# Variational principles for first-order wave functions



1969 J. Phys. A: Gen. Phys. 2 193

(http://iopscience.iop.org/0022-3689/2/2/006)

J. PHYS. A (GEN. PHYS.), 1969, SER. 2, VOL. 2. PRINTED IN GREAT BRITAIN

# Variational principles for first-order wave functions

P. D. ROBINSON
Mathematics Department, University of York
*MS. received 9th September* 1968

**Abstract.** Complementary variational principles are developed for approximate solutions of the first-order Rayleigh–Schrödinger perturbation correction to the wave equation, yielding upper and lower bounds for the second-order energy correction. The upper bound is the same as Hylleraas's; the complementary lower bound is related to Temple's result for eigenvalues, and (unlike previous lower bounds) is shown to be unconditional. The analysis extends to cover the first-order Brillouin–Wigner correction. As a by-product of the theory it is shown how the Rayleigh–Ritz upper bound and the Temple lower bound for eigenvalues arise in a complementary manner.

## 1. Introduction

In recent years there has been considerable interest in approximate variational solutions of the first-order Rayleigh-Schrödinger perturbation correction

$$(h-\epsilon_0)\Phi = (E_1-V)\psi_0 \tag{1}$$

to the wave equation

$$(h-\epsilon_0)\psi_0 = 0 \tag{2}$$

for the ground state $\psi_0$ of a Hamiltonian $h$. The perturbation is $V$ and the first two orders of correction to the unperturbed ground-state energy $\epsilon_0$ are

$$E_1 = \int \psi_0 V \psi_0 \,\mathrm{d}\boldsymbol{r} \tag{3}$$

$$E_2 = \int \phi(V-E_1)\psi_0 \,\mathrm{d}\boldsymbol{r} \tag{4}$$

where $\psi_0$ is normalized and $\Phi = \phi$ is the exact solution of (1) (we assume real functions throughout for simplicity).

Hylleraas (1930) was the first to exploit a variational principle to obtain approximate solutions of (1); his principle leads to an upper bound for $E_2$ (see also Sharma 1967) and is the basis of the so-called perturbation-variation method (see e.g. Scherr *et al.* 1966). Prager and Hirschfelder (1963, to be referred to as I) discuss a constrained principle which gives a lower bound for $E_2$, while Arthurs and Robinson (1968, to be referred to as II) have developed complementary variational principles which under certain circumstances yield simultaneous upper and lower bounds for $E_2$. Neither of these lower-bound principles is really satisfactory, the Prager–Hirschfelder principle because of the inconvenient constraint which a trial function must satisfy (this cannot be met at all in one-dimensional situations, see II) and the Arthurs–Robinson principle because it is not usually clear *a priori* whether it will give a lower bound for $E_2$ at all.

In the present paper, which is inspired by the work of Pomraning (1967), we give alternative complementary principles which provide a lower bound free from any restrictions. As in II, the upper bound is identical with that of Hylleraas. The new feature is the decomposition of the operator $(h-\epsilon_0)$ into the form

$$(h-\epsilon_0) = (\epsilon_1-\epsilon_0)+(h-\epsilon_1) \tag{5}$$

where $\epsilon_1$ is the unperturbed first excited-state energy. As a by-product of the theory we see how the Rayleigh–Ritz upper bound and the Temple lower bound for $\epsilon_0$ arise in a complementary manner.

Some of the analysis extends to the Brillouin–Wigner equation

$$(h-E)\Phi = (E_1 - V)\psi_0 \tag{6}$$

which is similar to (1) but more difficult to deal with (Meath and Hirschfelder 1964).

## 2. Basic theory

It is convenient to consider an inhomogeneous equation of basic type

$$(q+T^{\dagger}T)\Phi = f \tag{7}$$

where $q$ and $f$ are functions of coordinates, $T$ is a linear operator and $T^{\dagger}$ is its adjoint defined by

$$\int (T^{\dagger}U)\Phi \, \mathrm{d}\boldsymbol{r} = \int U(T\Phi)\, \mathrm{d}\boldsymbol{r}. \tag{8}$$

First we obtain the variational principles appropriate to (7) and then, in subsequent sections, we look at the different ways in which equations (1) and (7) can be identified with each other.

The procedure is to decompose (7) into a pair of canonical Euler equations

$$T\Phi = U = \frac{\partial H}{\partial U} \tag{9}$$

$$T^{\dagger}U = f - q\Phi = \frac{\partial H}{\partial \Phi} \tag{10}$$

with generalized 'classical Hamiltonian'

$$H = \tfrac{1}{2}U^2 - \tfrac{1}{2}q\Phi^2 + \Phi f. \tag{11}$$

Then variation of the generalized 'action' functional

$$S(\Phi, U) = \int \{U(T\Phi) - H\}\, \mathrm{d}\boldsymbol{r} = \int \{(T^{\dagger}U)\Phi - H\}\, \mathrm{d}\boldsymbol{r} \tag{12}$$

round $\Phi = \phi$ and $U = u = T\phi$ leads to complementary functionals

$$G(\Phi) \equiv S(\Phi, U(\Phi)) = -\tfrac{1}{2}\int \{-\Phi(q+T^{\dagger}T)\Phi + 2\Phi f\}\, \mathrm{d}\boldsymbol{r} \tag{13}$$

$$J(U) \equiv S(\Phi(U), U) = -\tfrac{1}{2}\int \{q^{-1}f^2 + U^2 + (T^{\dagger}U - 2f)q^{-1}(T^{\dagger}U)\}\, \mathrm{d}\boldsymbol{r} \tag{14}$$

which are stationary at $(\phi, u)$ with value

$$S(\phi, u) = -\tfrac{1}{2}\int \phi f \, \mathrm{d}\boldsymbol{r}. \tag{15}$$

The functional $G(\Phi)$ depends on a trial function $\Phi$ with the corresponding $U$ given by (9), whereas $J(U)$ depends on a trial $U$ and $\Phi$ obtained from (10).

We note that

$$G(\Phi) - S(\phi, u) = \tfrac{1}{2}\int (\Phi - \phi)(q+T^{\dagger}T)(\Phi - \phi)\, \mathrm{d}\boldsymbol{r} \tag{16}$$

and

$$S(\phi, u) - J(U) = \tfrac{1}{2}\int (U-u)(1+Tq^{-1}T^{\dagger})(U-u)\, \mathrm{d}\boldsymbol{r}. \tag{17}$$

Thus

$$(q+T^{\dagger}T) \text{ is positive definite} \Leftrightarrow G \geqslant S \tag{18}$$

and

$$(1+Tq^{-1}T^{\dagger}) \text{ is positive definite} \Leftrightarrow S \geqslant J. \tag{19}$$

So in order to obtain the complementary upper and lower bounds

$$G(\Phi) \geqslant S(\phi, u) \geqslant J(U) \tag{20}$$

it is necessary and sufficient that the operators $(q+T^{\dagger}T)$ and $(1+Tq^{-1}T^{\dagger})$ be positive definite. In particular, we see that (20) holds if

$$q > 0. \tag{21}$$

The functional

$$J(T\Theta) = -\tfrac{1}{2}\int \{q^{-1}f^2+q^{-1}(TT^{\dagger}\Theta)^2+(\Theta-2q^{-1}f)(TT^{\dagger}\Theta)\}\,\mathrm{d}\mathbf{r}$$

$$= G(\Theta)-\tfrac{1}{2}\int q^{-1}\{f-(q+T^{\dagger}T)\Theta\}^2\,\mathrm{d}\mathbf{r} \tag{22}$$

is slightly less general than (14), but is more convenient since it does not require a knowledge of the individual $T$ and $T^{\dagger}$.

## 3. The Hylleraas upper bound for $E_2$

If we set

$$f = (E_1-V)\psi_0 \tag{23}$$

and

$$(q+T^{\dagger}T) = (h-\epsilon_0) \tag{24}$$

so that equations (1) and (7) become the same, we find that

$$2S(\phi, u) = E_2 \tag{25}$$

and

$$2G(\Phi) = \int \{\Phi(h-\epsilon_0)\Phi-2\Phi(E_1-V)\psi_0\}\,\mathrm{d}\mathbf{r} = \mathrm{Hyl}(\Phi) \tag{26}$$

which is the Hylleraas functional. Since $(h-\epsilon_0)$ is positive definite it follows from (18) and (24) that

$$\mathrm{Hyl}(\Phi) \geqslant E_2. \tag{27}$$

## 4. Conditional lower bounds for $E_2$

Arthurs and Robinson (II) considered the situation when

$$h = -\tfrac{1}{2}\nabla^2+v \tag{28}$$

so that

$$(h-\epsilon_0) = -\tfrac{1}{2}\,\mathrm{div}\,\mathrm{grad}+(v-\epsilon_0) \tag{29}$$

which is immediately identifiable with (24) when the choice of adjoint operators

$$T = 2^{-1/2}\,\mathrm{grad}, \qquad T^{\dagger} = -2^{-1/2}\,\mathrm{div} \tag{30}$$

is made. In this case

$$q = (v-\epsilon_0) \tag{31}$$

and in order that $2J(U)$ should provide a lower bound for $E_2$ which is complementary to $\mathrm{Hyl}(\Phi)$ it follows from (19) that the operator

$$1+\tfrac{1}{2}\,\mathrm{grad}(v-\epsilon_0)^{-1}\,\mathrm{div} \tag{32}$$

should be positive definite. This condition depends crucially on the potential $v$, and evidently does not hold in general.

Prager and Hirschfelder (I) use (28) to rewrite (1) as

$$-\,\mathrm{div}\{\psi_0{}^2\,\mathrm{grad}(\Phi/\psi_0)\} = 2(E_1-V)\psi_0{}^2 \tag{33}$$

which can actually be regarded as an example of (7) with $\Phi/\psi_0$ for $\Phi$ and

$$q = 0. \tag{34}$$

In this case the Hylleraas upper bound is still obtained (II), and the $J$ functional always furnishes a lower bound since there is now no $q^{-1}$ term in (19). But from (10) we see that the trial function $U$ is subject to the constraint

$$T^{\dagger}U = f = 2(E_1 - V)\psi_0{}^2 \tag{35}$$

which can be difficult to meet (II); in particular a trial

$$U = T\Theta \tag{36}$$

would not be allowed unless $\Theta$ provided the exact solution.

## 5. Unconditional complementary bounds for $E_2$

### 5.1. *An unconditional lower bound*

To obtain an unconditional lower bound for $E_2$ which is complementary to the upper bound Hyl($\Phi$), we employ the decomposition (5) and think of (1) in the form

$$\{(\epsilon_1 - \epsilon_0) + (h - \epsilon_1)\}\Phi = (E_1 - V)\psi_0 \tag{37}$$

noting that:

(i) the right-hand side of (37) is orthogonal to $\psi_0$;

(ii) the solution $\Phi = \phi$ is not unique, for if $\phi$ is a solution then so is $\phi$ plus a constant multiple of $\psi_0$.

Thus we may consider (37) as an equation on the domain $D_0$ orthogonal to $\psi_0$, when it will have a unique solution $\phi$. The operator $(h - \epsilon_1)$ is positive definite on $D_0$ and so we can set

$$(\epsilon_1 - \epsilon_0) = q, \text{ a positive number} \tag{38}$$

and

$$(h - \epsilon_1) = T^{\dagger}T \tag{39}$$

without needing to know specifically the nature of $T$ and $T^{\dagger}$; any positive definite, self-adjoint operator can be decomposed in such form (Mikhlin 1964). The analysis of §§ 2 and 3 (equations (21), (22), (27)) now shows that

$$\begin{aligned} \mathrm{Hyl}(\Phi) &= \int \{\Phi(h - \epsilon_0)\Phi - 2\Phi(E_1 - V)\psi_0\}\, d\mathbf{r} \\ &\geqslant E_2 \geqslant \mathrm{Hyl}(\Theta) - \frac{1}{\epsilon_1 - \epsilon_0}\int \{(h - \epsilon_0)\Theta - (E_1 - V)\psi_0\}^2\, d\mathbf{r}. \end{aligned} \tag{40}$$

It is evident that any $\psi_0$ component in $\phi$ or in the trial functions $\Phi$ and $\Theta$ would not contribute to any term in (40), and so the restriction that $\Phi, \Theta \in D_0$ can be relaxed.

In practice, if $\epsilon_1$ is not known precisely, it can always be replaced by any quantity $\gamma$ which satisfies

$$\epsilon_0 < \gamma < \epsilon_1. \tag{41}$$

### 5.2. *Connections with Temple's formula*

If $\mathscr{E}_0$ and $\mathscr{E}_1$ are the two lowest eigenvalues of a Hamiltonian $H$, Temple's (1928) lower bound formula states that

$$\mathscr{E}_0 \geqslant \int \chi H \chi\, d\mathbf{r} - \frac{\int \chi H^2 \chi\, d\mathbf{r} - \{\int \chi H \chi\, d\mathbf{r}\}^2}{\mathscr{E}_1 - \int \chi H \chi\, d\mathbf{r}} \tag{42}$$

provided that $\chi$ is normalized and also assuming that

$$\int \chi H \chi\, d\mathbf{r} < \mathscr{E}_1. \tag{43}$$

By taking

$$H = h + V \tag{44}$$

$$\chi = \psi_0 + \phi \tag{45}$$

$$\mathscr{E}_0 = \epsilon_0 + E_1 + E_2 \tag{46}$$

and expanding (42) in orders of $V$, the lower bound in (40) can be derived from (42) (see I). However, our derivation shows that this lower bound holds *independently of the restriction* (43) (or of any other restriction), and that it is complementary to the Hylleraas upper bound.

An interesting by-product of the complementary bounds (40) is obtained by setting $V = 0$: we obtain

$$\int \Phi(h-\epsilon_0)\Phi \, \mathrm{d}\boldsymbol{r} \geqslant 0 \geqslant \frac{1}{\epsilon_1-\epsilon_0} \int \Theta(h-\epsilon_0)(\epsilon_1-h)\Theta \, \mathrm{d}\boldsymbol{r}. \tag{47}$$

The left-hand member of (47) gives the Rayleigh–Ritz upper bound for $\epsilon_0$, and the right-hand member leads to Temple's lower bound for $\epsilon_0$ (cf. (42), with $\mathscr{E}_0 = \epsilon_0$, $\chi = \Theta$, $H = h$). Thus we see how the Rayleigh–Ritz and Temple bounds for $\epsilon_0$ arise in a complementary manner.

5.3. *Scaling and Ritz procedures*

The complementary bounds (40) can always be improved by scaling (I) or Ritz procedures because of their quadratic nature. The resulting formulae are simplest if we expand

$$\Phi = \sum_{n=1}^{m} a_n \psi_n, \qquad \Theta = \sum_{n=1}^{m} b_n \psi_n \tag{48}$$

in terms of the normalized eigenfunctions $\{\psi_n\}$ of $h$. Optimization with respect to the coefficients $a_n$ and $b_n$ yields

$$-\sum_{n=1}^{m} \frac{V_{n0}{}^2}{\epsilon_n-\epsilon_0} \geqslant E_2 \geqslant -\frac{1}{\epsilon_1-\epsilon_0}\left\{\int \psi_0(V-E_1)^2\psi_0 \, \mathrm{d}\boldsymbol{r} - \sum_{n=1}^{m}\left(\frac{\epsilon_n-\epsilon_1}{\epsilon_n-\epsilon_0}\right) V_{n0}{}^2\right\} \tag{49}$$

where

$$V_{n0} = \int \psi_n V \psi_0 \, \mathrm{d}\boldsymbol{r}. \tag{50}$$

More general formulae can be obtained with other basis sets. The left-hand member of (49) is well known, and taking $m = 1$ on the right gives a result due to Dalgarno (1961). Goodisman (1967) discusses various averaging processes which can be applied to similar formulae.

5.4. *Possible alternative lower bounds*

In certain circumstances other ways of meeting the decomposition (24) may be possible. If there exists a self-adjoint operator $L$ (which could be merely a function of coordinates) which is such that:

(i) $(h-L)$ is positive definite on $\mathrm{D}_0$;
(ii) $(L-\epsilon_0)$ is strictly positive on $\mathrm{D}_0$;
(iii) $(L-\epsilon_0)^{-1}$ exists;

then the decomposition

$$L-\epsilon_0 = q \tag{51}$$

$$h-L = T^\dagger T \tag{52}$$

would lead to the lower bound

$$E_2 \geqslant \mathrm{Hyl}(\Theta) - \int \{(h-\epsilon_0)\Theta-(E_1-V)\psi_0\}(L-\epsilon_0)^{-1}\{(h-\epsilon_0)\Theta-(E_1-V)\psi_0\} \, \mathrm{d}\boldsymbol{r}. \tag{53}$$

(The analysis of § 2 is readily modified to admit the possibility of $q$ being an operator; care must be taken with the order of certain terms.)

## 6. The Brillouin–Wigner equation

In the Brillouin–Wigner equation (6), $E$ is the energy of the total Hamiltonian $(h+V)$. It is customary to retain the definition (3) for $E_1$, in which case we see from (6) that

$$0 = \int \psi_0(E_1 - V)\psi_0 \, d\mathbf{r} = \int \psi_0(h-E)\Phi \, d\mathbf{r} = (\epsilon_0 - E)\int \psi_0 \Phi \, d\mathbf{r}. \tag{54}$$

Thus, if now $\Phi = \phi$ is the solution of (6), we have $\phi \in D_0$ and also $\phi$ is unique. Provided that

$$\epsilon_1 > E \tag{55}$$

then we can think of (6) in the form

$$\{(\epsilon_1 - E) + (h - \epsilon_1)\}\Phi = (E_1 - V)\psi_0 \tag{56}$$

and by an analogy with § 5.1. obtain complementary upper and lower bounds for the second-order Brillouin–Wigner energy

$$E^{(2)} = \int \phi V \psi_0 \, d\mathbf{r}. \tag{57}$$

With the condition

$$\Phi, \Theta \in D_0 \tag{58}$$

imposed on the trial functions, the bounds are

$$\int \{\Phi(h-E)\Phi + 2\Phi V\psi_0\} \, d\mathbf{r} \geqslant E^{(2)}$$

$$\geqslant \int \{\Theta(h-E)\Theta + 2\Theta V \psi_0\} \, d\mathbf{r} - \frac{1}{\epsilon_1 - E}\int \{(h-E)\Theta - (E_1 - V)\psi_0\}^2 \, d\mathbf{r}. \tag{59}$$

The formulae (49) are adapted to this situation if we replace $\epsilon_0$ by $E$ and $E_2$ by $E^{(2)}$. In practice $E$ may not be known exactly (unless from experiment), but it can play the role of a parameter to be adjusted iteratively (Meath and Hirschfelder 1964).

## 7. Discussion

We have derived complementary upper and lower bounds for $E_2$, the second-order energy correction for ground states. The upper bound is the same as Hylleraas's, and the lower bound is related to Temple's lower bound for eigenvalues. We have shown that this lower bound is (unlike others) an unconditional one, and in that sense the result is new. The closeness of these complementary bounds measures the accuracy of approximate solutions for the first-order wave-function correction. The same functional form can be used as a trial solution for each bound, and the disposable parameters chosen to optimize each bound separately. If the optimum bounds are close, and the pairs of optimum parameter values are close, then the solution is a good one.

As a simple illustrative example, consider the Stark effect for hydrogen where

$$\psi_0 = \pi^{-1/2} e^{-r}, \qquad V = -z, \qquad \epsilon_0 = -\tfrac{1}{2}, \qquad E_1 = 0. \tag{60}$$

Assuming that

$$\Phi = A\pi^{-1/2} z \, e^{-\alpha r}, \qquad \Theta = B\pi^{-1/2} z \, e^{-\beta r} \tag{61}$$

we obtain the bounds

$$-2{\cdot}238 \geqslant E_2 \geqslant -2{\cdot}271 \tag{62}$$

with the optimum parameter values

$$A = 1{\cdot}310, \alpha = 0{\cdot}797; \qquad B = 1{\cdot}455, \beta = 0{\cdot}662. \tag{63}$$

The exact result is

$$E_2 = -2{\cdot}25, \qquad \phi = \pi^{-1/2}(1+\tfrac{1}{2}r)z\,e^{-r}. \tag{64}$$

It would be of interest to compare the lower bounds and corresponding $\Theta$'s with the Hylleraas upper bounds and $\Phi$'s which have been used in various extensive calculations (cf. Scherr *et al.* 1966).

The results in this paper can be generalized to cover excited states (with suitable restrictions on the classes of trial functions, cf. Sharma 1967), or extended to deal with higher-order corrections to the wave equation.

## References


ARTHURS, A. M., and ROBINSON, P. D., 1968, *Proc. R. Soc.* A, **303**, 503–9.

DALGARNO, A., 1961, *Quantum Theory*, ed. D. R. Bates (New York: Academic Press) **1**, Ch. 5.

GOODISMAN, J., 1967, *J. Chem. Phys.*, **47**, 2707–11.

HYLLERAAS, E. A., 1930, *Z. Phys.*, **65**, 209–25.

MEATH, W. J., and HIRSCHFELDER, J. O., 1964, *J. Chem. Phys.*, **41**, 1628–33.

MIKHLIN, S. G., 1964, *Variational Methods in Mathematical Physics* (New York: Macmillan).

POMRANING, G. C., 1967, *J. Math. Phys.*, **8**, 2096–108.

PRAGER, S., and HIRSCHFELDER, J. O., 1963, *J. Chem. Phys.*, **39**, 3289–94.

SCHERR, C. W., SANDERS, F. C., and KNIGHT, R. E., 1966, *Perturbation Theory and its Applications in Quantum Mechanics*, ed. C. H. Wilcox (New York: John Wiley), 97–118.

SHARMA, C. S., 1967, *Proc. Phys. Soc.*, **91**, 50–3.

TEMPLE, G., 1928, *Proc. R. Soc.* A, **119**, 276–93.