Some aspects of quantum mechanics in chemical theory

Thomas Freeman Babson

Union College - Schenectady, NY
SOME ASPECTS OF QUANTUM MECHANICS IN CHEMICAL THEORY

by

Thomas F. Babson

Senior Thesis Submitted
in Partial Fulfillment
of the Requirements of Graduation

DEPARTMENT OF CHEMISTRY
UNION COLLEGE
DECEMBER 1971
This Thesis

Submitted by

Thomas A. Baker

to the

Department of Chemistry of Union College

in partial fulfillment of the requirements of the degree of

Bachelor of Science with a Major in Chemistry

is approved by
# Table of Contents

Preface ........................................................................................................ iv

Introduction .................................................................................................. 1

Chapter I: Considerations in Orbital Construction ........................................... 5
  Section 1; Quantum Numbers and Hydrogen Orbitals ................................. 5
  Section 2; The Pauli Exclusion Principle .................................................... 7
  Section 3; Hund's Rules ............................................................................ 8
  Section 4; The Variational Principle .......................................................... 10

Chapter II: Molecular Orbitals ...................................................................... 13
  Section 5; Self Consistent Fields ................................................................ 13
  Section 6; Molecular Orbitals ..................................................................... 14

Chapter III: Some Chemically Relevant Models of Atomic and
  Molecular Systems ..................................................................................... 21
  Section 7; Linnett's Model of the Electronic Structure of Molecules .......... 21
  Section 8; Daudel's Model of Loges .......................................................... 23

Conclusion .................................................................................................... 27

Appendix I: Hilbert Space and the Mathematics of Quantum Theory .......... 30

Appendix II: The Statistical Interpretation of Quantum Mechanics of the Copenhagen School ........................................................................................................ 33

Appendix III: The Hidden Variable Interpretation of Quantum Mechanics .... 40

Appendix IV: Quantum Numbers .................................................................. 45

Appendix V: Thomas-Fermi Statistical Approximation ................................ 46

Appendix VI: The Pauli Exclusion Principle ................................................ 49

Appendix VII: The Variational Principle ...................................................... 51

Bibliography ................................................................................................. 53
Preface

Herein do lie some of the mystical tenets of the Way of the Quanta.

Equality of time and space has not been allotted the different sects of the Way of the Quanta for time and space are such relative parameters and are therefore not suitable for exact measurement. The Way of the Quanta has an intellectual and empirical side as well as its more substantial mystical side, and in the words of the High Priest of Determinism (heed well ye chemists, for verily most of ye be determinists): "God doesn't play with dice."
Introduction

The physics of microscopic phenomena is called quantum mechanics. When quantum mechanics is extended to the macroscopic scale, the classical macroscopic physics results. Atomic interactions such as the formation of bonds between atoms fall into the range of physics described by quantum mechanics. Because of the interest of a chemist in such interatomic effects, it seems reasonable that a chemist would be interested in quantum mechanics because it is the physics governing these effects.

Quantum mechanics exists as a discipline in two different forms. These forms are: the analytical quantum theory (Copenhagen interpretation and hidden variable interpretation; for more detail, see Appendices II and III) and empirical quantum mechanics. The analytical theories are derived from putting initial assumptions as to the nature of empirical reality into a mathematical formalism. Using this formalism and therefore the mathematical techniques it allows, predictions are made concerning physical reality. There are two analytical theories and each has initial assumptions which contradict the initial assumptions of the other. Nevertheless they both make basically the same physical predictions.

The empirical form of quantum mechanics is the collection of empirical data of microscopic phenomena and some results of analytical quantum mechanics all expressed in the same type of mathematical formalism analytical quantum mechanics is expressed in.

Because the interests of chemists and physicists are different, though related, the forms of quantum theory each uses need not be the same. Physicists are interested in the physics of the processes which make up properties of systems. Chemists are interested in the physics of properties of systems and not so much in the processes which are the causes of the properties. Because physicists are interested in the cause of the effect, their reasoning must be analytical so that the properties are de-
duced from the processes and not have the processes induced from the properties. The former is logically valid whereas the latter is not. Hence the physicist uses the analytical theories of quantum mechanics and insists upon mathematical rigor. Chemists can be much more empirical than physicists because of the emphasis they put on the manifestation and interaction of properties. It is all right for the chemist's view of quantum mechanics to be inductive and quasi-analytical because he uses it as a tool and such a view serves the chemist's purpose and is ultimately validated by the analytical view. A chemist is able to operationally define the properties with which he is working and then work with them. These operational definitions are derivable from analytical quantum theory and it is enough for the chemist to know this. Hence the chemist's use of empirical quantum mechanics which is perfectly valid for the use he puts it to.

Quantum mechanics is basically a mathematical theory whether it is mathematically derived or whether empirical results are put into mathematical formalism. It is from this formalism that predictions of the nature of physical processes and attributes and the correlations between different processes and attributes are made. It is therefore important to have some qualitative idea of the basis of the mathematics of quantum mechanics. The mathematics takes place in Hilbert space, a vector space which is complex and infinite dimensional. Vectors (functions) are used to represent physical states and physical observables are represented by a real (that is, having no imaginary character) Hilbert space operator called the Hermitian operator. A mathematical description of this space can be found in Appendix I.

The mathematical formalism of quantum mechanics is the wave equation which describes a wave function (eigenvector) in Hilbert space. This is
known as the Schrödinger equation, \( \dot{\psi} = -i\psi \) and the quantum mechanics discussed in this paper is known as the Schrödinger picture. There are other formalisms derivable as other theories (Heisenberg and interaction pictures) but the formalisms of these can be shown to be mathematically equivalent to the Schrödinger picture. This is done in part by von Neumann in his work, *Mathematics of Quantum Mechanics*.

When an eigenvector in Hilbert space is transformed into coordinate space (laboratory space) some of the subtle properties manifested in Hilbert space are lost. Transformations are made so that a conceptual picture of the system can be constructed in terms of space as we consciously perceive it (i.e. in terms of laboratory space). If the mathematical description in Hilbert space is assumed to be "reality" then its transformation into coordinate space must be "less real" due to the effects lost in the transformation. Therefore it should be remembered that the concepts of orbitals and other physical models in coordinate space based on a mathematical formalism in Hilbert space cannot be perfect models including all of the characteristics of Hilbert space representation.

Quantum mechanics is deficient in describing molecular systems due to the insolvability of the many-body problem (only the interaction between two bodies has been worked out mathematically, the interactions between many bodies is as yet unsolved). To adjust quantum considerations to the many-body problem further approximations must be made which overpower some of the subtle effects of quantum mechanics. Nevertheless, quantum theory can be of some use in determining the nature of chemical systems in several ways. Quantum mechanics derives a number of concepts such as Pauli exclusion and Heisenberg uncertainty which tell what the electron can or cannot do.
Using the results of quantum mechanics, both analytical and empirical, atomic models can be constructed and then these models can be used as a basis for techniques by which one can calculate bond energies, bond lengths, and other properties associated with molecules.

The point of this thesis is to present some of those aspects of quantum mechanics that are directly applicable to molecular systems and that can aid the chemist's understanding of molecular systems. The text of this paper will be primarily qualitative and therefore necessarily very general. The purpose in writing it is to point out some possibilities of applicability to chemical conceptions, not to rigourously prove all statements made. There may, however, be some readers who are purists enough to want to know where these statements come from and what the whole basis of quantum theory is itself, not just how it is applied to chemical systems. For the satisfaction of these purists, (the author included), the appendices following the text are much more mathematically developed. However, if one is willing to accept the results of the appendices, one need not understand them in order to understand the text.
Chapter I: Considerations in Orbital Construction

Section 1: Quantum Numbers and Hydrogen Orbitals

In the Schrödinger equation three compatible attributes of an electron (i.e. attributes which can be measured simultaneously, see Appendix IV), the energy, the total angular momentum and the angular momentum along a certain axis are described.

In the case of the hydrogen atom, the Schrödinger's equation can be written most simply, for the purpose of separation of variables, in spherical polar coordinates. This gives the result that the wave function for the hydrogen atom can be described as the product of the radial and angular functions \( R \) and \( \Theta \Phi \) when the said functions are properly normalized. That is

\[
\Psi_{n, l, m_l}(r, \theta, \phi) = R_{n, l}(r) \Theta_{l, m_l}(\theta) \Phi_{m_l}(\phi).
\]

The integers \( n, l, \) and \( m_l \) arise in the mathematical formalism associated with the separation of the variables of the Schrödinger's equation which can be found in any text on quantum chemistry. In the sense of the measured values of attributes, the integers \( n, l, m_l \) are quantum numbers (see Appendix IV). Due to the fact that the functions \( R_{n, l}(r), \Theta_{l, m_l}(\theta), \) and \( \Phi_{m_l}(\phi) \) are orthogonal, it follows that the one-electron wave functions are also mutually orthogonal, i.e.

\[
\langle \Psi_{n, l, m_l} | \Psi_{n', l', m_{l'}} \rangle = \sqrt{n n'} \sqrt{l l'} \sqrt{m_l m_{l'}}.
\]

In the simple system thus described by the Schrödinger's equation, there is associated with the energy level of the electron the quantum number, \( n \), called the principle quantum number. This corresponds to the number of nodes in the wave function (there are \( n-1 \) nodes). Associated with the total angular momentum of an electron is the quantum number, \( l \), known as the azimuthal quantum number. The functions \( \Theta(l, m_l) \Phi(m_l) \) are eigenfunctions of the Hermitian operator \( \hat{L}^2 \) corresponding to the total orbital angular momentum with the resulting
eigenvalues of \( \ell (\ell + 1) \hbar^2 \) as in
\[
\hat{L}^2 \mathbf{E}(\ell, m_\ell)\mathbf{E}(m_\ell) = \ell (\ell + 1) \hbar^2 \mathbf{E}(\ell, m_\ell)\mathbf{E}(m_\ell).
\]
The specific vectorial component of the angular momentum is described by placing the atom in a magnetic field and from the direction of this field, the Z-axis is defined. Associated with this attribute is the \( m_\ell \) or magnetic quantum number. The functions \( \mathbf{E}(m_\ell) \) are eigenfunctions of the Hermitian operator \( \hat{L}_z \) which is the operator of the angular momentum along the Z-axis and corresponding to these eigenfunctions are the eigenvalues of \( m_\ell \hbar \). The relationship between these three quantum numbers is:

\[
\begin{align*}
n &= 1, 2, 3, \ldots \quad & n & \text{is any positive integer} \\
n-1 & \leq l \leq 0 \quad & l & \text{is any positive integer or zero, but is less than } n \\
m_\ell &= 1, \ldots, 0, \ldots, -1 \quad & m_\ell & \text{is an integer and can take on values from the } l\text{-value to the negative of the } l\text{-value.}
\end{align*}
\]

The \( l\)-values of 0, 1, 2, 3, are called s, p, d, f, \ldots modes when used to describe the electronic structure of an atomic or molecular state.

With each \( l\)-value, there can be associated an \( m_\ell \) value as shown above, each one corresponding to a different direction.

An orbital is a wave function associated with a one electron system. Some examples of orbitals which are mathematically constructed with regards to the quantum numbers are (in the notation most familiar to chemists):

\[
\begin{align*}
\psi_{n, l, m_\ell} \\
\psi_{1, 0, 0} &= \psi_{1s} \\
\psi_{2, 0, 0} &= \psi_{2s} \\
\psi_{2, 1, \cos \psi} &= \psi_{2p_x} \\
\psi_{2, 1, \sin \psi} &= \psi_{2p_y} \\
\psi_{2, 1, 0} &= \psi_{2p_z}
\end{align*}
\]
in which the $m_l$ terms are written as the functions of angles, in the same way that they would be for the analogous case of spherical harmonics.

Section 2: The Pauli Exclusion Principle

There are two varieties of identical particles (this is according to the Copenhagen interpretation, see Appendices II and III). The first variety is known as fermions and the second as bosons. Fermions are particles which have spins (which are quantum numbers as defined in Appendix IV) that are odd multiples of $\frac{1}{2}$ and they obey Fermi-Dirac statistics. Bosons have spins that are even multiples of $\frac{1}{2}$ and obey Bose-Einstein statistics. According to Fermi-Dirac statistics, two fermions cannot occupy the same individual quantum state. This is commonly known as the Pauli Exclusion Principle. Electrons are fermions and hence they obey this principle. In terms of quantum numbers this means that no two electrons can have the same four quantum numbers. Therefore if two electrons in a given system have the same values of $n$, $l$, and $m_l$, their spin quantum numbers, $m_s$, must be different (one of them $+\frac{1}{2}$ the other $-\frac{1}{2}$).

The mathematical analysis of the Pauli exclusion principle have been left to Appendix VI.

In each one of the orbitals listed in Section 1 there can be two electrons. These electrons will have the same $n$, $l$, and $m_l$ quantum numbers (as these are what determine the orbital). However in light of the exclusion principle, the electrons in the orbital are in different spin states so that no two electrons are in the same overall state.
Section 3: Hund's Rules

Hund's rules are rules for determining the relative energies of atoms due to angular momentum couplings resulting from electronic configurations. These rules apply when the coupling involved is of the Russell Saunders type, also known as LS coupling. The explanations of the differences in energies are rationalizations and not theoretically derived. Some aspects of Hund's rules are derivable from relativistic considerations.

Hund's rules are of use to the chemist because they tell him which electronic configuration is lowest in relative energy. From this the ground state of the atom can be determined. Consideration of these effects is important in constructing the wave function for an atomic state.

Russell-Saunders coupling is the coupling which takes place in light atoms. It is the coupling of orbital angular momentum forming a total orbital angular momentum \( L \) and the coupling of spin angular momentum to form a total spin angular momentum \( S \). The magnetic interaction of these is small in light atoms. Thus the major components of the total angular momentum for light atoms are primarily these electrical couplings. The total angular momentum is denoted by \( J \) and is found from the different possible couplings of \( L \) and \( S \), i.e.,

\[
J = L+S, L+S - 1, \ldots, 0, \ldots, |L - S|.
\]

These rules state that:

1. The energy state is lowest for that state with the highest multiplicity. This means that for the given eigenvalues \( \vec{\lambda} \) for the different electronic states, the one with the greatest number of independent eigenvectors of the form \( \vec{\phi} \) associated with the eigenvalue \( \vec{\lambda} \), will have the lowest energy.

In other words, the state with the greatest \( S \) value has the
greatest number of eigenvectors associated with the eigenvalue used to describe spin of a given direction and thus the greatest multiplicity. According to the Pauli exclusion principle, fermions of the same energy states (i.e. same quantum numbers) cannot occupy the same space at the same time. The greater the number of electrons in an energy state, the greater the multiplicity and also the greater the avoidance of each other. The more electrons avoid each other, the smaller the Coulombic forces and hence the smaller the energy. If the spins are opposite then the electrons may occupy the same region of space and thus increase the Coulombic forces and raise the energy.

2. If there is more than one electronic system with a given multiplicity, then of these, the one with the highest orbital angular momentum is the lowest in energy. Angular momentum is a vectorial quantity and therefore has a direction associated with it. The greater the orbital angular momentum associated with an electronic system, the greater the number of electrons whose orbital angular momentum is associated in the same direction. Thus there is less interference between individual electrons in terms of trying to occupy the same region at one time than if their angular momenta were in opposite directions tending towards electron-electron "collisions." The result of this is a reduction in Coulombic forces and energy. Hence, the greater the orbital angular momentum, the smaller the energy.

3. a.) In electronic configurations where the shell is half or less than half filled with electrons, the states with the lower total angular momentum also have the lower energy.

3. b.) In electronic configurations in which the shell is more than half filled, the lower the total angular momentum, the higher the energy.

This third rule is concerned with the difference in energy between
two electron configurations with the same spin and orbital angular momenta (i.e., they are degenerate states). This energy difference is derived from the relativistic spin-orbit interactions. The result of these interactions is that the difference in energy between the two states due to spin-orbit interactions is proportional to the total angular momentum, \( J \). Thus

\[
\Delta E_{\mathcal{J}, \mathcal{J}'} = A(J).
\]

The constant \( A \) is either positive or negative and the value of \( J \) depends upon the algebraic sign of \( A \). Within the degenerate level, the lowest energy state for \( A > 0 \) is the lowest value of \( J \) or \( J = |L - S| \). The lowest value of energy for \( A < 0 \) is the highest value of \( J = |L + S| \). The algebraic sign of \( A \) depends on the completeness of the electronic shell due to the relativistic considerations of the spin-orbit interactions. When the shell is half or less than half filled, the sign of \( A \) is positive. When the shell is more than half filled \( A \) is negative. Thus the third rule is derivable from these relativistic considerations.

**Section 4: The Variational Principle**

In the mathematical construction of orbitals, it is important to have a function in the wave equation that will accurately represent the parameters of interest to the chemist (distances, energies, etc.). One method of achieving this is shown below.

Remembering the Schrödinger's equation,

\[
\frac{\mathcal{H}}{\mathcal{X}} \mathcal{X} = E \mathcal{X}
\]

consider the functional

\[
W[\phi] = \frac{\langle \phi | \mathcal{H} | \phi \rangle}{\langle \phi | \phi \rangle}
\]

where \( \mathcal{H} \) is the hamiltonian operator as defined in the Schrödinger equation (4-1) and \( \phi \) can be normalized and is a function of the coor-
coordinates of the space defined by the basis vectors of the eigenfunctions ($\psi$) as in (4-1). This functional is called the variational functional and use of it to determine the accuracy of an eigenfunction in terms of energy is called the variational principle. When $\psi = \phi$, then $W = E$, where $E$ is the exact energy of the system described by the Schrödinger's equation.

This functional is used to obtain approximate values for the eigenenergy of an electron (or electrons) in an atomic or molecular system described or approximately described by the Schrödinger equation if the exact nature of the wave equation is not known.

Assume that $\phi$ can be made up of some linear combination of the true orthonormal eigenfunctions of ($\mathcal{A}$). Then it follows that $\phi = \sum_j a_j \psi_j^0$ and if $\phi$ is normalized, then $\sum_j a_j^* a_j = 1$ and $\langle \phi | \phi \rangle = 1$. Substituting into the variational functional, the result is

$$W = \sum_j a_j^* a_j \langle j | H | j \rangle.$$  

Because the $\psi_j^0$'s are orthonormal and are eigenfunctions of ($\mathcal{A}$) with eigenvalues of $E_j$, then (4-3) becomes

$$W = \sum_j a_j^* a_j E_j.$$  

If $E_0$ is the lowest possible energy level, then $W - E_0 = \sum_j a_j^* a_j (E_j - E_0)$ is the difference in value between the approximate energy and the actual energy. $E_j - E_0$ must be zero or positive and therefore $W \geq E_0$. This means that whatever value obtained (for the energy) using the variational principle is equal to or greater than the actual value.

There is a definite correlation between the Schrödinger equation and the variational functional in that the Schrödinger equation is derivable from the functional. This is a fairly standard treatment and has been included in Appendix VII.

The danger in using the variational principle is that the estimated eigenenergy may effectively average out the error in energy over
the coordinates and thereby emphasize, to a disproportional extent, some
other operator which is coordinate dependent in some other fashion. In
other words, the total energy is not the most sensitive parameter con-
tained in a wave equation. As the energy functions are varied slightly,
other functions (e.g. the electric field gradient corresponding to elec-
tron distribution in regions of high nucleus probability density) may
change disproportionately.

The virial theorem states that any system made up of electrically
charged particles between which only electrostatic forces are signifi-
cant is subject to the following relation,

\[ E = -\bar{T} + \bar{V}/2 \]

where \( E \) is the total energy, \( \bar{T} \) is the average kinetic energy, and \( \bar{V} \)
is the average potential energy. Thus in the choosing of a trial func-
tion for the variational functional, the virial theorem should be sat-
ished. This can be done by scaling the trial function with variational
parameters that allow the energy to go to a minimum.

When a chemist is trying to determine the proper wave function to
use to describe an electronic system he will apply the variational func-
tional (4-2). This will give the worth of the trial function in terms
of energy. However because the chemist is interested in other properties
of the system, he must be aware of the severe limitations of the useful-
ness of this method. Hence the wave function used may not be the one
with the smallest \( W - E_0 \) value.
Chapter II: Molecular Orbitals

Section 5: Self Consistent Fields

Having some idea as to how the atomic orbitals of an atom are constructed (how an arbitrary wave equation is chosen so that it fulfills certain requirements, some of which were described in Chapter I) it is now necessary to see how this concept of the orbital is used in describing the molecular systems which are of interest to the chemist.

With the exception of the electrons in the outermost incomplete shell of a complex atom, the effect of the potential energy on any electron can be approximated as being due to a centrally symmetric potential energy. Because the part of the Schrödinger equation corresponding to the potential energy is only dependent upon the radial parameter, this approximation would result in solutions for the angular wave functions of complex atoms that would be the same as those for hydrogen. Thus the only difference between the wave function for hydrogen and the wave function for complex atoms would be in the radial function. The angular momentum quantum numbers $l$, $m_l$, and $m_s$ would be unchanged and therefore the structure of the complex atom could be described in the same terms as the structure of hydrogen is.

The Hartree self-consistent field method is based on the above assumptions and the additional assumption that the averaging of all non-spherically symmetric electron distributions over all angles is a valid approximation. This method proceeds by first constructing a possible electronic structure of the atom. Then using the modelistic concept of an electron cloud in the atomic space, the potential energy of the atom due to the electrons can be calculated at any point in the electron space. This potential energy is solely a function of the radius.
The mathematical theory used in the SCF method is of no great complexity. However the actual calculations are extremely hard. Furthermore, there are certain important physical affects the method, as it now stands, ignores and the inclusion of these affects in the method increases the complexity of the calculations. The interactions between electrons and the indistinguishability of individual electrons are two such effects. On the other hand the SCF method provides a basis for explaining the shielding of outer electrons from the full nuclear charge by inner electrons. This is done by the introduction of the shielding constant and is fundamental to the philosophy of the SCF method in that it further reduces the overall state of the complex atom to a collection of one electron systems.

The calculations of the SCF method are too complex to be of much use in atoms with a large atomic number. The Thomas-Fermi approximation method is a semi-classical statistical method that is useful for determining the potential energy of atoms with \( Z \gg 1 \). This method uses Fermi-Dirac statistics to predict the ground state energy and electron distribution of an atom. The details of the Thomas-Fermi method are fairly involved mathematically. A detailed account can be found in Appendix V.

Section 6; Molecular Orbitals

The SCF method is used to reduce the complex problem of many electron atoms to a series of one electron problems, involving the outermost electrons (the ones in the incomplete shell). It is in this light that the wave equations are found. Most chemists do not view many electron systems in terms of an overall wave equation as such an analysis would be meaningless. If this kind of wave equation could be constructed, it would be impossible to solve (more unknowns than equations relating the
unknowns). Instead many electron systems are thought of as combinations of one electron systems (in the SCF way) which are held together by certain interactions.

When the electron system is a molecule and the problem of the complex electron system is dealt with in a way similar to the SCF method, then the assumption is being made (albeit incorrectly) that the electronic structure of a molecule is the same as the electronic structure of an atom.

Using the variational principle, the energy of a proposed system can be calculated. In these calculations, three interaction integrals are conceived of. These are, the coulomb integral, $J$, the resonance integral, $K$, and the overlap integral, $S$. Solution of the variational principle for a $H_2^+$ molecule yields

$$W_{\text{symmetric}} = E_{1s} + \frac{e^2}{r_{AB}} + \frac{J + K}{1 + S}$$

$$W_{\text{antisymmetric}} = E_{1s} + \frac{e^2}{r_{AB}} + \frac{J - K}{1 - S}$$

when the trial function used was a linear combination of two hydrogen atom functions, i.e. $\\Psi = C_1 (1s_A) + C_2 (1s_B)$. The effect of this function is to account for the possibility of the electron being associated with either nucleus. The values obtained for radii and energy using this wave function are not very good. If other hydrogen orbitals are also linearly combined to construct the trial function, the results get better and better. Whatever wave function is used, it must satisfy the antisymmetry principle and therefore be antisymmetric. However the energy due to relativistic effects is small enough to be lost in the gross approximations made to construct the molecular orbitals (one electron wave functions for the molecule). Thus the spin function can be attached to the spatial function in such a way as to make the overall wave function antisymmetric.
Disregarding spin effects for the moment, the wave function of $H_2$ according to the linear combination of atomic orbitals (LCAO) method is

$$\psi_{\text{LCAO}} = \frac{1}{\sqrt{2}} \left[ \phi_A \phi_A + \phi_A \phi_B + \phi_B \phi_A + \phi_B \phi_B \right]$$

where the $\phi$'s are the $1s$ orbital wave functions, $A$ and $B$ stand for the two nuclei and the first function in every term stands for the first electron and the second stands for the second electron. The first and fourth terms of this equation are the probability of both electrons being close to one or the other nucleus. These represent the ionic character of the $H_2$ molecule. The second and third terms are the probability that one electron will be close to one nucleus and the other electron will be close to the other nucleus. These are representative of the covalent character of the $H_2$ molecule.

The ionic terms predict that the structure of $H_2$ is either $H_A^+ H_B^-$ or $H_A^- H_B^+$. These ionic terms are given footing equal to that of the covalent terms even though it is empirically known that the molecule is primarily covalent in nature. Hence the poor agreement between experimental values and values based on this approximation.

The valence bond approach to the construction of molecular orbitals for the $H_2$ molecule includes only the covalent terms. Again disregarding spin

$$\psi_{\text{VB}} = N (\phi_A \phi_B - \phi_B \phi_A)$$

where $N$ is the normalization constant. The values obtained from this approximation are better than the values obtained from the LCAO method (for this particular example).

Comparing the two methods, it can be seen that the valence bond approach is fairly accurate in its representations of electronic systems.
in which there is a fairly large distance between the nuclei. However in doing this, it over emphasises the positioning of the electrons between the nuclei, whereas the ionic term in the LCAO method has the effect of under emphasising the central positioning of the electrons.

There are ways of modifying the defects of both of these systems. One of these is to add an ionic-like function $\mu [\phi_A \phi_B + \phi_B \phi_A]$ to the valence bond approximation where the first function corresponds to the first electron and the second function to the second electron and $\mu$ is a scaling parameter. Equivalent to this would be to add the function $\frac{|\phi_A - \phi_B|^2}{2}$ to the MO configuration from the LCAO method which would decrease the ionic-like character. Both of these methods yield the Weinaum wave function which is

$$E = |\phi_A \phi_B| + |\phi_B \phi_A| + \lambda (|\phi_A \phi_A| + |\phi_B \phi_B|)$$

where the $\lambda$ is a scaling parameter called the interaction parameter and the first and second functions in each term correspond to the first and second electrons respectively.

Another form of this wave function is in terms of semilocalized atomic orbitals $\phi'_A$ and $\phi'_B$, defining $\phi''_A$ and $\phi''_B$ from

$$\phi''_A = \phi_A + k \phi_B$$
$$\phi''_B = \phi_B + k \phi_A$$

where the $k$ is a variational parameter.

Then the following relation, (6-3) holds true;

$$E = \frac{1}{|\phi'_A \phi'_B|} - \frac{1}{|\phi''_B \phi''_A|}.$$ 

In this case the similarity to the valence bond method is obvious with the addition of the small effect each electron has on the nucleus around which the other electron is primarily centered (or the effect each nucleus has on the electron primarily centered around the other nucleus). The
result is a decrease in the ionic character of the LCAOMO approximation. Substitution of the values of the semilocalized atomic orbitals \((6-2a;b)\) into equation \((6-3)\) yields

\[
\begin{align*}
\varepsilon &= \left[ (\Phi_A + k \Phi_B) (\Phi_A + k \Phi_B) \right]^{1/2} \\
&= \frac{1}{\varepsilon_k} \left( 1 + \frac{k}{k^2} \right) \psi_{\text{cov}} + 2k \psi_{\text{iouon}}
\end{align*}
\]

where \((\Phi_A \Phi_A + \Phi_B \Phi_B) = \psi_{\text{cov}}\) and \((\Phi_A \Phi_B + \Phi_B \Phi_A) = \psi_{\text{iouon}}\). The elimination of the coefficient in front of the covalent wave function of \((6-4)\) by multiplying the equation by \((\frac{1}{1+k^2})\) yields equation \((6-1)\) with a value of \(\lambda = \frac{\varepsilon_k}{1+k^2}\). Thus showing the equivalence of the two forms (valence bond and LCAOMO) Weinbaum's equation \((6-1)\) and \((6-3)\).

Arguments for the polarization of the electron distribution in molecules have been made on the following grounds. The electron distribution around the nucleus of a hydrogen atom is spherically symmetrical. However this distribution is pulled out of shape towards the other atomic nucleus in the two atom hydrogen molecule. Thus the electron distribution of the electron primarily associated with a particular nucleus is polarized toward the other nucleus. The 1s atomic orbital cannot account for this polarization and is therefore not a sufficient basis orbital from which to determine the molecular orbital. By including a little of the character of the 2p orbital in the basis orbital, this polarization is built into the wave function for the system. A wave function according to this method would then look like (again ignoring the spin effects)

\[
\psi_A = N \left[ (\frac{1}{1+k^2}) \left( 1 S_B + \lambda \frac{z P_z A}{1 S_B + \lambda \frac{z P_z B}{} \right) \right]
\]

where \(N\) is the normalization constant, \(\lambda\) is a variational parameter, and the first and second functions refer to the first and second electrons. This idea was first introduced by N. Rosen in Physical Review, 38, 2099 (1931).

The most accurate nonempirical calculation made on the hydrogen
molecule is that of the James-Coolidge wave function as extended to fifty terms by Kolos and Roos. This function is a variation function and the distance is included in each of the fifty terms of the equation. Because of the equation's lack of empirical and model dependent basis, it is not surprising that the concepts of orbitals (atomic and molecular) and interaction integrals (Coulomb, overlap, and resonance) are not present as these are theory based.

In the LCAO-MO and the valence bond methods, there is a fundamental difference in the interpretation of the nature of the electron. Mathematically, the functions are of the same nature and are interpreted probabilistically. The LCAO-MO model is primarily concerned with the overall symmetry of the molecule and as such interprets the probabilistic function of the electron as an electron cloud whose density varies over the volume in which the cloud is said to exist. Thus the greater the probability of finding an electron in a certain volume, the greater the density of the electron in that volume. Thus if the function speaks of a 90% probability within a given volume, this is interpreted as meaning 90% of the electron is present in the volume indicated.

The valence bond method is more attentive to the localized effects of electrons in atoms and molecules and therefore bases its model on these localized effects. As such, the model tends to think of the electron as a fairly discrete particle. It interprets the probabilistic function as being the probability that the particle electron will be at a given place at any time. If the probability is given as 90%, the physical interpretation is that 90% of the time, the particle electron will be within the volume specified.

Hybrid orbitals are mathematical constructs used to explain empirical
data. For example, carbon has an electronic structure of $1s^2$, $2s^2$, $2p_x$, $2p_y$. However, it forms bonds as if its structure was $1s^2$, $2sp^4$, that is, the electrons in the outer shell act as if they all had $1/4$ the character of $2s$ electrons and $3/4$ the character of $2p$ electrons.

According to the atomic orbital theory, atomic orbitals are real whereas hybrid orbitals only mathematically describe what is present in molecular geometry and do not have the same degree of reality associated with them.

The theories of Linnett and Daudel which are summarized in the next chapter predict molecular geometry fairly accurately without recourse to empirically based constructs, i.e. mathematical constructs empirically derived.
Chapter III: Some Chemically Relevant Models of Atomic and Molecular Systems

Section 7: Linnett's Model of the Electronic Structure of Molecules

Linnett's theory of the electronic structure of molecules is primarily based on the effects stemming from the exclusion principle and charge correlation between electrons and between electrons and protons. This is not a theory in the same way quantum mechanics is a theory. It is more of a hypothetical model based, in part, on the theory of quantum mechanics and empirical considerations.

In the theory he shows it is reasonable to think of atoms ionizing to the electronic structure of the inert element most similar to them by either gaining or losing electrons (i.e. fulfilling the so called octet rule); see (14).

Assuming the exclusion principle, charge correlation, and the octet rule, Linnett theorizes that the most stable configuration of the four electrons of a particular spin in an atom will be in the form of a tetrahedron with the nucleus in the center and the electrons at the corners. The other four electrons of the opposite spin will also form a tetrahedron. Now these two sets of four electrons are correlated will depend upon such things as neighboring charges. If the system is isolated, there will probably be little attraction due to charge effects. The tetrahedron model assumes the equivalence of all the electrons in a given tetrahedron, analogous to hybrid orbitals of the atomic orbital model.

If it is energetically favorable, atoms will try to alter their number of electrons so as to fulfill the octet rule. In the case of sodium and chlorine, the sodium tries to eliminate an electron and the chlorine is looking for another electron so that they can both have the electronic structure of an inert element. Due to the electrostatic energy produced
when the positive Na\(^+\) and the negative Cl\(^-\) ions either crystallize or solvate, it is energetically favorable for these two atoms to ionize and then form an ionic compound. Linnett points out that the octet rule is not universal and it is really the energy considerations that are of importance. The octet rule is a general qualitative parameter.

This type of chemical bond is the ionic bond, so called because of the ion like character of the atoms involved, due to the lending and borrowing of electrons between atoms. The atoms are held together by electrostatic forces.

Another type of chemical bond is the covalent bond. This bond involves the sharing of electrons between two (or more) atoms so that each of the atoms will have an octet structure. For example, chlorine needs one electron to fill its outer shell. That is, chlorine has one orbital available for an additional electron, the spin of which is opposite the spin of the electron already in that orbital. If two chlorine atoms come together each one needing an electron of opposite spin to fill their shells, then they will combine by sharing their odd electrons with each other. The two atoms are then bonded to each other because of the concentration of electric charge between them. The electrons around each of the nuclei still correspond to two sets of tetrahedrons.

This model is similar in many ways to the one proposed by Lewis except that the tetrahedrons of electrons are emphasised and electron pairs (which are of great importance in the Lewis theory) are of secondary importance. Linnett shows how one electron bonds exist and are sometimes energetically favored over electron pairs bonds.

LCAO-MO models stress the overall symmetry of the molecule and consider this to be the most important aspect in the construction of molecular orbitals. Linnett's model is more like the valence bond method which is
mostly concerned with local electron effects in that they both interpret electrons to be particle-like in nature.

**Section 8; Daudel's Model of Loges**

The purpose of this model is to be able to describe the structure of atoms and molecules based on the positioning of the electrons involved in each system. The model is based on and derived from wave mechanics, see (5).

The electron system is broken up into a series of volumes \( V_i \), called loges. Then the probability, \( P_i \), that one and only one electron in a given spin state will be found in \( V_i \) is calculated. If \( V_i \) is large, then there will probably be more than one electron of a given spin state in it. On the other hand if \( V_i \) is small, then it may be empty. Thus it seems reasonable that there is an optimum value to the size of \( V_i \). If \( p \) is the number of electrons of a given spin, then the electron space is divided up into \( p \) volumes, \( V_i \), all of which connect, none of which overlap. \( P_i \) is the probability of finding one and only one electron of a given spin in \( V_i \).

\[
\eta = \sum_i P_i
\]

is the sum of the probabilities, over the entire system, of finding one and only one electron of a given spin state in each of the volumes \( V_i \). \( \eta \) must have a value between 0 and \( p \).

Defining the lack of localization, \( \eta \) to be

\[
\eta = \frac{(\rho - \eta)}{\rho}
\]

one can see that the smaller \( \eta \) is, the more localized the electron effects.

Allowing the wave function of the system to be

\[
\psi = (x_1, w_1, x_2, w_2, \ldots, x_j, w_j, \ldots, x_n, w_n)
\]
then the value of \( P_i \) in loge \( V_i \) for electrons in the \((\text{arbitrarily}) + \frac{1}{2} \hbar \) state is

\[
8-3 \quad P_i = C_p^p(\psi) \int_{V_i} d\rho^{(j)} \int_{\text{space} - \psi_i} d\rho \cdots d\rho_p \int_{\text{space} - 1/2} d\rho_{p+1} \cdots d\rho_n \times \int_{-\frac{1}{2} \hbar} \cdots d\psi_p \int_{\frac{1}{2} \hbar} \cdots d\psi_n / \psi_i^2.
\]

\( C_p^p \) is a normalization constant, \( p \) is the number of electrons in the \(+ \frac{1}{2} \hbar \) spin state, \( \int_{V_i} d\rho^{(j)} \) is the volume of the loge \( V_i \) in three-dimensional coordinate space, \( \int_{\text{space} - \psi_i} d\rho \cdots d\rho_p \) is the rest of the volume in relation to the same spin state \((+ \frac{1}{2} \hbar)\), \( \int_{\text{space} - 1/2} d\rho_{p+1} \cdots d\rho_n \) is the volume of the entire space in relation to the other spin state \((- \frac{1}{2} \hbar)\).

The integration over a specific character (spin) shows the dependence of \( V_i \) on the specific spin state. Physically there is more information available concerning the ensemble of electrons if the probability of \( P_i \) is known as a function of both coordinate space and character space (1).

The model is used as shown below. First the different possibilities of dividing up the electron space are postulated. In the case of atoms this is usually dividing the space up by passing planes through the nucleus or concentric spheres around the nucleus. In each case the value of \( P_i \) for each loge \( V_i \) is calculated from equation (8-3) and then the lack of localization \( \eta \) is computed. The different \( \eta \) -values are then compared and the system of loges corresponding to the smallest value of \( \eta \) is chosen as the best system to represent the electronic structure of the atom.

The energy of any proposed system of loges can be calculated by finding (see Section IV and Appendix VII)

\[
8-4 \quad \tilde{E}_j = \frac{\langle \psi / H_j / \psi \rangle}{\langle \psi / \psi \rangle}
\]

where \( \tilde{E}_j \) is the average energy associated with electron \( j \), \( H_j \) is the Hamiltonian corresponding to the \( j \)th electron, and the integration takes place over the volume of the loge electron is contained in. The
total energy is then
\[ E_T = \sum_{j=1}^{n} E_j. \]

When the best system of loges is found to be concentric spheres, this corresponds to the classical concept of electrons in shells. If the best system of loges is found to be the separation of the space by a plane passing through the nucleus, this corresponds to the classical idea of two electrons in the same shell. There can be a combination of these two systems in a given atom.

If the average electrical potential energy in a loge \( V_i \), where \( V_i \) is any arbitrary loge, is \( E_i \) and the volume of \( V_i \) is \( V_i \), then \( (V_i)(E_i)^{3/2} \) is a constant. This fact is consistent with the theoretical conclusions of the Thomas-Fermi statistical model (see Appendix V).

In the molecular case, two types of loges are distinguished. These are loges of the core, which are pretty much like atomic loges, and bonding loges. The core loges are determined in the same fashion as the atomic loges were. The rest of the electron space of the molecule is considered to be taken up by the bonding loge. Depending upon the values obtained, the proposed systems of loges can be compared and the number of electrons in the bond can be determined. The bonding loges corresponding to \( \sigma \) bonds and \( \pi \) bonds are not distinguishable by any characteristics intrinsic to the particular loges. The determination of these must come from energy and symmetry considerations.

It is not unreasonable to assume that loge \( V_i \) of character \(+\frac{1}{2}\alpha\) extends over the same volume of electron space as loge \( V_{i+p} \) of character \(-\frac{1}{2}\alpha\). If the volume over which a loge extends is thought to be an electron orbital, then the two electrons of opposite spin within that volume would correspond to the two electrons of different spin allowed in the same orbital by the exclusion principle. Thus the Lewis concept of electron pairs which is fundamental to most chemical thinking (whether or not
it has a basis in physical reality) can be accounted for.

This model combines some of the attributes of different models and theories. First of all, it is similar to valence bond in its emphasis on the importance of localized electron effects. Secondly, the electron is best interpreted from this model as being an electron cloud of varying density (hence all $p_i \neq 1$). Thirdly, it is consistent in many ways with the other theoretical models derived from wave mechanics along different lines. An example is the consistency with the Thomas-Fermi model in the aspect previously mentioned.
Conclusion

From the preceding text, it can be seen that quantum mechanics is indeed applicable to chemical theory. Quantum mechanics provides the basis of the theory behind the concepts of electron orbitals, bonding, molecular geometry, and atomic interactions. Chemical theories are based on the results and theory of quantum mechanics and the theoretical predictions are compared to empirically known facts.

In its present form, quantum mechanics cannot exactly describe atomic and molecular systems for primarily two reasons. 1.) When it attempts to describe systems on so large a scale, some of the subtle quantum effects are approximated over due to the approximations necessary to describe so large a system. 2.) Atomic and molecular systems of interest usually contain more than two interacting "bodies" (i.e. protons, electrons, neutrons). It is possible to mathematically solve for the effect of one body on another (and vice versa). This is a two-body problem. It is not, as yet, possible to solve for the exact effect of many bodies on each other (the many body problem). Hence it is possible to solve Schrödinger's equation exactly for the hydrogen atom (two bodies) whereas approximations are needed for any system of greater complexity (many bodies). Until the many body problem is solved, it will not be possible to apply quantum mechanics directly (without approximations) to problems of chemical interest.

Theories such as the LCAOMO and the valence bond models are based on quantum mechanical effects and empirical results. These theories are constructed so that they explain (mathematically) how the electrons in atoms and molecules behave and from this properties of chemical interest (bonds, bond lengths, relative energies, etc.) are predicted. These are not the only two molecular orbital models. The free electron
molecular orbital model (FEMO), for example, is used to explain conjugated organic molecules. The two models presented are, however, the most commonly used.

Molecular orbital models are useful constructs which can be applied to unknown systems and yield meaningful predictions.

Models such as Linnett's and Daudel's differ from the molecular orbital models in many ways. The molecular orbital models are primarily mathematical models and not supposed to be interpreted as being "reality." They are an empirically conceived mathematical description from which "reality" is predicted. Linnett's and Daudel's models are physical models. They attempt to give a physical description of the electronic structure of atomic and molecular systems. These models are based on the theory of quantum mechanics and other physical effects. They are used to predict molecular geometries and other properties associated with molecular systems. Their advantage is in that they are able to consider the "outer most shell" of the atom as a whole entity instead of breaking it up into a series of one electron problems.

The mathematically descriptive models are more successful in their predictions than the physically descriptive models. However, the mathematical models have been extended about as far as they can go. In the present state of affairs, in order to get more accurate results from these models an unreasonable number of variables must be included in the equations. In effect a mathematical model of a particular system is constructed by using the great number of variables and the ability to make optimum generalizations is then lost. The physically descriptive theories, on the other hand, have much potential. Although they do not have the same general accuracy as the mathematical models, they have not been so highly refined as those models. When they are further refined
and made more mathematically precise, their use will greatly strengthen the predictive ability of theoretical chemistry. The author feels that the future of theoretical chemistry lies in this type of physical theory (not necessarily the two presented in this paper) and that of the Thomas-Fermi atomic model (see Appendix V).
Appendix I: Hilbert Space and the Mathematics of Quantum Theory

Hilbert Space

Hilbert Space is an infinite dimensional vector space the elements of which are complex-valued functions of a real variable \( x \), defined on the interval \([a, b]\) (\([a, b]\) is a closed interval and is the set of points \( x \) such that \( a \leq x \leq b \)) and which are square integrable (i.e. 
\[
\int_a^b |F(x)|^2 \, dx
\]
exists and is finite). Some other properties of this space are that it is linear (i.e. if \( \psi_1 \) and \( \psi_2 \) are two complex square integrable functions, and \( \alpha \) and \( \beta \) arbitrary complex scalars, then \( \psi = \alpha \psi_1 + \beta \psi_2 \) is also a complex square-integrable function) and within this space a scalar or inner product can be defined. Therefore the following hold.

1.) \( \langle \psi, \psi \rangle = \langle \psi, \psi \rangle^* \)

where \( \langle \psi, \psi \rangle^* \) is the complex conjugate of \( \langle \psi, \psi \rangle \).

2.) The scalar product of one function by another is linear or

\[
\langle \psi, \alpha \psi_1 + \beta \psi_2 \rangle = \alpha \langle \psi, \psi_1 \rangle + \beta \langle \psi, \psi_2 \rangle.
\]

3.) The norm of a function within the space exists and is a real, non-negative

\[
\langle \psi, \psi \rangle = N \psi > 0
\]

If \( \langle \psi, \psi \rangle = 0 \), then \( \psi = 0 \).

Another property of Hilbert space is its completeness. The criteria for completeness is that every Cauchy sequence of vectors must converge to a limiting vector within the space.\(^1\)

1. A Cauchy sequence is a sequence of vectors called \( \psi_n \) which have the property of \( \| \psi_n - \psi_m \| \to 0 \) as \( m, n \to \infty \). If a sequence of vectors converges to a limit, then it is a Cauchy sequence. If in the limit \( \psi_n \to \psi \), then \( \| \psi_n - \psi_m \| = \| \psi_n - \psi + \psi - \psi_m \| \leq \| \psi_n - \psi \| + \| \psi - \psi_m \| \to 0 \) as \( m, n \to \infty \). The converse is also true (i.e. all Cauchy sequences converge).
The convergence must be, of course, towards a square-integrable function, by the very nature of being a complete space; manipulations of the functions in the space will result in functions which are also in the space.

The properties of vector spaces in general of which a Hilbert space is one type are to be found in any textbook on linear algebra (8).

A Hilbert space is said to be separable if it has a basis which is orthonormal in nature and the orthonormal vectors in the basis are countable. That is, if it has a basis \((\varphi_1, \varphi_2, \ldots)\) and \(\varphi = (\varphi_1, \varphi_2, \ldots)\), the number of vectors in the basis may be finite or infinite and if there exists one countable orthonormal basis for a Hilbert space, all orthonormal bases are countable for that Hilbert space.

Hilbert space is the mathematical background of quantum mechanics. Functions in Hilbert space are used to analytically describe physical states and operators in Hilbert space are analytically used to describe physical observables. The completeness of Hilbert space is essential to quantum theory.

Projection Operator

Consider a Hermitian operator, \(A\) whose eigenvalues and eigenvectors are \(\lambda_n\) and \(\varphi_n\) (where \(n = 1, 2, \ldots, N\)) respectively. Define the operator \((I)\) to be an operator such that

\[
I \varphi = \sum_{n=1}^{N} \lambda_n \varphi_n (\varphi_n, \varphi)
\]

for any \(\varphi\) in a finite dimensional vector space. Because \(A\) is a Hermitian matrix, there exists a unitary matrix \(U\) such that \(U^{-1}AU\) is a diagonal matrix (the elements of the diagonal of \(U^{-1}AU\) are the characteristic roots, \(\lambda_n\)). Thus there always exists a complete set of orthonormal basis vectors which will diagonalize a given Hermitian
linear transformation $A$. It follows that $\sum_{n=1}^{N} \Phi_n (\Phi_n, \psi) = \psi$ which means that $\psi$ is an eigenfunction of $A$. (I) therefore is the identity operator.

Defining $p_n \psi = \Phi_n (\Phi_n, \psi)$ then $\mathbf{I} = \sum_{n=1}^{N} p_n$. From this it can be seen that $p_n$ is idempotent ($p_n^2 = p_n$), Hermitian ($p_n^+ = p_n$), and orthogonal ($p_m p_n = 0$ if $m \neq n$). $p_n$ is called a projection vector because it projects any vector into the one-dimensional subspace (of vector space $V$) spanned by $\Phi_n$.

If an operator $B$ is defined by $B = \sum_{n=1}^{N} \mathcal{P}_n \Phi_n$ and because $B \psi = \sum_{n=1}^{N} \mathcal{P}_n \Phi_n (\Phi_n, \psi)$ it is obvious that $B$'s eigenvalues are $\mathcal{P}_n$ and its eigenvectors $\Phi_n$. $B$ is therefore equal to $A$.

Spectral Theorem (taken primarily from (4))

The spectral theorem is of great importance to the developing of the theory of quantum mechanics. It is the formalism for one of the axioms.

Given $A$, a Hermitian operator, then there exists an unique operator-valued function, $\mathcal{E}(\lambda)$ such that:

1.) $E(\lambda) E(\lambda) = E(\lambda)$, $\bar{\lambda} = \min (\lambda_1, \lambda_2)$
2.) $\lim_{\lambda \to -\infty} E(\lambda) = 0$, $\lim_{\lambda \to \infty} E(\lambda) = \mathbf{I}$
3.) $\mathbf{I} = \int_{-\infty}^{\infty} \lambda \, dE(\lambda)$
4.) $\lambda = \int_{-\infty}^{\infty} \lambda \, dE(\lambda)$.

$E(\lambda)$ is the resolution of the identity corresponding to $A$. The set of points where $E(\lambda)$ is inconsistent is called the spectrum of $A$ and for all $(\lambda)$, $E(\lambda)$ commutes with $A$ on any transformation that commutes with $A$.

In finite space this reduces to the completeness theorem.
Appendix II: The Statistical Interpretation of Quantum Mechanics of the Copenhagen School

The first physical interpretation of the Schrödinger formalism for quantum theory \( \psi = \psi' \) was proposed by Schrödinger. It said that \( \psi \) was electromagnetic in nature and corresponded to a continuous distribution of electricity in space. Experiments in the scattering processes showed that it was incorrect to conclude that \( \psi \) corresponded to a group of waves. These experiments showed that the electron had the attributes of a particle as it was always maintained as a whole.

Following this erroneous interpretation of the Schrödinger equation, Born and Heisenberg came out with a probabilistic interpretation. This conception determined that \( |\psi|^2 \) corresponds to the probability of finding an electron in a given space at a given time. The theory of interpretation was based on three assumptions. The first was the particle nature of the electron, and thus the wave equation must be based on the fact that the particles are in existence.

The second basis for the theory was that this "quantum mechanical" probability of wave functions was not quite the same as the classical probability. That is, classical probability gives a total probability of \( P_A + P_B \) for the superpositioning of two wavefunctions \( \psi_A \) and \( \psi_B \) of probability \( P_A \) and \( P_B \) (probability of \( \psi = \psi_A + \psi_B \) is \( P_A + P_B \) respectively. Quantum mechanical probability gives a total value of \( (\psi_A^* \psi_B + \psi_A^* \psi_B^*)^2 = P_A^2 + P_B^2 + \psi_A^* \psi_B + \psi_A^* \psi_B^* \), where the two new terms are usually not equal to zero.

Thus probability is a fundamentally real entity (being) and is not a construct. Nevertheless, it is not real in the sense of being affected by physical processes such as energy or momentum transmission. If this is so, then this probability must not be real in the sense of being an attribute or property of a process, rather it must be real in the sense of being an attribute or property of a process, rather it must be real in the sense of being an attribute or property of a process, rather it must be real in the sense of being an attribute or property of a process.
being a fundamental process itself.

The final assumption follows from this. The probability in the quantum mechanical sense is a reality which is intermediate between the realities of that which exists as a physical attribute (matter) and that which could possibly exist (the total lack of attributes or the addition or subtraction of any attributes presently existing).

In this interpretation, quantum mechanics does not measure individual occurrences, rather it measures the logically deduced probability of such occurrences. This leads to two of the axioms of quantum mechanics; spectral decomposition and the interference of probabilities, (10), (4).

It was Born's probabilistic interpretation that led to the transformation theory of quantum mechanics. Transformation theory is the study of transformations which must exist between the different mathematical formalisms of quantum mechanics because of the identity of their empirically significant results. Von Neumann developed a transformation theory which results in a mathematically formulated quantum mechanics that is based on a separable Hilbert space and therefore the orthonormal vectors of the orthonormal basis are countable. It also follows from this that variables forming this space must be quadratically integrable functions.

This transformation theory is based on the statistical formalism of quantum mechanics. In essence, it says that if $R = \{R_1, R_2, \ldots R_j\}$ and if $S = \{S_1, S_2, \ldots S_j\}$ are two sets of commuting Hermitian operators corresponding to real variables and the first set has the $(E_1(R), \ldots)$ as its resolution of the identity and the second set has the $(E_2(R), \ldots)$, then the probability that the elements of $S$ lie in the intervals $\{J_1, \ldots\}$ if the elements of $R$ lie in the intervals $\{I_1, \ldots\}$ is given by

$$\left[ E(I_1) \ldots \ldots E(I_n) F(J_1) \ldots \ldots F(J_m) \right].$$
Unfortunately this formalism applied to systems with continuous energy states leads to results not closed by a separated Hilbert space. These results do not normalize to quadrat cally integratable basis vectors.

The formalism as applied strictly to non-relativistic quantum mechanics (which works with eigenenergies) is logically consistent and complete. There is a definite correlation between the theoretical predictions of this theory and empirical results. Unfortunately the premises of the theory are not all so firmly grounded in empirical experience. Because of the anticommuting nature of the position and momentum operators and thus the unclassical nature of position and momentum, concepts like position and velocity in quantum mechanics could not have their classical meanings. Therefore the results of Born's statistical interpretation which are dependent upon the ability to localize particles, are questionable because the ability to localize particles is based on questionable assumptions concerning the nature of position.

Realizing this state of affairs, Heisenberg developed a formalism that restricted the classical definitions of these, anticommuting attributes and was less demanding in its assumptions concerning the fundamental nature of space.¹

Incorporated into Heisenberg's formalism is the Heisenberg uncertainty principle which is the physical interpretation of the mathematical fact that the Hamiltonian operators corresponding to position and momentum do not commute, i.e., $[p, q] = -i\hbar$. This uncertainty

¹. The Schrödinger and interaction mathematical formalisms of quantum theory assume that space in continuous. The Heisenberg picture does not restrict space to either a continuous or a discontinuous nature.
also exists between any two anticommuting Hermitian operators corresponding to real variables.

Heisenberg quantum mechanics says that the electron mathematically resembles a wave packet in configuration space that is described by the superpositioning of eigenfunctions of the possible states of the electron. Heisenberg intuitively believed that an electron is a quantum mechanical particle as opposed to being a classical particle in the same way momentum and position are different in the quantum mechanical and classical senses. In the case of an electron in an atomic system, the wave packet mathematically describes a classical orbit which spreads with time. As the position of the electron is more fully determined, the time spreads with time. As the position of the electron is more fully determined, the time spread is reduced and a smaller wave packet produced. Thus the orbit of an electron can be thought of as the time dependent sequence of the spatial locations where the wave packet was observed.

The Heisenberg interpretation disavows the determinacy of reality. It says that because all we can measure and therefore know about is restricted by the uncertainty principle, any speculations on a hidden, more fundamentally real existence behind the statistical reality of perception is meaningless. The determinists' view of reality is that exact knowledge of the present can be found and is ruled by causality. In the Heisenberg approach, predictability is restricted by the uncertainty of exact values.

The principle of causality, which is that natural phenomena obey exact laws, follows from predictability and therefore has no operational significance. This means that the difference between states of a system in two succeeding moments are linked by cause and effect but there are a number of equally probable effects due to a given cause.
For Bohr, the exact physical nature of the electron in space presented a paradox. On one hand he believed that the electron had a particle nature due to the nature of the Bothe-Geiger experiments. These experiments showed that the electron had a definite particle manifestation. On the other hand he felt that only physical description of electrons, consistent with the mathematics of the analytical approach of Schrödinger's radiation theory, was a wave picture. The consideration in the corpuscular case is that by performing experiments the phenomena being measured is influenced. Hence there can be no relation of a specific cause to a specific effect if, as in quantum mechanics, the very act of measuring changes the relationship between the suspected cause and the suspected effect. The quantum mechanical observations (experiments) of an object show those attributes which are particle like in nature. Nevertheless, in the analytical approach, the mathematical formalisms of Schrödinger which mathematically describe non-relativistic quantum effects cannot be physically interpreted in terms of relations between objects in space and time. To quote Heisenberg (7), "There exists a body of exact mathematical laws, but these cannot be interpreted as expressing simple relationships between objects existing in space and time. The observable predictions of this (quantum) theory can be approximately described in such terms, but not uniquely. The wave and corpuscular pictures both possess the same approximate validity."

1. Mathematically the limit of the well determined wave packet spread out in time is the quanta (particle) whose position is a function of time. However this limit is not approached unless some measurement is made.
De Broglie's equation for the electron, $\frac{h}{\lambda}$ relates momentum, which is associated with particles, to wavelength which is associated with the wavelike propagation of electromagnetic radiation and Planck's equation for the photon, $E=hf$ relates frequency (waves) and energy (particles). This relation and the above considerations lead Bohr to postulate his complementarity concept. When two sets are mutually exclusive, but both are needed in order to completely describe some system, the logical relation between them is a complementarity relation. This uncertainty as to the model of the particular process is caused by the fundamental nature of observation. For one thing, the concept of observation is primarily a classical notion and hence certain limits must be on it so that it can be meaningful in quantum mechanics. These limitations are governed by the uncertainty principle.

The interpretation presented so far in this chapter is what was eventually formalized into the "Copenhagen" interpretation of statistical quantum mechanics. This interpretation maintains that classically the structural properties of a closed system in phase space can be determined from the constituent parts of momentum and position in that space. The quantum postulate says that in order to measure a system, the system must be disturbed and therefore must be in an open state. Thus the structural properties cannot be determined for the quantum mechanical case in the classical sense. If the closed system restrictions are lifted, then the classical notion of causality is invalidated.

This is the interpretation of quantum mechanics accepted by most physicists. However, it is of great importance to point out that some of its bitterest enemies (alive and dead) were Einstein, de Broglie, Schrödinger, Landé, Duane, and Bohm. Their objection is not on the basis
of what the theory predicts, it is that the theory does not go far enough in the deterministic sense and precludes the possibility of anything existing beyond it and hence cannot be tested. These critics argue that if some hidden variable existed that governed the indeterminacy principle, one could not conceive of that variable within the system of physics that preached indeterminacy. Landé (12) further questions whether or not the wave–particle duality in this interpretation was originally meant to be taken for fact or if it was at first a construct device.
Appendix III: The Hidden Variable Interpretation of Quantum Mechanics

At the same time that Heisenberg, Bohr, and Born came out with the Copenhagen approach, de Broglie suggested an alternative interpretation of quantum mechanics (1927). De Broglie's physical interpretation of the Schrödinger equation is that the equation lends itself to two different types of solutions, both of which are physically valid in some way. The first is of a statistical or probabilistic realism and the second is a description (measurement) of the attributes by which a particle is known. These two solutions are linked by the conceptualization that though there is a localization of particles in space, these particles are dependent in their motion on all objects that interfere with the free propagation of the wave phenomena associated with the particle. This causal interpretation was rejected at the time it was presented because of its seemingly sloppy approach compared to the high powered Copenhagen approach. Partially due to its cold reception, de Broglie abandoned it and supported the Copenhagen interpretation, not because he believed it, but because it was the best around. However, twenty five years after its presentation, de Broglie went back to his causal interpretation. This reversal on his part was largely due to general relativistic considerations (not discussed here) and the hidden variable theory of Bohm.

Bohm's theory (3) attempts to show that the indeterminant picture of the Copenhagen interpretation is just a special case of a more general reality.

The Copenhagen interpretation is based on the assumptions—

1. The most that the state of any system can be determined is given by the probability interpretation of the wave function.
2. The measurement of a quantum affects the state of the quantum and it is impossible to know what happened to the quantum between two successive measurements.

Although these assumptions and the mathematical formalism of the Copenhagen interpretation form a consistent view which is experimentally valid down to the limit of $10^{-13}$ cm., it does not mean that no other interpretations can be made from these assumptions.

If the assumptions of a theory give rise to a unique mathematical formalism, then the experimental results testing this theory either agree with or contradict the theoretical predictions. If the results contradict then the assumptions are shown to be wrong. Agreement between results and predictions is an indication of the theory's accuracy. If no unique mathematical formalism results, then contradiction shows that the formalism is incorrect and agreement indicates that the formalism may be the correct one or a more specific case of the correct one.

The assumptions of the Copenhagen theory do not imply a unique mathematical formalism. Hence the theoretical predictions resulting from the mathematical formalism used (which is not the only one possible) are not absolute and so experimental results cannot be used to judge the ultimate validity of the theoretical predictions or the assumptions upon which they are based.

Unfortunately the Copenhagen theory is often viewed as being the unique mathematical apparatus of the assumptions. For example, these assumptions do not restrict changes in the Hamiltonian operator such as the possibility of new varieties of meson fields with few restrictions on their attributes. Other changes in the mathematical formalism can be made and these can direct the predictions of the theory (and still not
have any basic differences from the original physical picture), so that they will be consistent with experimental results.

Bohm maintains that the only way to stay out of this sort of circular postulating is to determine the effect of assumptions that violate these assumptions. Let it be assumed that there is a deterministic mechanism that governs the action of a quantum between two successive measurements and this mechanism is dependent upon "hidden" variables. Experiments could be devised to determine the state of these "hidden" variables and if the predictions are shown to be valid, then there is good reason to think these hidden variables exist as they are theorized. If the predictions are shown to be invalid, then the hidden variables are incorrectly theorized (the specific character of the hidden variables must be changed in the theory formulation).

The assumptions behind the hidden variable interpretation are:

1. The wave function \( \psi(x) \) corresponding to a corpuscular electron is the mathematical representation of an objectively real field. The action of the \( \psi \)-field on a particle is comparable to the action of an electro-magnetic field on a charge. Just as the electromagnetic field is subject to Maxwell's equations, the \( \psi \)-field is subject to Schrödinger's equation. \[1\]

2. The momentum of the particle is restricted to \( \hat{p} = -i \hbar \nabla \psi(x) \). This value for the momentum comes from classical mechanics. Consider an en-

1. There are also many differences between the electromagnetic field and the \( \psi \)-field. Maxwell's equations are inhomogeneous in the electromagnetic field whereas Schrödinger's equation is homogeneous in the \( \psi \)-field. Inhomogeneity is a necessary condition for the existence of radiation and thus the homogeneous \( \psi \)-field has no radiation associated with it.
sembl of particle trajectories, all of which are solutions to the equation of motion. If all of these particle trajectories are normal with respect to any of the surfaces of $S$, then they are normal to all surfaces of $S$ (where $S$ is a sphere). It follows then that $\nabla S(x)_{/m} = \frac{\partial \xi}{\partial \xi}$, the velocity of any particle passing point $(x)$. $S(x)$ is the solution of the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial \xi} + \left(\frac{\partial S}{\partial \xi}\right)^2 + \sqrt{\mathbf{V}^2} = \frac{\hbar^2}{4m} \left(\frac{\partial^2 \rho}{\partial \xi^2} - \frac{1}{2} \frac{(\nabla \rho)^2}{\rho^2}\right) = 0.$$ 

This equation and the Schrödinger equation have a definite correspondence.

The derivation of the Hamilton-Jacobi equation is too involved to go into here but can be found in (16). This equation is interpreted to mean that the surface of constant action of a system of particles and the surface of constant phase in optics are analogous in their propagation where the wave vector corresponds to momentum and the frequency corresponds to energy.

3. Statistical analysis is used to describe the particle (i.e. its location and momentum, or energy and time) but the statistical formalism is not fundamentally real; it is used only because of our ignorance concerning the initial conditions surrounding the particle. The statistical ensemble has a probability density $\rho(x) = |\psi(x)|^2$.

Bohm (3) shows how the use of these assumptions yields the same predicted results for experiments as does the Copenhagen interpretation. The Heisenberg uncertainty principle holds true in this interpretation not as a fundamental physical axiom, but rather as a practical limit of the ability to observe because of the ways by which we observe. This limit will be surpassed when the hidden variables have been discovered.

One of the basic objections to this interpretation is that the hidden variables involved are completely unknown. Bethe says (2), "It has been
suggested (Bohm, de Broglie) that there exist hidden variables in terms of which causal descriptions could be effected. These variables are 'hidden' in the sense that they do not affect the energy eigenvalues. The existence of identical particles and composites show that such hidden observables cannot have any observable consequences (as in the example of radioactivity) and are therefore empty. It assures us that the present description must be complete," (i.e. the statistical Copenhagen interpretation).

The fallacy in Bethe's argument lies in his initial assumption of the ultimate validity of the Copenhagen approach. Using the identical particle of the Copenhagen school, he does not allow for the possibility that these particles might be shown to be non-identical in terms of a more basic physics, and hence distinguishable from one another by some hidden variable. In terms of the physics allowed by quantum mechanics as interpreted by the Copenhagen school, hidden variables could not be detected. In order to detect them if they did exist, one would have to go beyond this physics to a physics that allows for their possible existence (whether or not they exist). In one sense, the insistence upon indeterminancy by the Copenhagen school borders on mysticism. The origin of the concept may have both empirical and analytical (anticommuting Hermitian operators) bases, but the restrictions in the physics built around it makes it an absolute. As such there is no conceivable experiment to disprove the assumption of no-hidden variables in the realm of quantum theory (Copenhagen interpretation) and therefore it has a quasi-mystical status.
Appendix IV: Quantum Numbers

When simultaneous measurements are made on different attributes, $A_1, A_2, \ldots, A_k$, of an object (i.e. particle) and these measurements do not affect each other, the linear operators of these attributes are commutable. In such a case these attributes are called compatible. A complete set of compatible attributes is one in which every attribute in the set is compatible with every other attribute in the set and there is no attribute outside the set which is compatible with every one of the attributes in the set. There is, at present, no experimental criteria for determining whether or not the requirements for a complete set of attributes have been fulfilled.

A complete measurement on an object is a series of measurements which give definite values $a_i^{(k)}$ of a maximum number of compatible attributes $A_k$. When the compatible attributes, $A_k$, of an object are known to have the values $a_i^{(k)}$, the object is in a pure quantum state characterized by the quantum numbers $a_i^{(k)}$. By the nature of the attributes, the state of the object can be described in terms of the quantum numbers.
Appendix V: Thomas–Fermi Statistical Approximation

According to Fermi–Dirac statistics of classical particles, the number of stationary states within an energy band of \((E, E + \sqrt{\varepsilon})\) is proportional to the volume in phase space occupied by the band with a proportionality factor of \(\frac{2}{\hbar^3}\). The density in phase space of \(Z\) classical electrons, \(n(r,p)\) is

\[
\text{AV 1} \quad n(r,p) = \frac{2}{\hbar^3} \quad \text{if} \quad \varepsilon < \varepsilon_0
\]

or

\[
\text{AV 1} \quad n(r,p) = 0 \quad \text{if} \quad \varepsilon > \varepsilon_0
\]

where \(\varepsilon_0\) is the energy of the highest energy level occupied.

In the case of an atom in the ground state, let \(\rho(r)\) be the fraction of the electrons that can be found in a volume element \((r, r + dr)\). Thus \(\rho(r)\) is the probability density for finding an electron in \((r, r + dr)\). Assuming spherical symmetry and normalization to give the total number of electrons, i.e.

\[
\text{AV 2} \quad 4\pi \int_0^\infty \rho(r) r^2 dr = Z
\]

the following can be said concerning this system. Because of the charged particles in the atom an average electrical potential, \(\overline{E}(r)\), is created. The effect of the nucleus is that of a point charge at the origin equal to \(Ze\) and the effect of the electrons (assumed to be in a continuous distribution over the atomic space) is to produce a continuous distribution of negative electrical charge density equal to \(-e\rho(r)\).

Despite the assumption of the continuous distribution of electrons, when \(Z \gg 1\), the effect of one electron is small compared to the effect of all the electrons. In the approximation of making one of these electrons a classical electron, the potential that would act on the resulting classical electron is \(-e\overline{E}(r)\).

In the ground state of an atom, the \(Z\) electrons occupy the \(Z\) lowest energy levels. Therefore the overall probability density \(\rho(r)\) is just
the sum of all the probability densities, $|\Psi|^2$, each corresponding to one of the $Z$ lowest levels, i.e.,

$$A\,V\,3 \quad \rho(r) = \frac{Z}{\rho^2} / \mathcal{V}_x / \mathcal{V}_y / \mathcal{V}_z.$$

Because these levels are determined by the field $-e\mathcal{E}$ in which they exist, it follows that there must be some sort of functional relation between the field $-e\mathcal{E}$ and the probability density $\rho(r)$.

Using Fermi-Dirac statistics to describe the atomic system with

$$E = \frac{p^2}{2m_e} - e\mathcal{E}$$

and arbitrarily setting $E_0 = 0$, then

$$A\,V\,4 \quad \rho(r) = Z \int n(r, \rho) \ d\rho = \frac{Z}{\hbar^3} \int_{r < \rho} d\rho.$$

Substituting for the value of $\rho(r)$ and performing the integration results in

$$A\,V\,5 \quad \rho(r) = \frac{8\pi^2}{3h^3} \left( \frac{2m_e e}{\hbar^3} \right)^{3/2}, \mathcal{E} > 0$$

$$r = 0, \mathcal{E} < 0$$

The electrical potential, $\mathcal{E}$, can be described mathematically as in

$$A\,V\,6 \quad \Delta \mathcal{E} = \frac{1}{r} \left( \frac{d^2}{dr^2} r \right) \mathcal{E} = 4\pi Ne \rho$$

noting that at the origin,

$$A\,V\,7 \quad \lim_{r \to 0} (r \mathcal{E}) = Ze,$$

where the inhomogeneous differential equation is of the form of a Poisson equation.

Substituting the value of equation $A\,V\,5$ into the right most term of $A\,V\,6$ gives a second order differential equation for $\mathcal{E}$,

$$A\,V\,8 \quad \Delta \mathcal{E} = (4\pi Ne) \left( \frac{8\pi^2}{3h^3} \right) \left( \frac{2m_e e}{\hbar^3} \right)^{3/2}$$

which in atomic units of $e = m_e = h = 1$ is

$$A\,V\,8 \quad \Delta \mathcal{E} = \left( 8 \frac{2\pi^2}{3h^3} \right) \left( \frac{2m_e e}{\hbar^3} \right)^{3/2},$$

Equation $A\,V\,8$ is the fundamental equation of the Thomas-Fermi model (method).

The distribution of electrons in the atomic system is given by this equation providing the solution is symmetric with regard to the center and obeys the boundary conditions of as $r \to 0$, $\mathcal{E} r \to \mathcal{Z}$ and as $r \to \infty$, $\mathcal{E} r \to \infty$. 
In order to determine $\mathcal{E}$ and $\rho$, $r$ and $\mathcal{E}$ are replaced by $x$ and $\kappa$ which are found from

\begin{align*}
A V 9 & \quad r = x \cdot b \cdot \frac{2}{z} \cdot \frac{1}{\sqrt[3]{\sqrt{b} \cdot \kappa}} \\
A V 10 & \quad b = \frac{2}{z} \cdot \frac{3}{4} \cdot \frac{\kappa}{\rho} \cdot \frac{1}{\sqrt[3]{\sqrt{b} \cdot \kappa}} \\
A V 11 & \quad \mathcal{E}(r) = \frac{\pi}{\kappa} \cdot \left( \frac{r \cdot \frac{2}{z} \cdot \frac{1}{\sqrt[3]{\sqrt{b} \cdot \kappa}}} {b} \right) = \frac{\pi}{\kappa} \cdot \frac{\frac{2}{z} \cdot \frac{1}{\sqrt[3]{\sqrt{b} \cdot \kappa}}}{x} \cdot \kappa(X)
\end{align*}

After mathematical manipulations the equation

\begin{align*}
A V 12 & \quad \kappa^{1/2} \left( \frac{\sigma^2 \kappa}{\sigma x^2} \right) = \kappa^{s/2}
\end{align*}

is derived in place of equation A V 9. This equation has the boundary conditions of $\kappa = 1$ at $x = 0$ and $\kappa = 0$ at $x = \infty$. In terms of the model the second boundary condition means there is no boundary to the atom.

At $x = 0$, $\kappa'(x) = \kappa'(0) = -1.59$ and from this the potential $\mathcal{E}(r)$ is found to be $\mathcal{E}(r) = -1.80 \cdot \frac{2}{z} \cdot \frac{1}{\sqrt[3]{\sqrt{b} \cdot \kappa}}$ in atomic units or

$$
\mathcal{E}(r) = \frac{2}{r} - (1.80 \cdot \frac{e}{m_e} \cdot \frac{c}{2} \left( \frac{\frac{2}{z} \cdot \frac{1}{\sqrt[3]{\sqrt{b} \cdot \kappa}}} {x} \right))
$$

in cgs units. The electron density is found to be

$$
n(x) = \int r \cdot \frac{2}{z} \cdot \frac{1}{\sqrt[3]{\sqrt{b} \cdot \kappa}} \cdot \frac{1}{x} \cdot f(x) \quad \text{where} \quad f(x) = \left( \frac{\frac{2}{z} \cdot \frac{1}{\sqrt[3]{\sqrt{b} \cdot \kappa}}} {x} \right)^{3/2}
$$

These results show that the maximum electron density is at the same distance from the nucleus for all atoms. The semi-classical approximations lose their meaning at distances on the order of the inverse of the atomic number on the small side and at distances on the order of 1 a.u. on the large side. Thus the Thomas-Fermi equation is restricted to distances between these two extremes.

The ionisation energy of neutral atoms can be calculated from the Thomas-Fermi model with fairly good agreement to empirical data. This model does not account for the interactions between electrons.

Overall, the Thomas-Fermi method is primarily used to calculate effective potentials to be used in the SCF method as trial potentials.
Appendix VI: Pauli Exclusion

Given two eigenfunctions $\Psi_A$ and $\Psi_B$ and rotate the coordinate system which they are in. Allow them to undergo a linear transformation of the type

\[ \Psi_A = \alpha \Psi_A' + \beta \Psi_B' \quad \Psi_B = \gamma \Psi_A' + \delta \Psi_B' \]

where the $\alpha, \beta, \gamma, \delta$ and $\gamma$ are complex functions to the angle of rotation. If eigenfunctions $(\Psi_A', \Psi_B')$ and another pair of eigenfunctions $(\Phi_A', \Phi_B')$ were both transformed as in A VI 1, consider the equation

\[ \Phi_A' \Phi_B' - \Phi_B' \Phi_A' = (\alpha \sqrt{\beta}) (\Psi_A' \Phi_B' - \Psi_B' \Phi_A') \]

A VI 1 is a linear transformation that leaves A VI 2 invariant and as such is called a binary transformation. A spinor is a quantity that has two components both of which undergo a binary transformation upon rotation of the coordinate system. Spinor algebra has a spin component of $\frac{1}{2}$, $-\frac{1}{2}$, or zero associated with its eigenfunctions (in the case of the zero spin component, the spinor is really a scalar). This algebra is used in quantum mechanics to describe the dichotomic phenomena of the electron spin component in the direction of an applied magnetic field. A spinor can be thought of in terms of it being a vector with complex components in a two dimensional vector space.

The general spin operator

\[ \vec{\sigma} \cdot \vec{e} = \begin{pmatrix} \cos \gamma & \sin \gamma e^{-i\epsilon} \\ \sin \gamma e^{i\epsilon} & -\cos \gamma \end{pmatrix} \]

has the more specific forms corresponding to the Cartesian coordinates of

\[ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \]
These are called the Pauli spin matrices. Some of the properties of these matrices are:

1. \( \sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1 \)
2. \( \sigma_x \sigma_y = i \sigma_z \), \( \sigma_y \sigma_z = i \sigma_x \), \( \sigma_z \sigma_x = i \sigma_y \) (and therefore are unitary matrices)
3. \( \text{trace} \sigma_x = \text{trace} \sigma_y = \text{trace} \sigma_z = 0 \)
4. \( \sigma_i \sigma_j + \sigma_j \sigma_i = 0 \), \( (i,j) = (x,y,z) \), \( i \neq j \) (and therefore anticommuting)

These matrices are used as operators in quantum mechanics. They operate upon the part of the overall wave function corresponding to the electrons' spin states. Their function is to make the overall wave equation antisymmetric by making the spin part antisymmetric if the spatial part is symmetric or vice-versa. This is so that the wave function will change signs upon the interchange of two electrons which are degenerate except for their respective spins.
Appendix VII: Variational Principle

The Schrödinger equation can be derived from the variational principle as follows. Allow \( \phi = \psi + \psi' \) to be an approximation to the eigenfunction \( \psi \). This approximation varies from \( \psi \) on the first order. Considering the operator \( (H - E) \) it can be seen that

\[
(H - E) \phi = (H - E)(\psi + \psi') = (H\psi + H\psi' - E\psi - E\psi') = (H\psi - E\psi) - E\psi
\]

(because \( H\psi - E\psi = 0 \)).

Then the expectation value associated with this operator is

\[
\langle H - E \rangle = \frac{\langle \phi | (H - E) | \phi \rangle}{\langle \phi | \phi \rangle}
\]

Hence \( \langle H \rangle = E + \frac{\langle \psi | (H - E) | \psi \rangle}{\langle \phi | \phi \rangle} = \langle \psi | \psi \rangle \). This means that the trial function \( W(\phi) \) is only in error to the energy by the second-order term. However the original error in the eigenfunction chosen was in the first order and thus

\[
\sqrt{W(\phi)} = \sqrt{\langle H \rangle} = 0
\]

By showing that \( W(\phi) \) is an eigenfunction of the Hamiltonian operator, it can be seen that this leads to Schrödinger's equation. Start by finding the functions \( \phi \) such that

\[
Q[\phi] = \langle \phi | \psi' | \phi \rangle
\]

has a maxima or a minima with respect to small arbitrary changes in (i.e. it is stationary). Assuming that normality is obeyed \( \langle \phi | \phi \rangle \).

Let \( G[\phi] = \langle \phi | \phi \rangle - 1 \) be another functional and find the functional(s) \( \phi \) such that

\[
Q[\phi] = Q[\phi] - \lambda G[\phi]
\]

is stationary where \( \lambda \) is real and is known as the lagrangian multiplier. \( Q[\phi] \) will be stationary only when \( \sqrt{Q[\phi]} = \sqrt{\langle H \rangle} = 0 \). Substitut-
ing the values for \( A[\phi] \) and \( C[\phi] \) one gets that the first order variation of \( A[\phi] \), \( \sqrt{A[\phi]} \) is \( \sqrt{A[\phi]} = \langle \psi_1 | \psi_1 \rangle + \langle \psi_2 | \psi_2 \rangle - \sqrt{\langle \psi_1 | \psi_1 \rangle + 1} \langle \psi_1 | \phi \rangle - \langle \psi_2 | \phi \rangle \).

Calling \( \langle \psi_1 | \phi \rangle - \sqrt{\langle \psi_1 | \phi \rangle} = A \) and noting that \( \left[ \langle \psi_1 | \phi \rangle - \sqrt{\langle \psi_1 | \phi \rangle} \right]^* = \left[ \langle \psi_1 | \phi \rangle - \sqrt{\langle \psi_1 | \phi \rangle} \right] \),

so that \( \sqrt{A[\phi]} = A + A^* - \sqrt{\langle \phi | \phi \rangle + 1} \). Remembering that \( \langle \phi | \phi \rangle \) is normalized and thus \( \langle \phi | \phi \rangle = 1 \), then \( \sqrt{A[\phi]} = A + A^* \).

Thus for \( \sqrt{A[\phi]} \) to be stationary \( A \) and \( A^* \) must be equal to zero, i.e., \( \langle \psi_1 | \phi \rangle - \sqrt{\langle \psi_1 | \phi \rangle} = 0 \). Because the value of \( \sqrt{A[\phi]} \) is arbitrarily small, as previously assumed as a requirement of being stationary, \( \delta[\phi] - \sqrt{\delta[\phi]} = 0 \) or \( \delta[\phi] = \sqrt{\delta[\phi]} \) which is Schrödinger's equation.
Bibliography

Prefatory note—Not all of the sources in the bibliography were directly cited in the paper. Most of the sources listed are general sources, but they were the ones used to piece together the information for the paper.


