The determination and interpretation of molecular wave functions

ERICH STEINER

The determination and interpretation of molecular wave functions

ERICH STEINER
Lecturer in Chemistry, University of Exeter

Cambridge University Press
CAMBRIDGE
LONDON · NEW YORK · MELBOURNE

Published by the Syndics of the Cambridge University Press
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
Bentley House, 200 Euston Road, London NW1 2DB
32 East 57th Street, New York, NY 10022, USA
296 Beaconsfield Parade, Middle Park, Melbourne 3206, Australia

Library of Congress catalogue card number: 75–18120

ISBN: 0 521 21037 2

First published 1976

Printed in Great Britain at the
University Printing House, Cambridge
(Euan Phillips, University Printer)

Preface

It is a fact of some historical interest that although a comprehensive and detailed theory of the electronic structure and bonding in molecules had been developed by the year 1960 (Pauling 1960; Coulson 1961), very few non-empirical ('*ab initio*') quantum-mechanical calculations for polyatomic molecules were performed before that date. The theory was built up almost wholly intuitively and empirically from an experimental knowledge of the physical and chemical properties of molecules, coupled with an extrapolation to polyatomic molecules of the results of (i) accurate solutions of the Schrödinger equation for the hydrogen molecule and molecular ion, (ii) highly approximate non-empirical calculations for small, and mainly diatomic, molecules, and (iii) semi-empirical calculations, mainly of the Hückel molecular-orbital type, for some larger molecules. It was not until the late 1950s that the advent of the electronic computer brought with it the possibility of performing accurate non-empirical calculations for polyatomic molecules of chemical interest. Much of the subsequent evolution of the theory of molecular structure has paralleled, and has to a certain extent depended on, the development of computing machines and techniques.

The advent of the computer has led to the creation of a new branch of theoretical chemistry, computational quantum chemistry, with its own specialized language, and with concepts that are increasingly influenced by questions of mathematical tractability and computational expedience. One consequence has been the creation of new problems of communication between theoretical chemist and experimental chemist, and this book can be regarded as an attempt to bridge the widening gap between the two. The primary concern of the book is the exposition of some of the more important theoretical and computational techniques that have been developed in recent years for the determination and interpretation of molecular wave functions, with particular emphasis on the non-empirical molecular-orbital approach. A feature of the evolution of the theory since 1960 has been the declining importance of the valence-bond approach as a practical tool of the computational quantum chemist. Although valence-bond theory is as valid as molecular-orbital theory, and merely represents an alternative method of constructing molecular wave functions, molecular-orbital theory has been found to be the more con-

venient for computational purposes and, to a lesser extent, for interpretative purposes.

It has been assumed that the reader is a graduate or advanced undergraduate in chemistry, with the appropriate knowledge of mathematics and of the applications of quantum mechanics in chemistry. Chapter 1 is devoted to a brief discussion of the non-relativistic time-independent Schrödinger equation for the motion of electrons in molecules, and of the general techniques available for its solution. Chapter 2 deals with the symmetry properties of electronic wave functions. Nearly all of the material in these first two chapters can be found discussed at greater length in standard undergraduate texts, but it has been included both to make the book more or less self-contained and to introduce the notation, units, and those general concepts that are used in the subsequent discussion. Chapter 3 is devoted to the Hartree–Fock model of electronic structure, and to its relation to what is commonly known as molecular-orbital theory. Methods of proceeding beyond the orbital approximation towards the exact solution of the Schrödinger equation are considered in chapter 4, which also includes a brief discussion of relativistic effects. Chapter 5 is concerned with some of the computational techniques that have been developed for the practical implementation of the theory developed in the previous chapters. The analysis and interpretation of molecular wave functions is discussed in the final two chapters. Chapter 6 is concerned with the electron distribution and chapter 7 with the nature of the chemical bond.

I wish to acknowledge with gratitude the encouragement and advice given to me by several of my colleagues in the Chemistry Department. Particularly I would like to thank Dr B. J. Skillerne de Bristowe for reading much of the manuscript during the earlier stages of preparation, and for many helpful suggestions.

E. S.

University of Exeter
May 1975

Contents

1 Introduction

1.1 THE SCHRÖDINGER EQUATION

We are concerned in this book with the non-empirical theory of the electronic structure of molecules. By a non-empirical (or *ab initio*) calculation in molecular quantum chemistry is normally meant the solution of a time-independent Schrödinger equation

$$\mathscr{H}\Psi = E\Psi \tag{1.1}$$

in which the Hamiltonian $\mathscr{H}$ is that appropriate to any model of the system which does not depend, either explicitly or implicitly, on the properties of any finite number of states of the system. In the simplest and most widely used model of a molecule, the nuclei and electrons are assumed to be non-relativistic point charges interacting through electrostatic (Coulomb) forces only, the force acting between charges q_1 and q_2 separated by distance r being $q_1 q_2/4\pi\epsilon_0 r^2$, where ϵ_0 is the permittivity of a vacuum.

Consider a molecule containing ν nuclei, with charges $Z_\alpha e$ and masses M_α ($\alpha = 1, 2, \ldots, \nu$), and N electrons, with charges $-e$ and masses m_e. Let the position of nucleus α be given by the vector $\boldsymbol{R}_\alpha$, whose components are the coordinates of the nucleus in a fixed coordinate system, and let $\boldsymbol{r}_i$ be the position vector of electron i ($i = 1, 2, \ldots, N$) The Hamiltonian for this system of point charges is

$$\mathscr{H} = \sum_{\alpha=1}^{\nu} \frac{-h^2}{8\pi^2 M_\alpha} \nabla_\alpha^2 + \sum_{i=1}^{N} \frac{-h^2}{8\pi^2 m_e} \nabla_i^2$$
$$- \sum_{i=1}^{N} \sum_{\alpha=1}^{\nu} \frac{Z_\alpha e^2}{4\pi\epsilon_0 r_{i\alpha}} + \sum_{i>j=1}^{N} \sum \frac{e^2}{4\pi\epsilon_0 r_{ij}} + \sum_{\alpha>\beta=1}^{\nu} \sum \frac{Z_\alpha Z_\beta e^2}{4\pi\epsilon_0 R_{\alpha\beta}} \tag{1.2}$$

where h is Planck's constant and, for example, $r_{i\alpha} = |\boldsymbol{R}_\alpha - \boldsymbol{r}_i|$ is the distance between electron i and nucleus α. The corresponding Schrödinger equation for this model has been found, through many applications, to form a satisfactory basis for the description of a very wide variety of properties of molecules, and we will be concerned in this book almost wholly with the methods that have been developed for its solution and with the interpretation of the solutions. Apart from a brief discussion

in §4.5 of the magnitudes of the relativistic corrections to the model, we will therefore consider as outside the scope of this book all time-dependent phenomena, the effects of external fields, as well as all magnetic interactions and other relativistic effects. Inclusion of these requires modification either of the Hamiltonian or of the form of the wave equation itself, although these modifications are often treated as perturbations on the model system (§1.5).

The Hamiltonian (1.2) contains terms which describe not only the motion of the electrons about the nuclei, but also the motion of the nuclei with respect to each other and of the molecule as a whole in space. The corresponding Schrödinger equation is very difficult to solve in general, and accurate solutions have been obtained only for the simplest molecules, H_2^+ and H_2 (Kolos and Wolniewicz 1964). The problem is simplified considerably however by including in the model further assumptions which result in the separation of the electronic and nuclear motions. These assumptions rely on the observation that the ratio m_e/M_α of the electronic to nuclear masses is a small number compared with unity, and they are therefore consistent with our definition of a non-empirical theory. The nuclear and electronic motions may be separated exactly for a one-electron atom, and to a very good approximation for other atoms (Bethe and Salpeter 1957). The resulting Hamiltonian for the internal motion of an atom, whose nucleus has charge Ze and mass M, is

$$\mathscr{H} = \frac{-h^2}{8\pi^2\mu}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\frac{Ze^2}{4\pi\epsilon_0 r_i} + \sum_{i>j=1}^{N}\sum\frac{e^2}{4\pi\epsilon_0 r_{ij}} \tag{1.3}$$

where $\mu = m_e M/(m_e + M)$ is the reduced mass of an electron in the atom, and $\boldsymbol{r}_i$ is the position vector of electron i relative to the nucleus as origin. The separation for molecules is of a somewhat different kind, the Hamiltonian for electronic motion being obtained by assuming that the nuclei have infinite masses and, therefore, have fixed positions relative to a fixed coordinate system. The Hamiltonian for the motion of the electrons in this 'fixed-nuclei' or Born–Oppenheimer approximation is

$$\mathscr{H} = \frac{-h^2}{8\pi^2 m_e}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{\alpha=1}^{\nu}\frac{Z_\alpha e^2}{4\pi\epsilon_0 r_{i\alpha}} + \sum_{i>j=1}^{N}\sum\frac{e^2}{4\pi\epsilon_0 r_{ij}} + \sum_{\alpha>\beta=1}^{\nu}\sum\frac{Z_\alpha Z_\beta e^2}{4\pi\epsilon_0 R_{\alpha\beta}} \tag{1.4}$$

The solutions of the corresponding Schrödinger equation depend on the nuclear positions, and a separate calculation of any electronic state must

be performed for each assumed molecular geometry. The stable geometry (in the Born–Oppenheimer approximation) for any state is that with the lowest energy.

The eigenfunctions of the Hamiltonian, (1.3) for an atom and (1.4) for a molecule, describe the stationary electronic states of the system. They are functions of the coordinates of the electrons, and they can always be chosen to be normalized and orthogonal (orthonormal); if Ψ_m and Ψ_n are any two eigenfunctions,

$$\int \Psi_m^* \Psi_n \, d\tau = \begin{cases} 1 & \text{if } m = n, \quad \text{for normalization} \\ 0 & \text{if } m \neq n, \quad \text{for orthogonality} \end{cases}$$

where Ψ_m^* is the complex conjugate of Ψ_m, and $\int \ldots d\tau$ implies integration over all the coordinates of the electrons. An important property of the eigenfunctions is that they form a *complete set* of functions in the sense that any arbitrary wave function Ψ, which is not an eigenfunction of $\mathscr{H}$ but which satisfies the same boundary conditions as the eigenfunctions, can be expressed as a linear combination of the eigenfunctions:

$$\Psi = \sum_n C_n \Psi_n$$

If the eigenfunctions are orthonormal, the coefficients are given by

$$C_n = \int \Psi_n^* \Psi \, d\tau$$

1.2 ATOMIC UNITS

The Schrödinger equation with Hamiltonian (1.4) for a Born–Oppenheimer molecule may be freed of the experimentally determined quantities e, h, ϵ_0 and m_e by the substitutions

$$\mathscr{H} = (m_e e^4/4h^2\epsilon_0^2)\mathscr{H}', \quad E = (m_e e^4/4h^2\epsilon_0^2)E'$$

and, for example,

$$r_{ij} = (\epsilon_0 h^2/\pi m_e e^2) r'_{ij}$$

The conversion factors are often treated as units, atomic units (Shull and Hall 1959). They are the Bohr radius a_∞ (or a_0) and the Hartree energy H_∞:

$$\left.\begin{aligned} a_\infty &= \epsilon_0 h^2/\pi m_e e^2 = 5.2918 \times 10^{-11}\,\text{m} \\ H_\infty &= m_e e^4/4h^2\epsilon_0^2 = e^2/4\pi\epsilon_0 a_\infty = 4.3598 \times 10^{-18}\,\text{J} \end{aligned}\right\} \qquad (1.5)$$

The resulting dimensionless Schrödinger equation and Hamiltonian are

$$\left.\begin{aligned} \mathscr{H}'\Psi' &= E'\Psi' \\ \mathscr{H}' &= -\tfrac{1}{2}\sum_{i=1}^{N}\nabla_i'^2 - \sum_{i=1}^{N}\sum_{\alpha=1}^{\nu}\frac{Z_\alpha}{r'_{i\alpha}} + \sum_{i>j=1}^{N}\sum\frac{1}{r'_{ij}} + \sum_{\alpha>\beta=1}^{\nu}\sum\frac{Z_\alpha Z_\beta}{R'_{\alpha\beta}} \end{aligned}\right\} \quad (1.6)$$

and the primes, which convert energies and lengths to numbers, are in practice omitted.

The use of the symbols a_∞ and H_∞ for the atomic units of length and energy requires a few words of explanation. The most commonly used symbol for the Bohr radius is a_0, but no generally accepted symbol exists for the Hartree energy, and many authors avoid the use of special symbols by denoting the atomic unit of every physical quantity by the abbreviation a.u. The symbols a_∞ and H_∞ proposed here have been chosen to be consistent with the SI-recommended symbol R_∞ for the Rydberg constant, which is related to the Hartree energy by $H_\infty = 2hcR_\infty$, where c is the speed of light. In the case of an atom whose Hamiltonian (1.3) involves the reduced mass μ instead of the electronic mass m_e, the Schrödinger equation is reduced to the dimensionless form (1.6), not involving μ, by the conversion factors obtained from a_∞ and H_∞ by replacing m_e by μ. The new conversion factors can be distinguished from a_∞ and H_∞ by a change of subscript to specify the nature of the nucleus. Thus the Rydberg constant for the normal (^{1}H) hydrogen atom is

$$R_H = R_\infty m_p/(m_p + m_e)$$

where m_p is the mass of the proton, and the corresponding symbols for the length and energy are a_H and H_H. In general for any subscript X,

$$H_X = 2hcR_X \quad \text{and} \quad a_X H_X = e^2/4\pi\epsilon_0$$

In this book we shall ignore, for simplicity, the small difference in value between m_e and the reduced mass μ of an electron in an atom; this corresponds to the Born–Oppenheimer assumption of infinite nuclear mass.

Any four of the five quantities m_e, e, $\hbar = h/2\pi$, a_∞ and H_∞ may be regarded as base atomic units for the construction of the atomic units of all those other physical quantities which, in SI, involve only the units of length mass, time and electric current. A list of some of the more important quantities is given in table 1.1. Other quantities used as units in this book include the ångström $\text{Å} = 10^{-10}\text{m} = 1.8897a_\infty$; the debye $\text{D} = 3.3356 \times 10^{-30}\,\text{C m}$, which is related to the atomic unit of electric dipole moment by $ea_\infty = 2.5418\,\text{D}$; the electron volt

TABLE 1.1 *Atomic units*

Physical quantity	Atomic unit	Value in SI units
Mass	m_e	9.1096×10^{-31} kg
Charge	e	1.6022×10^{-19} C
Angular momentum	$\hbar = h/2\pi$	1.0546×10^{-34} J s
Length	$a_\infty = 4\pi\epsilon_0\hbar^2/m_e e^2$	5.2918×10^{-11} m
Energy	$H_\infty = m_e e^4/16\pi^2\epsilon_0^2\hbar^2$	4.3598×10^{-18} J
Time	$\hbar/H_\infty$	2.4189×10^{-17} s
Linear momentum	$\hbar/a_\infty$	1.9928×10^{-24} kg m s^{-2}
Electric current	$eH_\infty/\hbar$	6.6237×10^{-3} A
Electric potential	H_∞/e	2.7211×10^{1} V
Electric dipole moment	ea_∞	8.4784×10^{-30} C m
Electric charge density	e/a_∞^3	1.0812×10^{12} C m^{-3}

eV $= 1.6022 \times 10^{-19}$ J, with $H_\infty = 27.211$ eV; and the molar energy $LH_\infty = 2.6255 \times 10^6$ J mol^{-1}, where L is the Avogadro constant.

1.3 THE VARIATION PRINCIPLE

In a non-empirical calculation one attempts to find eigenfunctions and eigenvalues of the 'exact' model Hamiltonian (1.6). Except for the simplest systems however, a complete solution of the Schrödinger equation is still an intractable problem, and it is therefore always necessary to resort to methods of finding approximate solutions. Almost all of these methods are based on the variation principle.

A simple expression of the variation principle is that, given any trial N-electron wave function Ψ which satisfies the necessary boundary conditions for the system, an upper bound to the exact ground-state energy E_0 is

$$E = \frac{\int \Psi^* \mathscr{H} \Psi \, d\tau}{\int \Psi^* \Psi \, d\tau} \geqslant E_0 \tag{1.7}$$

Analogous inequalities exist for excited states. A consequence of the principle is that if a trial wave function depends on a number of arbitrary parameters, $\lambda_1, \lambda_2, \ldots, \lambda_n$,

$$\Psi = \Psi(\boldsymbol{r}; \lambda_1, \lambda_2, \ldots, \lambda_n)$$

where $\boldsymbol{r}$ represents the dependence of Ψ on the coordinates of the electrons, then the values of these parameters can be chosen to give the lowest possible, and hence the most accurate, value of the energy. The energy is a

function of the parameters, and the values of the parameters which give the lowest value of the energy are obtained by solving the equations

$$\frac{\partial E}{\partial \lambda_i} = 0 \quad (i = 1, 2, \ldots, n)$$

The most general type of approximate wave function commonly used has the form

$$\Psi = C_1\Phi_1 + C_2\Phi_2 + \ldots + C_n\Phi_n = \sum_{j=1}^{n} C_j\Phi_j \tag{1.8}$$

where the coefficients C_j are parameters, and the Φ_j are given N-electron functions which satisfy the same boundary conditions as Ψ and which may or may not depend on further parameters. The corresponding energy (1.7) is

$$E = \sum_{i,j=1}^{n}\sum C_i^* C_j H_{ij} \Big/ \sum_{i,j=1}^{n}\sum C_i^* C_j S_{ij} \tag{1.9}$$

where

$$H_{ij} = \int \Phi_i^* \mathscr{H} \Phi_j \, d\tau, \quad S_{ij} = \int \Phi_i^* \Phi_j \, d\tau \tag{1.10}$$

and the minimization of E with respect to the n coefficients gives a set of n 'secular' equations

$$\sum_{j=1}^{n} (H_{ij} - ES_{ij}) C_j = 0 \quad (i = 1, 2, \ldots, n) \tag{1.11}$$

One (trivial) solution of the equations is obtained by setting all the coefficients equal to zero. Non-trivial solutions are obtained only if the energy E is chosen such that the secular determinant, whose elements are $(H_{ij} - ES_{ij})$, vanishes:

$$\det(H_{ij} - ES_{ij}) = 0$$

or

$$\begin{vmatrix} H_{11}-ES_{11} & H_{12}-ES_{12} & \ldots & H_{1n}-ES_{1n} \\ H_{21}-ES_{21} & H_{22}-ES_{22} & \ldots & H_{2n}-ES_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ H_{n1}-ES_{n1} & H_{n2}-ES_{n2} & \ldots & H_{nn}-ES_{nn} \end{vmatrix} = 0 \tag{1.12}$$

The secular determinant is a polynomial of degree n in the energy, and it has n roots, not necessarily all different,

$$E_1 \leqslant E_2 \leqslant E_3 \leqslant \ldots \leqslant E_n$$

Corresponding to each energy E_i, a wave function

$$\Psi_i = \sum_{j=1}^{n} \Phi_j C_{ji} \tag{1.13}$$

may now be obtained by solving the secular equations and normalization. The resulting wave functions are orthonormal:

$$\int \Psi_i^* \Psi_j \, d\tau = \sum_{k,l=1}^{n} \sum C_{ki}^* C_{lj} S_{kl} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The lowest root E_1 is an approximate ground-state energy, and the corresponding function Ψ_1 is an approximate wave function for the ground state. In fact, the set of n solutions are approximations for the first n states of the system. If $E_i^{(e)}$ is the exact energy of the ith state, then $E_i \geqslant E_i^{(e)}$ as shown in fig 1.1, and $E_i = E_i^{(e)}$ only if Ψ_i is the exact wave function for the ith state. The magnitudes of the separations $(E_i - E_i^{(e)})$ depend on the functions Φ_j included in the wave function (1.8). If the functions depend on further parameters,

$$\Phi_j = \Phi_j(\boldsymbol{r}; \lambda_1, \lambda_2, \ldots)$$

then the parameters can be chosen to minimize one of the roots E_i of the secular problem.

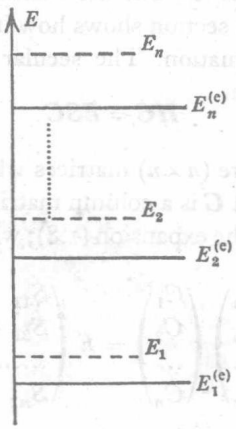

Fig. 1.1

It is generally true that increasing the number of variational parameters in a wave function results in improved accuracy of the corresponding energy. If the wave function has a general enough form, the exact solution is approached as the number of parameters is increased indefinitely. In particular, the 'method of linear combinations' outlined above provides better approximations to more and more states as the number n increases. In practice however the form of the wave function is often

constrained, and cannot lead to an exact solution of the Schrödinger equation. An example is the orbital approximation, discussed in some detail in chapter 3, in which the wave function and energy approach definite (Hartree–Fock) limiting values as the number of variational parameters increases, but these limits do not represent an exact solution of the Schrödinger equation. As we shall see, the wave function obtained in this way is an eigenfunction not of the Hamiltonian (1.6) but of a different Hamiltonian, the Hartree–Fock Hamiltonian. In this way a new model based on the *form* of the wave function is obtained as an approximation to the original 'exact' model, to which it reduces when the constraints imposed on the form of the wave function are relaxed.

1.4 MATRIX REPRESENTATION OF THE SCHRÖDINGER EQUATION

The Schrödinger equation for an N-electron system is a partial differential equation in $3N$ variables, but the method of linear combinations discussed in the previous section shows how it may be transformed into an equivalent matrix equation. The secular equations (1.11) can be written in the matrix form

$$\boldsymbol{HC} = \boldsymbol{ESC} \tag{1.14}$$

where $\boldsymbol{H}$ and $\boldsymbol{S}$ are square $(n \times n)$ matrices whose elements are H_{ij} and S_{ij} defined by (1.10), and $\boldsymbol{C}$ is a column matrix (vector) whose elements are the coefficients C_j of the expansion (1.8); written out in full, (1.14) is

$$\begin{pmatrix} H_{11} & H_{12} & \dots & H_{1n} \\ H_{21} & H_{22} & \dots & H_{2n} \\ \dots & \dots & \dots & \dots \\ H_{n1} & H_{n2} & \dots & H_{nn} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \dots \\ C_n \end{pmatrix} = E \begin{pmatrix} S_{11} & S_{12} & \dots & S_{1n} \\ S_{21} & S_{22} & \dots & S_{2n} \\ \dots & \dots & \dots & \dots \\ S_{n1} & S_{n2} & \dots & S_{nn} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \dots \\ C_n \end{pmatrix}$$

The matrix $\boldsymbol{H}$ is a representation of the Hamiltonian $\mathscr{H}$ in terms of the *basis* of n functions Φ_j. As has already been remarked, the solutions of the matrix equation are approximations for n states of the system, and increasing the number of basis functions leads to better approximations for more and more states. When n becomes infinitely large and the basis becomes complete,† the matrix equation (1.14) becomes entirely equiva-

† The definition of a complete set given on p. 3 is sufficient for our purposes. Given a set of functions Φ_n $(n = 1, 2, \dots, M)$ satisfying certain boundary conditions, the set is said to be complete if any arbitrary function Φ, which satisfies the same boundary

lent to the differential Schrödinger equation, having the same set of eigenvalues and eigenfunctions, the latter being expressed in the form (1.13) as linear combinations of the basis functions.

Given a basis therefore, the problem of solving the Schrödinger equation is reduced to the evaluation of the matrix elements H_{ij} and S_{ij}, and to the solution of the corresponding matrix equation. The only serious problem is concerned with the choice of basis functions. Ideally, we would like to be able to use a basis which gives both easily evaluated matrix elements and a rapid convergence (small n) of the expansion (1.13) of the wave function. This is not always possible, particularly for polyatomic molecules.

1.5 PERTURBATION THEORY

It is often the case that the Hamiltonian $\mathscr{H}$ for the system of interest differs only slightly from the Hamiltonian $\mathscr{H}_0$ of a related system. One example is a molecule in a weak external electric field for which the Hamiltonian can be written as

$$\mathscr{H} = \mathscr{H}_0 + \lambda V \tag{1.15}$$

where $\mathscr{H}_0$ describes the unperturbed system which is the molecule in the absence of the field, and λV is a 'small' perturbation term which describes the interaction of the molecule with the field. λ is a parameter, called the perturbation parameter (for example, the field strength), which is a measure of the strength of the perturbation.

It is generally assumed in perturbation theory that the eigenfunctions and eigenvalues of $\mathscr{H}_0$ are known,

$$\mathscr{H}_0 \Psi_n^{(0)} = E_n^{(0)} \Psi_n^{(0)} \tag{1.16}$$

or, since in practice only a few eigenfunctions may be known, that at least the unperturbed wave function $\Psi_n^{(0)}$ for the state of interest is known. We shall also assume, for simplicity, that the state of interest is one for which the energy $E_n^{(0)}$ is non-degenerate. Then, if Ψ_n and E_n are the

conditions, can be expressed as a linear combination $\Phi = \sum_{n=1}^{M} C_n \Phi_n$. If the basis functions are orthonormal, $\int \Phi_m^* \Phi_n \, d\tau = \delta_{mn}$, the coefficients are given by

$$C_n = \int \Phi_n^* \Phi \, d\tau,$$

Such complete sets of functions are almost always infinite, notable exceptions being the sets of N-electron spin functions containing 2^N members (§2.7).

(unknown) eigenfunction and eigenvalue of $\mathscr{H}$ for this state,

$$\mathscr{H}\Psi_n = E_n\Psi_n \tag{1.17}$$

it follows that in the limit $\lambda \to 0$,

$$\mathscr{H} \to \mathscr{H}_0, \quad E_n \to E_n^{(0)}, \quad \Psi_n \to \Psi_n^{(0)}$$

The basic assumption of perturbation theory is that the energy and wave function for the perturbed state may be expanded as power series in λ about the corresponding energy and wave function of the unperturbed state,

$$\left.\begin{aligned} E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots \\ \Psi_n &= \Psi_n^{(0)} + \lambda\Psi_n^{(1)} + \lambda^2\Psi_n^{(2)} + \ldots \end{aligned}\right\} \tag{1.18}$$

and that these expansions are valid for the whole range of values of λ between zero and the value of interest. The quantity $E_n^{(i)}$ is called the ith-order energy and $\Psi_n^{(i)}$ is the ith-order wave function. Substitution of the expansions for the energy and wave function in the Schrödinger equation (1.17) gives

$$(\mathscr{H}_0 + \lambda V)(\Psi_n^{(0)} + \lambda\Psi_n^{(1)} + \lambda^2\Psi_n^{(2)} + \ldots)$$
$$= (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots)(\Psi_n^{(0)} + \lambda\Psi_n^{(1)} + \lambda^2\Psi_n^{(2)} + \ldots)$$

or

$$(\mathscr{H}_0 - E_n^{(0)})\Psi_n^{(0)} + \lambda[(\mathscr{H}_0 - E_n^{(0)})\Psi_n^{(1)} + (V - E_n^{(1)})\Psi_n^{(0)}]$$
$$+ \lambda^2[(\mathscr{H}_0 - E_n^{(0)})\Psi_n^{(2)} + (V - E_n^{(1)})\Psi_n^{(1)} - E_n^{(2)}\Psi_n^{(0)}] + \ldots = 0 \tag{1.19}$$

In order that the equation be satisfied for arbitrary values of λ, it is necessary that the coefficient of each power of λ be separately zero. The zeroth-order equation is

$$(\mathscr{H}_0 - E_n^{(0)})\Psi_n^{(0)} = 0$$

which is simply the original Schrödinger equation for the unperturbed state. The first- and second-order equations are

$$(\mathscr{H}_0 - E_n^{(0)})\Psi_n^{(1)} + (V - E_n^{(1)})\Psi_n^{(0)} = 0 \tag{1.20}$$

$$(\mathscr{H}_0 - E_n^{(0)})\Psi_n^{(2)} + (V - E_n^{(1)})\Psi_n^{(1)} = E_n^{(2)}\Psi_n^{(0)} \tag{1.21}$$

We therefore obtain, if the perturbation λV is small enough, a set of equations which can be solved in sequence to give progressively more accurate solutions of the Schrödinger equation for the perturbed state. In many applications it is not necessary to go beyond the first-order wave function as this determines the energy to third order.

The solution of the perturbation equations is discussed in standard