THE VARIATION METHOD IN QUANTUM CHEMISTRY

SAUL T. EPSTEIN

PHYSICAL CHEMISTRY

A Series of Monographs

The Variation Method in Quantum Chemistry

SAUL T. EPSTEIN
Theoretical Chemistry Institute and Department of Physics
University of Wisconsin
Madison, Wisconsin

1974

ACADEMIC PRESS New York San Francisco London
A Subsidiary of Harcourt Brace Jovanovich, Publishers

To my mother and father, secure in the knowledge that they would have liked it sight unseen.

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1

Library of Congress Cataloging in Publication Data

Epstein, Saul T
The variation method in quantum chemistry.

(Physical chemistry, a series of monographs,)
Bibliography: p.
1. Quantum chemistry. 2. Calculus of variations.
I. Title. II. Series.
QD462.E67 541'.28 73-18960
ISBN 0-12-240550-1

PRINTED IN THE UNITED STATES OF AMERICA

The Variation Method in Quantum Chemistry

This is Volume 33 of
PHYSICAL CHEMISTRY
A series of monographs
Edited by ERNEST M. LOEBL, *Polytechnic Institute of New York*

A complete list of the books in this series appears at the end of the volume.

Preface

What follows is very far from being a treatise on the variation method. Rather it represents an attempt to bring together into a coherent whole those aspects of the variation method as applied to bound states problems in quantum mechanics which have been of especial interest to me. Hopefully the results will be of some interest to others as well.

A fair appraisal of the contents of the book can be gained simply by inspecting the table of contents. As will be gathered, most of the text is devoted to generalities, general points of view, general theorems, etc. To balance this to some extent, we do describe in detail, in Sections 6, 8, and 9, three standard variational procedures which then serve in Sections 27–30 and also here and there in other sections as concrete examples to which we can apply the general theorems. Also we give a number of references to papers in the literature where other examples may be found.

Since, as will be seen, the level of discussion, both as regards physics and mathematics, is fairly elementary (and certainly should be readily accessible to anyone with two semesters of quantum mechanics), we have omitted many steps, while at the same time making frequent use of phrases like "it is obvious that," "it is easily shown," or "one readily finds," with the expectation that the reader will be alerted by them to work out the omitted steps. Also, at the ends of sections we have appended selections of problems, some of them asking the reader more explicitly to work out details, others indicating extensions and generalizations of results in the text.

Although my general interest in the variation method certainly predated my arrival at the University of Wisconsin, the great intensification of that interest in recent years owes much to the stimulation provided by my association with Professor Joseph O. Hirschfelder, and with the Theoretical Chemistry Institute generally. In particular, in addition to acknowledging my great debt to Professor Hirschfelder, I would like also to express my thanks to Professors W. Byers Brown, P. R. Certain, and P. D. Robinson for many discussions of these matters through the years. Also special thanks are due to my wife, Jean Hoopes Epstein, who has assisted in every phase of the preparation of this manuscript. Finally thanks are due to the National Aeronautics and Space Administration and the National Science Foundation which have supported the Institute.

Contents

Chapter I **General Theory of the Variation Method**

Chapter II **Applications of the Variation Method**

Chapter III **The Generalized Brillouin Theorem**

Chapter IV **Special Theorems Satisfied by Optimal Trial Functions**

Chapter V **Perturbation Theory and the Variation Method: General Theory**

Chapter VI **Perturbation Theory and the Variation Method: Applications**

Chapter VII **The Hylleraas Variation Method**

Chapter VIII **Special Theorems Satisfied by Optimal First-Order Trial Functions**

Chapter IX **Corrections to Approximate Calculations**

Chapter I / General Theory of the Variation Method

1. Some Background

Principles of "least this" and "maximum that" have for a long time fascinated scientists and philosophers alike.[1] Moreover, even principles of "stationary this or that," though they may have somewhat less philosophical appeal, nonetheless continue to interest scientists because on the one hand they usually provide a very compact way of stating the mathematical essentials of a theory, and on the other provide a useful avenue to the development of new theories.

Thus it is not surprising that in his first paper on wave mechanics [3], Schrödinger presented his equation in the form of a variation principle, and only later [4] revealed some of the background which led to his writing down of the equation in the first place.

An obvious reason for the philosophical fascination of minimal (or maximal) principles is that they seem to make physical laws less arbitrary, more rational, suggest a purpose, etc. Indeed, it is tempting to take a variation principle quite literally and then imagine that the variations are actually taking place; that at each moment or at each point, the system, whatever it is, is sampling all sorts of possible behavior, with the actual behavior then being selected on the basis that it will make the least change in something, that is, it will make some quantity stationary. Moreover, Feynman [5], following on some earlier work of Dirac, has shown that in a very real sense quantum mechanics can be "derived" from classical mechanics by omitting the final selection processes and assuming that all behavior, not just the classical one, is possible.

In a not dissimilar vein, Ruedenberg and co-workers [6] have argued that one can get real insight into the nature of chemical binding by imagining that as a molecule forms, it actually does, so to speak, try one wave function,

---

[1] See, for example, Yourgrau and Mandelstam [1] and Born [2].

then another, relaxing a bit here, tightening a bit there, before finding the most suitable one.

Although these general theoretical and philosophical considerations concerning variation principles are of great importance, probably of more practical significance are the associated variation methods for approximating the solutions of actual physical problems. The use of such methods has a long history in science,[2] and in particular it is not much of an exaggeration to say that in applied quantum mechanics, most approximation procedures are either direct applications of the variation method or can profitably be related to it in one way or another.

In this book, we will be concerned with the variation method for approximating the energy eigenvalues and energy eigenfunctions of a bound quantum mechanical system. Thus we consider situations in which at least the low lying energy levels are discrete, and we concentrate our attention on these discrete levels. More particularly, our standard example will be that of *N* electrons interacting with one another and with various fixed nuclei via Coulomb interactions, and possibly also with external static electric and magnetic fields. Thus the Hamiltonian for the system, including various nuclear interactions as well, is

$$H = \sum_{s=1}^{N} \left\{ \frac{[\mathbf{p}_s + (1/c)\mathscr{A}(\mathbf{r}_s)]^2}{2} - \sum_a \frac{\mathscr{Z}_a}{|\mathbf{r}_s - \mathbf{R}_a|} - \Phi(\mathbf{r}_s) \right\} + \sum_{\substack{s=1 \\ s>t}}^{N} \sum_{t=1}^{N} \frac{1}{|\mathbf{r}_s - \mathbf{r}_t|} + \sum_{a>b}\sum \frac{\mathscr{Z}_a \mathscr{Z}_b}{|\mathbf{R}_a - \mathbf{R}_b|} + \sum_a \mathscr{Z}_a \Phi(\mathbf{R}_a) \tag{1}$$

where we have used atomic units [mass of electron = −charge of electron = (Planck's constant)/2π = 1]. In Eq. (1), c is the velocity of light, $\mathscr{A}$ is the vector potential, and Φ is the scalar potential; the electric and magnetic fields $\mathscr{E}$ and $\mathscr{B}$ are then given by

$$\mathscr{E} = -\boldsymbol{\nabla}\Phi \tag{2}$$

$$\mathscr{B} = \boldsymbol{\nabla} \times \mathscr{A} \tag{3}$$

Also, $\mathbf{r}_s$ is the coordinate of the sth electron, $\mathbf{p}_s$ is the associated momentum operator, $\mathbf{R}_a$ is the coordinate of the ath nucleus, and $\mathscr{Z}_a$ is its charge.

The Hamiltonian (1) is, of course, the Hamiltonian for a molecule in the so-called "clamped nuclei approximation" with complete neglect of all spin (electronic and nuclear) and relativistic effects. As such, it provides

---

[2] See, for example, Mikhlin [7], Courant [8], Finlayson [9], and Birkhoff [10].

a conceptually simple but physically meaningful example to which we can apply the results of our general considerations. Nevertheless, it should be clear to the reader, without our making the point each time, that many of our general conclusions are quite independent of the detailed form of H. Also, although our language will be appropriate to the coordinate representation, most results will be representation independent, holding equally well in momentum space, for example, with appropriate interpretation of the symbols.

Our general notation will be quite a standard one. The scalar product of two wave functions ψ_1 and ψ_2 will be denoted by

$$(\psi_1, \psi_2) \tag{4}$$

with the "integration," for the N-electron example, being over both space and spin. It has the standard properties:

$$(\psi_1, \psi_2) = (\psi_2, \psi_1)^*, \quad (\psi_1, \alpha\psi_2) = \alpha(\psi_1, \psi_2), \quad (\alpha\psi_1, \psi_2) = \alpha^*(\psi_1, \psi_2) \tag{5}$$

where α is a number and the asterisk denotes complex conjugate. Also, we will assume that

$$(\psi, \psi) = 0 \tag{6}$$

implies that

$$\psi \equiv 0 \tag{7}$$

(that is, we will not worry about sets of measure zero). Finally, the statement that an operator θ is Hermitian means that

$$(\psi_1, \theta\psi_2) = (\theta\psi_1, \psi_2) \tag{8}$$

In our discussions, we will have frequent use for matrix and vector notation. A particular vector will usually be distinguished by a subscript in parentheses; thus $V_{(j)}$ with components $V_{(j)k}$ or $V_{k(j)}$, and the scalar product of two vectors will be denoted by

$$W_{(k)}^{+} \cdot V_{(j)} \equiv \sum_{l} W_{(k)l}^{*} V_{(j)l} \tag{9}$$

If this scalar product multiplies another vector X, we will, however, usually write the result as $(W_{(k)}^{+} \cdot V_{(j)})X$ rather than simply $W_{(k)}^{+} \cdot V_{(j)}X$. Finally, as usual, a Hermitian matrix is one for which

$$M_{lj} = M_{jl}^{*} \tag{10}$$

and which therefore has the property

$$V_{(j)}^{+} \cdot MV_{(i)} = (MV_{(j)})^{+} \cdot V_{(i)} \tag{11}$$

while a symmetric matrix is one for which

$$M_{lj} = M_{jl} \tag{12}$$

PROBLEMS

1. What is the numerical value of c in atomic units?
2. If in atomic units $\mathscr{E} = \mathscr{B} = 1$, what are the field strengths in more practical units? Are they large or small by laboratory standards?
3. Presumably you are familiar with the standard properties of Hermitian operators and matrices and of their eigenfunctions and eigenvalues—that they can be diagonalized by a unitary transformation, the orthogonality of the eigenfunctions, reality of the eigenvalues, etc. If not, please consult almost any graduate level quantum mechanics text. Also, you should be familiar with the fact that the average value of any Hermitian operator is greater than or equal to its smallest eigenvalue.
4. A Hermitian matrix M is said to be positive definite if all its eigenvalues are positive. Show that a necessary and sufficient condition for this is that $X^{+} \cdot MX > 0$ for all X.
5. Show that a positive-definite Hermitian matrix always has an inverse (which is also positive definite), and a positive square root.

2. The Variation Principle

Given any function ψ for which the requisite integrals exist (we will refer to such functions as *trial functions*), we can calculate the real number

$$E \equiv (\psi, H\psi)/(\psi, \psi) \tag{1}$$

Since E would be the average energy of the system if the system were in the state described by the function ψ, we have that E cannot be less than the smallest possible energy, that is, E cannot be smaller than the smallest eigenvalue of H.

To learn more, we consider another trial function ψ' and Δ, the difference between ψ and ψ'; thus

$$\psi = \psi' + \Delta \tag{2}$$

Then with the definition

$$E' \equiv (\psi', H\psi')/(\psi', \psi') \tag{3}$$

one finds upon inserting (2) into (1) that

$$E = \frac{E'(\psi', \psi') + (\psi', H\Delta) + (\Delta, H\psi') + (\Delta, H\Delta)}{(\psi, \psi)}$$

or, since

$$(\psi', \psi') + (\psi', \Delta) + (\Delta, \psi') + (\Delta, \Delta) = (\psi, \psi) \tag{4}$$

we have

$$E = E' + \frac{(\psi', (H - E')\Delta) + (\Delta, (H - E')\psi') + (\Delta, (H - E')\Delta)}{(\psi, \psi)}$$

Finally, using the hermiticity of $(H - E')$, we rewrite this as

$$E = E' + \frac{((H - E')\psi', \Delta) + (\Delta, (H - E')\psi') + (\Delta, (H - E')\Delta)}{(\psi, \psi)} \tag{5}$$

We will now draw several important conclusions from this result.

Result 1. Suppose that

$$(H - E')\psi' = 0 \tag{6}$$

that is, suppose that ψ' and E' are an eigenfunction and the corresponding eigenvalue of H. Then (5) becomes

$$E = E' + \frac{(\Delta, (H - E')\Delta)}{(\psi, \psi)} \tag{7}$$

from which we see that E differs from E' by terms which are of at least second order in the difference between ψ and ψ'. Therefore if we think of ψ as being a continuously variable quantity, then it follows that *the eigenvalues of H are stationary points of E as a functional of ψ.*

Result 2. We will now show that E has no other stationary points. Thus suppose that E' is a stationary point of E as a functional of ψ. Then the first-order terms on the right hand side of (5) must vanish for all $\varDelta$ and therefore the first-order terms in

$$((H - E')\psi', \varDelta) + (\varDelta, (H - E')\psi')$$

must vanish for all $\varDelta$. In particular this must be true for

$$\varDelta = \alpha(H - E')\psi' \tag{8}$$

where α is a continuously variable number. Thus we must have

$$(\alpha + \alpha^*)((H - E')\psi',\ (H - E')\psi') = 0$$

which can be satisfied only if

$$(H - E')\psi' = 0$$

Therefore *if E' is a stationary point, then E' is an eigenvalue and the corresponding ψ is an eigenfunction.* The characterization of the eigenvalues and eigenfunctions of H provided by Results 1 and 2 constitutes a statement of the *variation principle.*

Result 3. Suppose now that E' is the smallest eigenvalue of H. Then $(H - E')$ has only nonnegative eigenvalues and thus its average value is always nonnegative. In particular, then, this means that however large $\varDelta$ may be, still

$$(\varDelta, (H - E')\varDelta) > 0$$

and hence from (7) we see, consistent with our initial observation, that *the smallest eigenvalue of H* is not just a stationary point, *it is the absolute minimum of E as a functional of ψ.* Conversely, *E is always an upper bound to the smallest eigenvalue.*

On the other hand, if E' is a higher eigenvalue, then by choosing $\varDelta$ to be a linear combination of the lower eigenfunctions (still higher eigenfunctions—and this can include continuum states if there are any), we can evidently make $(\varDelta, (H - E')\varDelta)$ less than (greater than) zero. Thus the intermediate eigenvalues of H are only saddle points of E as a functional of ψ, and even locally are neither maxima nor minima.

Result 4. That the smallest eigenvalue is the absolute minimum of E is a very striking result. However, it does not in general serve to characterize the energies of the actual ground states of atoms or molecules since, because of the requirements of the Pauli principle, these ground states are usually not the lowest eigenstates of the appropriate Hamiltonians; for example, the ground state of the lithium atom is $(1s)^2 2s$ and not $(1s)^3$.

Happily, however, there is a similar theorem which is applicable to the physical ground states and to the various excited states as well: often the eigenstates of H can be classified into types according to their symmetries or, more generally, according to the eigenvalues of other operators K which commute with H. If then E' is the smallest eigenvalue of H for states of a certain type (for example, for states satisfying the Pauli principle), it follows that $(H - E')$ will have a nonnegative expectation value with respect to functions of that type because such functions will be orthogonal to all eigenfunctions of H associated with smaller eigenvalues.

If now we confine attention to ψ of that same type, then $\Delta = \psi - \psi'$ will also be of that type and therefore however large Δ may be, still

$$(\Delta, (H - E')\Delta) \geq 0$$

Thus *the smallest eigenvalue of H for states of a given type is the absolute minimum of E as a functional of trial functions of that type*, and conversely, if one uses only trial functions of that type, then *E is always an upper bound to that lowest eigenvalue.*

Result 5. We now observe that if H commutes with K, it will follow that if ψ' is of a definite type, then the Δ, which played the decisive role in Result 2, namely $(H - E')\psi'$, will also be of that type. Thus we may generalize the result found there as follows: If H commutes with K, if ψ' is of a certain type, and if E' is a stationary point with respect to all variations of that type, then E' is an eigenvalue and ψ' is an eigenfunction. In summary, combining this last result with Result 1, *the variation principle applies separately to the eigenfunctions and eigenvalues of each type.*

Result 6. A further generalization of Result 4 is evidently the following. Let E' be an arbitrary eigenvalue of H and confine attention to ψ which are orthogonal to all eigenfunctions of H whose associated eigenvalues are less than E'. Then we will have

$$(\Delta, (H - E')\Delta) \geq 0$$

and thus *an arbitrary eigenvalue of H is an absolute minimum of E as a functional of trial functions orthogonal to eigenfunctions associated with smaller eigenvalues.*

Before concluding this section, it is appropriate to make some remarks about the existence of the integrals in (1) and hence about the functions which can be trial functions. For one thing, of course, as befits a bound state function, they must be normalizable so that the denominator exists. Further, if we are working in configuration space, then, for the usual Hamiltonians of atomic and molecular physics, for example for (1-1), the existence of $(\psi, H\psi)$ requires that ψ be twice differentiable. However, if one uses the formally equivalent $(\boldsymbol{\nabla}\psi, \cdot\boldsymbol{\nabla}\psi)$ instead of $(\psi, -\boldsymbol{\nabla}^2\psi)$ it can be shown [11] that the results which we have found in this section continue to hold even if ψ is only once differentiable.[3] Further, the former form is usually much more convenient to deal with numerically. Also, it was in this form that Schrödinger [3] gave his original variational formulation of quantum mechanics, and finally, as Schrödinger has emphasized [13], it is an especially useful one for transforming the Schrödinger equation from Cartesian coordinates to arbitrary coordinates (it is easier to transform $\boldsymbol{\nabla}\psi$ than to transform $\boldsymbol{\nabla}^2\psi$). Nevertheless, in spite of all these virtues of $(\boldsymbol{\nabla}\psi, \cdot\boldsymbol{\nabla}\psi)$, we will continue to use the expression (1) because it is much easier to deal with formally, and because in most applications the ψ *are* twice differentiable.

PROBLEMS

1. If ψ and ψ' are physically distinct (for example, orthogonal to one another), can $E = E'$ in (7)?
2. It may seem odd that out of all possible Δ's the special choice (8) already sufficed to derive the general result. However, show that the first-order contributions of any physically distinct additions to this Δ would vanish identically.
3. Devise a projection operator (presumably you are familiar with the general notion of projection operators) which is such that if one knows the lower eigenvalues of H, then one can produce trial functions which are orthogonal to the lower eigenfunctions. (See, for example, Löwdin [14].)

3. The Variation Method: Introduction

The results of the previous section are, aside from their theoretical interest, of great practical importance because they suggest a soundly based method

---

[3] For variation principles which allow ψ to be discontinuous, see, for example, Rudge [12].