldots, \tag{10}$$

where

$$\Phi_k = \frac{\partial \Phi}{\partial c_k}, \quad \Phi_{kl} = \frac{\partial^2 \Phi}{\partial c_k\, \partial c_l} = \Phi_{lk}, \ldots \tag{11}$$

Substitution of this expression into (8) gives directly

$$B'\,\Delta I = 2\mathrm{Re} \sum_k \Delta c_k^* \langle \Phi_k | Q | \Phi \rangle + \\ + \mathrm{Re} \sum_{k,l} \Delta c_k^*\, \Delta c_l^* \langle \Phi_{kl} | Q | \Phi \rangle + \sum_{k,l} \Delta c_k^* \langle \Phi_k | Q | \Phi_l \rangle\, \Delta c_l \\ + \text{terms of higher order.} \tag{12}$$

At this point, we will assume that the variations Δc_k are small and that we will consider the quantity $B' \cdot \Delta I$ up to the second order. Separating the variations into their real and imaginary parts:

$$\Delta c_k = \Delta a_k + i\Delta b_k, \tag{13}$$

one obtains for the first-order variation in (12):

$$B' \cdot \Delta I = \sum_k \{(\Delta a_k - i\Delta b_k) \langle \Phi_k|Q|\Phi\rangle + (\Delta a_k + i\,\Delta b_k) \langle \Phi|Q|\Phi_k\rangle\}$$

$$= \sum_k \Delta a_k \{ \langle \Phi_k|Q|\Phi\rangle + \langle \Phi|Q|\Phi_k\rangle \} -$$

$$- i \sum_k \Delta b_k \{ \langle \Phi_k|Q|\Phi\rangle - \langle \Phi|Q|\Phi_k\rangle \} = 0, \tag{14}$$

for all possible real Δa_k and Δb_k. Hence, the extreme values of I are determined by the conditions

$$\langle \Phi_k|Q|\Phi\rangle = 0, \tag{15}$$

for $k = 1, 2, \ldots p$, which is a generalization of the well-known Brillouin theorem. Here $\Phi_k = \partial\Phi/\partial c_k$ are the first derivatives of Φ at those values of the parameters $\mathbf{c} = \{c_1, c_2, \ldots c_p\}$ which characterize a particular extreme value. In the Hartree-Fock and projected Hartree-Fock schemes (Löwdin *et al* 1981), one starts from a single Slater determinant D built up from N one-electron wave functions ψ_k, which are then varied according to the formula

$$\psi_k' = \psi_k + \Delta c_k \overline{\psi}_k, \tag{16}$$

where $\overline{\psi}_k$ is assumed to be orthogonal to all the functions ψ_k for $k = 1, 2, \ldots N$. It is evident that the relations (15) give the condition for the 'best' trial wave function obtainable from the variation principle.

As a simple example of the linear case, let us introduce a many-particle basis $\boldsymbol{\phi} = \{\phi_k\}$ and let us expand the trial function Φ in the form:

$$\Phi = \sum_k \phi_k c_k. \tag{17}$$

One gets immediately $\Phi_k = \phi_k$ and $\Phi_{kl} = 0$, and the condition (15) gives then

$$\langle \phi_k|Q|\Phi\rangle = \sum_l \langle \phi_k|Q|\phi_l\rangle c_l = 0, \tag{18}$$

which is the standard linear equation system for the coefficients c_l with the approximate eigenvalues I determined by the secular equation:

$$|\langle \phi_k|\,\Omega_1 - I\Omega_2|\phi_l\rangle| = 0. \tag{19}$$

The generalized Brillouin theorem (15), hence, gives a connection between the linear 'configurational interaction' method and the non-linear Hartree-Fock scheme.

2. Stability condition and instabilities

In order to treat the second-order variations in the right-hand side of expression (12), it is convenient to introduce the matrices $\mathbf{A} = \{A_{kl}\}$ and $\mathbf{B} = \{B_{kl}\}$ having elements defined through the relations:

$$A_{kl} = \langle \Phi_k|Q|\Phi_l\rangle, B_{kl} = \langle \Phi_{kl}|Q|\Phi\rangle, \tag{20}$$

and to separate them into their real and imaginary parts:

$$\mathbf{A} = \mathbf{A}_1 + i\mathbf{A}_2, \quad \mathbf{B} = \mathbf{B}_1 + i\mathbf{B}_2. \tag{21}$$

Since Q is a self-adjoint operator, one has $\mathbf{A}\dagger = \mathbf{A}$, which implies:

$$\tilde{\mathbf{A}}_1 = \mathbf{A}_1, \quad \tilde{\mathbf{A}}_2 = -\mathbf{A}_2, \tag{22}$$

where the symbol $\sim$ indicates the transposed matrix with rows and columns interchanged. Since $\mathbf{B}_{kl} = \mathbf{B}_{lk}$, the matrix $\mathbf{B}$ is a symmetric matrix with complex elements, and the relation $\tilde{\mathbf{B}} = \mathbf{B}$ gives:

$$\tilde{\mathbf{B}}_1 = \mathbf{B}_1, \quad \tilde{\mathbf{B}}_2 = \mathbf{B}_2. \tag{23}$$

For the last term in the right-hand side of (12), one obtains now—observing that the imaginary part vanishes identically:

$$\begin{aligned} t_1 &= \sum_{k,l} \Delta c_k^* \langle \Phi_k | Q | \Phi_l \rangle \Delta c_l = \Delta \mathbf{c}\dagger \, \mathbf{A} \, \Delta \mathbf{c} \\ &= (\Delta \tilde{\mathbf{a}} - i\Delta \tilde{\mathbf{b}})(\mathbf{A}_1 + i\mathbf{A}_2)(\Delta \mathbf{a} + i\Delta \mathbf{b}) \\ &= \Delta \tilde{\mathbf{a}} \cdot \mathbf{A}_1 \cdot \Delta \mathbf{a} - \Delta \tilde{\mathbf{a}} \cdot \mathbf{A}_2 \cdot \Delta \mathbf{b} + \Delta \tilde{\mathbf{b}} \cdot \mathbf{A}_2 \cdot \Delta \mathbf{a} + \Delta \tilde{\mathbf{b}} \cdot \mathbf{A}_1 \cdot \Delta \mathbf{b} \\ &= (\Delta \tilde{\mathbf{a}}, \Delta \tilde{\mathbf{b}}) \begin{pmatrix} \mathbf{A}_1 & -\mathbf{A}_2 \\ \mathbf{A}_2 & \mathbf{A}_1 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{a} \\ \Delta \mathbf{b} \end{pmatrix}. \end{aligned} \tag{24}$$

For the other second-order term in the right-hand side of (12), one obtains similarly:

$$\begin{aligned} t_2 &= \mathrm{Re} \sum_{k,l} \Delta c_k^* \langle \Phi_{kl} | Q | \Phi \rangle \Delta c_l^* = \mathrm{Re} \, \Delta c\dagger \, \mathbf{B} \, \Delta \mathbf{c}^* \\ &= \mathrm{Re} \, (\Delta \tilde{\mathbf{a}} - i\Delta \tilde{\mathbf{b}})(\mathbf{B}_1 + i\mathbf{B}_2)(\Delta \mathbf{a} - i\Delta \mathbf{b}) \\ &= \Delta \tilde{\mathbf{a}} \cdot \mathbf{B}_1 \cdot \Delta \mathbf{a} + \Delta \tilde{\mathbf{a}} \cdot \mathbf{B}_2 \cdot \Delta \mathbf{b} + \Delta \tilde{\mathbf{b}} \cdot \mathbf{B}_2 \cdot \Delta \mathbf{a} - \Delta \tilde{\mathbf{b}} \cdot \mathbf{B}_1 \cdot \Delta \mathbf{b} \\ &= (\Delta \tilde{\mathbf{a}}, \Delta \tilde{\mathbf{b}}) \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_2 & -\mathbf{B}_1 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{a} \\ \Delta \mathbf{b} \end{pmatrix}. \end{aligned} \tag{25}$$

Introducing the column vector

$$\mathbf{d} = \begin{pmatrix} \Delta \mathbf{a} \\ \Delta \mathbf{b} \end{pmatrix} \tag{26}$$

of order $2p$ and the *real symmetric matrix*

$$\overline{\mathbf{T}} = \begin{pmatrix} \mathbf{A}_1 + \mathbf{B}_1, & -\mathbf{A}_2 + \mathbf{B}_2 \\ \mathbf{A}_2 + \mathbf{B}_2, & \mathbf{A}_1 - \mathbf{B}_1 \end{pmatrix} \tag{27}$$

of order $2p \times 2p$, one gets finally the following result for the second-order part $(t_1 + t_2)$:

$$B' \cdot I = \tilde{\mathbf{d}} \, \overline{\mathbf{T}} \, \mathbf{d}. \tag{28}$$

The result is structurally the same as in the previous paper (Löwdin *et al* 1981), except that the matrices $\mathbf{A}$ and $\mathbf{B}$ now have different interpretations; we note particularly that $\mathbf{B}$ may have non-vanishing diagonal elements.

The extreme value I corresponding to the conditions (15) is hence a *true minimum*, if

the matrix $\overline{\mathbf{T}}$ is positive definite, i.e. if all the eigenvalues of $\overline{\mathbf{T}}$ are positive. To prove that this is the case is often both difficult and time-consuming, and most of the work so far has instead been devoted to studying instabilities, i.e. extreme values of a different character.

If there are any variations **d**, for which the right-hand side of (28) becomes *negative*, one is dealing with an unstable extreme value or 'instability.' This will, of course, always occur if the matrix $\overline{\mathbf{T}}$ has one or more negative eigenvalues. For real variations $\Delta\mathbf{c}$, it will also occur if the matrix $\mathbf{A}_1+\mathbf{B}_1$ has any negative eigenvalues, and, for pure imaginary variations $\Delta\mathbf{c}$, it will occur if the matrix $\mathbf{A}_1-\mathbf{B}_1$ has any negative eigenvalues, etc. In fact, the number of possibilities turns out to be very large. It remains to be investigated to what extent these instabilities in the case of a general trial function may be classified in the same way as in the Hartree-Fock scheme (Cizek and Paldus 1967, 1970; Fukutome 1968, 1971, 1972, 1973a, b, 1974, 1975; Paldus and Cizek 1969, 1970a, b, 1971; Paldus 1970; Paldus *et al* 1973, 1978; Laforgue *et al* 1973; Laidlaw 1973; Paldus and Veillard 1977, 1978; Ozaki and Fukutome 1978; Ozaki 1979, 1980; Benard and Paldus 1980; Fukutome 1981). It is also illustrative to discuss the standard 'configurational interaction' method as described by (17–19) from this point of view.

3. Search for the extreme values

It is evident that, if the conditions (15) for an extreme value are not satisfied for a specific set of parameter values $\mathbf{c}=\{c_1, c_2, \ldots c_p\}$:

$$f_k=\langle\Phi_k|Q|\Phi\rangle\neq 0, \tag{29}$$

one may use the technique developed here to search for a better approximation $\mathbf{c}'=\mathbf{c}+\Delta\mathbf{c}$ to an extreme value. One gets directly:

$$\begin{aligned} f'_k&=\langle\Phi'_k|Q|\Phi'\rangle=\langle\Phi_k+\Delta\Phi_k|Q|\Phi+\Delta\Phi\rangle\\ &=\langle\Phi_k|Q|\Phi\rangle+\sum_l \Delta c_l^*\langle\Phi_{kl}|Q|\Phi\rangle+\sum_l\langle\Phi_k|Q|\Phi_l\rangle\Delta C_l+\ldots\\ &=f_k+\sum_l B_{kl}\Delta c_l^*+\sum_l A_{kl}\Delta c_l+\ldots=0, \end{aligned} \tag{30}$$

or

$$\mathbf{f}'=\mathbf{f}+\mathbf{A}\,\Delta\mathbf{c}+\mathbf{B}\,\Delta\mathbf{c}^*+\ldots=0. \tag{31}$$

Separating the real and imaginary parts of each term, one obtains with $\mathbf{f}=\mathbf{f}_1+i\,\mathbf{f}_2$:

$$\begin{cases} \mathbf{f}_1+(\mathbf{A}_1+\mathbf{B}_1;\ -\mathbf{A}_2+\mathbf{B}_2)\begin{pmatrix}\Delta\mathbf{a}\\ \Delta\mathbf{b}\end{pmatrix}=0;\\ \mathbf{f}_2+(\mathbf{A}_2+\mathbf{B}_2;\ \mathbf{A}_1-\mathbf{B}_1)\begin{pmatrix}\Delta\mathbf{a}\\ \Delta\mathbf{b}\end{pmatrix}=0; \end{cases} \tag{32}$$

i.e.

$$\begin{pmatrix}\Delta\mathbf{a}\\ \Delta\mathbf{b}\end{pmatrix}=-(\overline{\mathbf{T}})^{-1}\begin{pmatrix}\mathbf{f}_1\\ \mathbf{f}_2\end{pmatrix}. \tag{33}$$

This means that the matrix $\overline{\mathbf{T}}$ and its inverse may be of essential importance also in the search for the extreme values.

In practice, one does not evaluate the inverse of the matrix $\overline{\mathbf{T}}$ but solves the equation system

$$\overline{\mathbf{T}}\begin{pmatrix}\Delta\mathbf{a}\\ \Delta\mathbf{b}\end{pmatrix} = -\begin{pmatrix}\mathbf{f}_1\\ \mathbf{f}_2\end{pmatrix}, \tag{34}$$

by some of the standard methods available. This leads to a better approximation for the vector $\mathbf{c}$, and one may then repeat the procedure, until the conditon $\mathbf{f} = \mathbf{0}$ becomes satisfied with the accuracy desired. Equation (33) gives the connection between some of the current algorithms for finding extreme values of I and the stability conditions.

References

Adams W 1962 *Phys. Rev.* **127** 1650
Bénard M and Paldus J 1980 *J. Chem. Phys.* **72** 6546
Brillouin L 1933 *Actual. Sci. Ind.* **71**
Brillouin L 1934 *Actual. Sci. Ind.* **159**
Cizek J and Paldus J 1967 *J. Chem. Phys.* **47** 3976
Cizek J and Paldus J 1970 *J. Chem. Phys.* **53** 821
Fukutome H 1968 *Prog. Theor. Phys.* **40** 998, 1227
Fukutome H 1971 *Prog. Theor. Phys.* **45** 1382
Fukutome H 1972 *Prog. Theor. Phys.* **47** 1156
Fukutome H 1973 *Prog. Theor. Phys.* **49** 22
Fukutome H 1973 *Prog. Theor. Phys.* **50** 1433
Fukutome H 1974 *Prog. Theor. Phys.* **52** 115, 1766
Fukutome H 1975 *Prog. Theor. Phys.* **53** 1320
Fukutome H 1981 *Int. J. Quantum Chem.* **20** 955
Laforgue A, Cizek J and Paldus J 1973 *J. Chem. Phys.* **59** 2560
Laidlaw W G 1973 *Int. J. Quantum Chem.* **7** 87
Löwdin P O, Calais J L and Calazans J M 1981 *Int. J. Quantum Chem.* **20** 1201
Ozaki M 1979 *Prog. Theor. Phys.* **62** 1183
Ozaki M 1980 *Prog. Theor. Phys.* **63** 84
Ozaki M and Fukutome H 1978 *Prog. Theor. Phys.* **60** 1322
Paldus J and Cizek J 1969 *Prog. Theor. Phys.* **42** 769
Paldus J and Cizek J 1970a *J. Chem. Phys.* **52** 2919
Paldus J and Cizek J 1970b *J. Polymer Sci.* **C29** 199
Paldus J and Cizek J 1970c *Phys. Rev.* **A2** 2268
Paldus J and Cizek J 1971 *J. Chem. Phys.* **54** 2293
Paldus J, Cizek J and Keating B A 1973 *Phys. Rev.* **A8** 640
Paldus J, Cizek J and Laforgue A 1978 *Int. J. Quantum Chem.* **13** 41
Paldus J and Veillard A 1977 *Chem. Phys. Lett.* **50** 6
Paldus J and Veillard A 1978 *Mol. Phys.* **35** 445
Thouless D J 1961 *The quantum mechanics of many-body systems* (New York: Academic Press)