CONTRIBUTORS TO THIS VOLUME

JEAN-MARIE ANDRÉ
DAVID M. BISHOP
RAMON CARBÓ
LAP M. CHEUNG
A. GAVEZZOTTI
ODD GROPEN
Y. J. I'HAYA
PER-OLOV LÖWDIN
ISTVÁN MAYER
TETSUO MORIKAWA
M. SIMONETTA

ADVANCES IN
QUANTUM CHEMISTRY

EDITED BY

PER-OLOV LÖWDIN

DEPARTMENT OF QUANTUM CHEMISTRY
UPPSALA UNIVERSITY
UPPSALA, SWEDEN
AND
QUANTUM THEORY PROJECT
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA

VOLUME 12—1980

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

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

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 64–8029

ISBN 0–12–034812–8

PRINTED IN THE UNITED STATES OF AMERICA

80 81 82 83 9 8 7 6 5 4 3 2 1

LIST OF CONTRIBUTORS

Numbers in parentheses indicate the pages on which the authors' contributions begin.

JEAN-MARIE ANDRÉ, Laboratoire de Chimie Théorique Appliquée, Facultés Universitaires N.D. de la Paix, Namur, Belgium (65)

DAVID M. BISHOP, Department of Chemistry, University of Ottawa, Ottawa, Canada (1)

RAMON CARBÓ, Departamento de Química Orgànica, Secció de Química Quàntica, Instituto Química de Sarrià, Barcelona-17, Spain (159)

LAP M. CHEUNG, Department of Chemistry, University of Ottawa, Ottawa, Canada (1)

A. GAVEZZOTTI, Istituto di Chimica Fisica e Centro CNR, University of Milan, Milan, Italy (103)

ODD GROPEN, Institute of Mathematical and Physical Sciences, Department of Chemistry, University of Tromsø, N-9001 Tromsø, Norway (159)

Y. J. I'HAYA, Department of Materials Science, The University of Electro-Communications, Chofu, Tokyo, Japan (43)

PER-OLOV LÖWDIN, Department of Quantum Chemistry, Uppsala University, S-751 20 Uppsala, Sweden and Quantum Theory Project, University of Florida, Gainesville, Florida 32611 (263)

ISTVÁN MAYER, Central Research Institute for Chemistry of the Hungarian Academy of Sciences, H-1525 Budapest, Hungary (189)

TETSUO MORIKAWA, School of Pharmaceutical Science, Toho University, Funabashi, Chiba, Japan (43)

M. SIMONETTA, Istituto di Chimica Fisica e Centro CNR, University of Milan, Milan, Italy (103)

PREFACE

In investigating the highly different phenomena in nature, scientists have always tried to find some fundamental principles that can explain the variety from a basic unity. Today they have not only shown that all the various kinds of matter are built up from a rather limited number of atoms, but also that these atoms are constituted of a few basic elements of building blocks. It seems possible to understand the innermost structure of matter and its behavior in terms of a few elementary particles: electrons, protons, neutrons, photons, etc., and their interactions. Since these particles obey not the laws of classical physics but the rules of modern quantum theory of wave mechanics established in 1925, there has developed a new field of "quantum science" which deals with the explanation of nature on this ground.

Quantum chemistry deals particularly with the electronic structure of atoms, molecules, and crystalline matter and describes it in terms of electronic wave patterns. It uses physical and chemical insight, sophisticated mathematics, and high-speed computers to solve the wave equations and achieve its results. Its goals are great, but perhaps the new field can better boast of its conceptual framework than of its numerical accomplishments. It provides a unification of the natural sciences that was previously inconceivable, and the modern development of cellular biology shows that the life sciences are now, in turn, using the same basis. "Quantum biology" is a new field which describes the life processes and the functioning of the cell on a molecular and submolecular level.

Quantum chemistry is hence a rapidly developing field which falls between the historically established areas of mathematics, physics, chemistry, and biology. As a result there is a wide diversity of backgrounds among those interested in quantum chemistry. Since the results of the research are reported in periodicals of many different types, it has become increasingly difficult for both the expert and the nonexpert to follow the rapid development in this new borderline area.

The purpose of this serial publication is to try to present a survey of the current development of quantum chemistry as it is seen by a number of the internationally leading research workers in various countries. The authors have been invited to give their personal points of view of the subject freely and without severe space limitations. No attempts have been made to avoid overlap—on the contrary, it has seemed desirable to have certain important research areas reviewed from different points of view. The response from the authors has been so encouraging that a thirteenth volume is now being prepared.

The Editor would like to thank the authors for their contributions which give an interesting picture of the current status of selected parts of quantum chemistry. The topics in this volume cover studies of atomic and molecular structure, collision and surface phenomena, electronic and photoelectron spectra, as well as method developments including some extensions of the self-consistent field approach. Some of the articles emphasize studies in fundamental quantum theory and quantum statistics, and others applications to comparatively complicated systems.

It is our hope that the collection of surveys of various parts of quantum chemistry and its advances presented here will prove to be valuable and stimulating, not only to the active research workers but also to the scientists in neighboring fields of physics, chemistry, and biology who are turning to the elementary particles and their behavior to explain the details and innermost structure of their experimental phenomena.

PER-OLOV LÖWDIN

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 12

Accurate One- and Two-Electron Diatomic Molecular Calculations

DAVID M. BISHOP and LAP M. CHEUNG

Department of Chemistry
University of Ottawa
Ottawa, Canada

I. Introduction

The Schrödinger equation, within the Born–Oppenheimer approximation, can be solved accurately for one- and two-electron diatomic molecules. Therefore, since its inception these species have played an important role in the development of quantum chemistry. One has only to consider the papers by Burrau (1927), Heitler and London (1927), and James and Coolidge (1933) to recognize this fact. These papers formed the foundations for the work which is reviewed here. The Born–Oppenheimer electronic Schrödinger equation for the hydrogen molecular ion is separable in elliptic coordinates and Burrau successfully solved the separated equations. Heitler and London placed the intuitive idea of the valence bond on a firm quantum mechanical basis and this led to a qualitative

ISBN 0-12-034812-8

understanding of the hydrogen molecule. Six years later James and Coolidge, using the method Hylleraas had applied to the helium atom, calculated the dissociation energy of the hydrogen molecule and obtained agreement with experiment to within almost three figures. This paper, a prototype of many to follow, showed that the molecular Schrödinger equation was capable of giving quantitative results. It was central to the conversion of those remaining "doubters" to "believers" in the ways of quantum mechanics.

The average reader might well ask why, approximately 50 years later, papers are still being published in this area of quantum chemistry: Have the problems not been solved? There are many answers to this question and we will now give those which, to us, seem the most pertinent.

The first answer concerns the accuracy of calculated and experimentally determined observable properties. The calculation of observables has a pivotal position in quantum chemistry. As Sir Arthur Eddington, the famous British astronomer, is reported to have said: "You cannot believe in (astronomical) observations before they are confirmed by theory." This interplay between theory and experiment is a major component in the progress of science and as often as the theoretician must revise his results in the light of new experiments, so must the experimentalist do the reverse (e.g., the dissociation energy of H_2, the van der Waal's well in He_2). For one- and two-electron molecules the most important observables are those obtained by spectroscopic methods and the accuracy for transition frequencies is usually at least ± 0.1 cm^{-1}. A dazzling example is that of Wing *et al.* (1976) who obtained transition frequencies for HD^+ to within ± 0.002 cm^{-1}. To use units more familiar to theoreticians (the only common denominator in the choice of units by theoreticians and spectroscopists is an abhorrence of SI units; in Table I conversion factors are given), an energy of 0.1 cm^{-1} is equivalent to 0.5×10^{-6} hartrees or 0.3

TABLE I

UNITS[a]

	Quantum mechanical (a.u.)	Spectroscopic	International system
Distance	1 bohr (a_0)	0.529177 Å	0.529177×10^{-10} m
Energy	1 hartree	219,474.64 cm^{-1}	4.359828×10^{-18} J
Mass	1 amu (m_e)	5.48580×10^{-4} unified amu (u)	9.10953×10^{-31} kg

[a] Based on the fundamental constants given by Cohen and Taylor (1973).

cal/mole, which, for the molecules under discussion, means energies must be calculated to at least seven significant figures. This is an accuracy 3000 times greater than so-called chemical accuracy (1 kcal/mole). To achieve it clearly requires that the Schrödinger equation be solved almost exactly and that only the most refined wave functions will be relevant. This then establishes work of this nature in the top left-hand corner of Pople's (1965) two-dimensional chart of quantum chemistry, and progress as being vertical movement (from bottom to top) on the chart. Improvement in calculating, e.g., the rovibronic energies of H_2^+ and H_2 is therefore not some form of arcane sport but rather a reaction to and a reflection of what is happening experimentally. Lest it should be thought that it is *only* spectroscopic observables that require near-exact wave functions for their evaluation, it should be pointed out (see Section II,C) that only with the very best wave functions can one even obtain *two* significant figures for the static hyperpolarizability of H_2.

It is apropos to mention here some recent papers by Woolley (1978) and by Woolley and Sutcliffe (1977). They draw attention to the fact that the Born–Oppenheimer approximation (assumption of infinite nuclear masses) cannot be justified in any simple way in a completely nonclassical theory. This then leads to the abandoning of molecular structures, potential curves, and surfaces, and leaves us with what is known as nonadiabatic theory. It is stressed by Woolley that, more and more, experiments will be concerned in the future with practically isolated molecules (very dilute gases) and, with the concomitant higher resolution (increased accuracy), there will be the need for nonadiabatic calculations. Though Woolley's articles are polemical and sometimes relate only to the philosophical and conceptual nature of quantum mechanics, they are useful in that they draw the attention of a wider audience to the dangers of the Born–Oppenheimer approximation; theoreticians have been aware of this for some time. Essén (1977) has taken a view somewhat contrary to that of Woolley and, with a new derivation of the Born–Oppenheimer approximation, has attempted to remove at least some of the conceptual difficulties concerning molecular structure.

Some of the papers we will review have their raison d'être in the prediction of quantities which have yet to be measured experimentally. These results can be especially useful in astrophysics. As an example, the HeH^+ ion is found in the mass spectrum of discharges through He–H_2 mixtures but no optical spectrum of it has yet been reported. Since He and H_2 occur in high abundance in the universe it is reasonable to suppose that HeH^+ exists in some regions of stellar atmospheres or interstellar clouds. Its detection has been made easier by recent calculations of its infrared spectrum. Another astrophysical example is the accurate prediction of the most likely spectrum of H_2^+ to be observed by a Jupiter satellite. An important property which has yet to be measured directly is the quadrupole moment of the deuteron. It can, however, be found from combining the measurement of the electric-quadrupole-interaction constant for D_2 with the calculation of the electric field gradient. It is of profound interest that the most recent value so obtained does not agree with that predicted by various nuclear theories.

Very often approximate theories are tested on one- and two-electron diatomics before being applied to larger molecules. Accurate results, as reviewed here, thereby become the yardstick for the quality of the approximations.

Some of the work we will discuss is concerned with the interpretation of highly complex results, particularly for the most accurate calculations. It is essential, for the process of communication, that the results be presented in an easily understood form. This is contrary, perhaps, to Woolley's view but necessary if scientific links are to be maintained.

Important review articles have previously been written by Kołos and Wolniewicz (1963), and by Kołos (1968, 1970, 1977). The intention of this article is not only to cover the progress of the last few years but also to serve as a compendium of the best "numbers" available; to this extent we will sometimes reference work which, though older, has not yet been improved upon. Since it is impossible to cover everything in one review (e.g., an important area we have omitted is that of collision theory), our main emphasis will be (a) on results which are sufficiently accurate to be considered final or nearly final answers, and (b) areas where the greatest strides have recently been made (e.g., nonadiabatic calculations of H_2^+ and H_2).

The review is divided into sections which parallel the theory: Born–Oppenheimer calculations (electronic-nuclear coupling ignored), relativistic and radiative corrections (arising from use of the Schrödinger equation rather than the Dirac equation), adiabatic calculations (electronic-nuclear coupling partly taken into account), and nonadiabatic calculations (electronic-nuclear coupling completely accounted for); each section is further subdivided according to molecular species. This order, though the reverse of what would be logically deduced from the theory by starting from the complete Schrödinger equation, corresponds to that of increasing accuracy. The theory and corresponding equations will be introduced as necessary in the appropriate sections. Most of them were developed by Kołos and Wolniewicz, and their coupled names appear so frequently that we have taken the liberty of referring to them by initials only, i.e., KW.

II. Born–Oppenheimer Calculations

The separation of the coupling of electronic and nuclear motions in a molecule by use of the Born–Oppenheimer (1927) approximation is known well enough not to require explanation. All that need be said is that derivations differing from the original one have been given by Longuet-Higgins (1961) and Essén (1977). The approximation leads to two equations in place of the single initial Schrödinger equation: one electronic and the other nuclear.

The electronic equation has the form

$$H^0\Psi_B = E^0(R)\Psi_B \tag{1}$$

(the subscript B indicating a Born–Oppenheimer wave function), where, in atomic units, the Hamiltonian H^0 is

$$H^0 = \sum_i - \tfrac{1}{2} \nabla_i^2 + V, \tag{2}$$

∇_i^2 is the Laplacian operator for electron i (for later work, where it must be defined, this is relative to the *geometric* center of the nuclei), and V is the Coulombic potential between the particles. The eigenfunction Ψ_B is the Born–Oppenheimer electronic wave function and the eigenvalues $E^0(R)$ are a function of internuclear separation R. Equation (1) can be solved exactly for one-electron diatomic molecules when elliptic coordinates are introduced for the position of the electron. The wave function Ψ_B is usually expressed in terms of mathematical polynomials, e.g., Laguerre and Legendre, with certain fixed coefficients. However, there is no analytic solution for two-electron diatomic molecules; usually Eq. (1) is solved variationally with some highly flexible trial form for Ψ_B.

The nuclear equation, after separating out the rotational coordinates, is

$$[-(2\mu)^{-1}d^2/dR^2 + E^0(R) + (2\mu R^2)^{-1}J(J + 1)]S_{v,J}(R) = E_{v,J}S_{v,J}(R), \tag{3}$$

where μ is the nuclear reduced mass, and v and J are the vibrational and rotational quantum numbers, respectively. $S_{v,J}(R)$ is the vibrational wave function and $E_{v,J}$ is the total rovibronic energy (having electronic, vibrational, and rotational components).

Equation (3), a one-dimensional Schrödinger equation, is always solved numerically and a whole review article could be written on just this one topic. In brief, there are three main approaches: (a) the iterative method of Numerov and Cooley; (b) the finite-difference boundary-value method (a matrix method); and (c) the variation method using harmonic-oscillator or Morse-eigenfunction basis functions. Wicke and Harris (1976) have studied and compared these three techniques and found that they *all* give accurate results when properly applied. We have found Johnson's (1977) modified (renormalized) Numerov method to be both reliable and efficient.

In general, the calculations that we will now report have been carried out as the first step of more accurate (adiabatic) investigations.

A. H_2^+ and Its Isotopes

Since Eq. (1) is independent of nuclear mass, $E^0(R)$ and Ψ_B will be the same, for a given internuclear separation, for all the isotopes of H_2^+. Peek (1965) has tabulated $E^0(R)$ and the eigenparameters which define Ψ_B for the electronic ground state $1s\sigma_g$ and the excited state $2p\sigma_u$ of H_2^+ for $R = 0.1, (0.01), 1.0, (0.05), 30.0$ bohrs. Likewise Madsen and Peek (1971) have given eigenvalues and eigenparameters at $R = 1.0, (0.5), 9.0$ bohrs for all the other states which correlate at large R to a hydrogen atom with principal quantum number $(n) \leq 3$. The eigenvalues are accurate to 10–12 significant figures and these tables should replace the better known ones of Bates *et al.* (1953). Murai and Takatsu (1974) have tabulated for $R = 0.0, (0.1), 5.0$ bohrs electronic energies for those states which have a quantum number $(m) \leq 4$ for angular momentum along the molecular axis and united atom quantum numbers (n and l) such that $m \leq l \leq 9$ and $l + 1 \leq n \leq 10$. In addition, results are given at $R = 0.0, (0.5), 10.0, (1.0), 40.0$ bohrs for $n \leq 4$. Ten significant figures are reported.

However, a computer program written by Power (1973) which evaluates $E^0(R)$ for any specified state of any one-electron diatomic molecule is readily available. This program is so fast that Power has said: "It is cheaper to run it to get the eigenparameters written onto tape or cards than it is to have one of the printed tables key-punched and verified." We concur.

Values of the rovibronic energies $E_{v,J}$ for H_2^+, found by solving Eq. (3), have been determined for the electronic ground state ($1s\sigma_g$) (Beckel *et al.*, 1970) and for the excited states $2p\pi_u$ (Beckel *et al.*, 1973) and $3d\sigma_g$ (Shafi and Beckel, 1973). The Dunham spectroscopic constants (Y_{ij}) as well as ω_e and B_e were obtained both by fitting these rovibronic energies to Dunham expansions and by expanding the original $E^0(R)$ curves in a Dunham series. The latter are more precise.

Montgomery (1977) has evaluated a number of expectation values for the $1s\sigma_g$ and $2p\sigma_u$ states of H_2^+ as a function of R. He lists the total energy, kinetic (T) and potential (V) components, the quadrupole moment, $\langle x^2 \rangle$, and $\langle z^2 \rangle$, where z is the electronic coordinate along the molecular axis, for $R = 1.0, (1.0), 20.0$ bohrs. The values of T and V are in agreement with Ruedenberg's conception of binding in H_2^+. Other expectation values, $\langle z^2 \rangle$, $\langle r^2 \rangle$, $\langle r_a^n \rangle$, and $\langle R^n \rangle$, have been found by Bishop and Cheung (1978a) and averaged over the three lowest vibrational states for H_2^+, HD^+, and D_2^+. The purpose of this exercise was to find out how they changed when adiabatic and nonadiabatic wave functions were used (see Sections IV,A and V,A).

The static dipole polarizabilities ($\alpha_\parallel$, $\alpha_\perp$) as a function of R have been determined by both Montgomery (1978) and Bishop and Cheung (1978d) and the two sets of results are in near-exact agreement. Some selected values are given in Table II. They were calculated by using variational-perturbation (first-order) theory; the zero-order wave functions were found exactly. These are the first (except for $R = 2$ bohrs) accurate values of $\alpha_\parallel$ and $\alpha_\perp$ to be published. It is interesting to note that as R increases $\alpha_\perp$ goes through a maximum and then approaches the value for a H atom, whereas $\alpha_\parallel$ continues to increase, becoming extremely large. This can be understood in the following way. The ground-state electronic wave function of H_2^+ is symmetric about the xy plane (perpendicular to and through the midpoint of the nuclear axis) and under the perturbation xF_x (where F_x is the field strength), which is applied for finding $\alpha_\perp$, this symmetry is maintained in the sense that the electron density remains balanced on either side of this plane. Hence, at large R values the perturbed wave function is an equal combination of the perturbed wave functions for the species $H \cdots H^+$ *and* $H^+ \cdots H$. The polarizability $\alpha_\perp$ will therefore approach the H atom value. However, if the perturbation destroys this symmetry, as zF_z does, then the electron density changes from being equally distributed about the xy plane (and about the two protons) to being, at large R, practically wholly centered on one of the protons, i.e., either $H \cdots H^+$ *or* $H^+ \cdots H$. This change implies a massive transfer of charge under the perturbation and a very large value of $\alpha_\parallel$. We might call this the "egg-timer effect": an egg timer carefully balanced horizontally needs only a slight perturbation for all the sand to end up in one bulb. A detailed discussion of this effect has been given by Bishop and Cheung (1979g).

Montgomery and Rubenstein (1978), using time-dependent variational perturbation theory, have calculated the *dynamic* dipole polarizabilities of H_2^+, i.e., the values of α in an oscillating field. This calculation simply required the zero-order Hamiltonian in the first-order perturbation equation to be modified by the inclusion of the term $\pm h\nu$, where ν is the frequency of the oscillating electric field.

Higher static polarizabilities have been discussed by Bishop and Cheung (1979e). Their importance lies in such things as, e.g., the computation of potential energy curves for H_2–H_2^+ interactions, pressure-induced

TABLE II

POLARIZABILITIES OF H_2^+ AS A FUNCTION OF INTERNUCLEAR SEPARATION (R)[a]

R	$\alpha_\parallel$	$\alpha_\perp$	γ_{zzzz}	γ_{xxxx}	$B_{zz:zz}$	$B_{xx:xx}$	$C_{zz:zz}$	$C_{xx:xx}$
0.2	0.32318(0)	0.31675(0)	0.16871(1)	0.16684(1)	−0.52583(0)	−0.51571(0)	0.93005(−1)	0.92271(−1)
1.0	0.11196(1)	0.77623(0)	0.12964(2)	0.11471(2)	−0.39444(1)	−0.27159(1)	0.39445(0)	0.34140(0)
2.0	0.50777(1)	0.17577(1)	−0.40935(2)	0.73038(2)	−0.41869(2)	−0.13249(2)	0.19113(1)	0.12670(1)
3.0	0.19700(2)	0.29366(1)	−0.76142(4)	0.25711(3)	−0.32957(3)	−0.38618(2)	0.66921(1)	0.33714(1)
4.0	0.70047(2)	0.39644(1)	−0.25926(6)	0.59604(3)	−0.21268(4)	−0.78630(2)	0.17888(2)	0.71992(1)
7.0	0.24889(4)	0.47548(1)	−0.39708(10)	0.14333(4)	−0.27187(6)	−0.15981(3)	0.85762(2)	0.25405(2)
10.0	0.72422(5)	0.45388(1)	−0.46432(14)	0.13806(4)	−0.19074(8)	−0.13948(3)	0.16025(3)	0.43845(2)

[a] The number in parentheses is the power of ten by which the entry is to be multiplied. All values are given in atomic units.

spectra, collision-induced rotational Raman scattering. Hyperpolarizabilities (γ), quadrupole polarizabilities (B), and field-gradient quadrupole polarizabilities (C) were calculated for H_2^+ in what was a "dry run" for H_2. These constants can be defined in the following way. Let Greek subscripts α, β, . . . denote vector or tensor components equal to x, y, or z, a repeated subscript denote summation over all three Cartesian components, and F_α, $F_{\alpha\beta}$ be the electric field and field gradient, respectively, at the origin due to external charges. Then, the total dipole moment is

$$\begin{aligned}\mu_\alpha = \mu_\alpha^0 &+ \alpha_{\alpha\beta}F_\beta + \tfrac{1}{2}\beta_{\alpha\beta\gamma}F_\beta F_\gamma + \tfrac{1}{6}\gamma_{\alpha\beta\gamma\delta}F_\beta F_\gamma F_\delta \\ &+ \tfrac{1}{3}A_{\alpha:\beta\gamma}F_{\beta\gamma} + \tfrac{1}{3}B_{\alpha\beta:\gamma\delta}F_\beta F_{\gamma\delta} + \cdots \end{aligned} \tag{4}$$

and the total quadrupole moment is

$$\theta_{\alpha\beta} = \theta_{\alpha\beta}^0 + A_{\gamma:\alpha\beta}F_\gamma + \tfrac{1}{2}B_{\gamma\delta:\alpha\beta}F_\gamma F_\delta + C_{\alpha\beta:\gamma\delta}F_{\gamma\delta} + \cdots . \tag{5}$$

The values of γ, B, and C were found by applying single and double perturbation theory in order to obtain the required second- and fourth-order energies and the first- and second-order wave functions. The results approach the known He^+ values as $R \to 0$, and behave according to the pattern established by $\alpha_\parallel$ and $\alpha_\perp$ as $R \to \infty$. Some selected values are given in Table II. In the light of the many experimental advances that are being made to determine these properties (Bogaard and Orr, 1975), it seems to us that there will be many more calculations of this type in the future.

Finally, Bishop and Cheung (1978d) have published values of the ground-state moment functions S_k ($k = -3, -2, -1, 1, 2$) for $R = 0.2$, (0.2), 4.0, (1.0), 10.0 bohrs. These are required in order to obtain radiative corrections (see Section III,A) and are defined by

$$S_k = \sum_n{}' f_{0n}(E_n - E_0)^k, \tag{6}$$

where the subscript zero refers to the ground state, the subscript n refers to the nth excited state, E_0 and E_n are the energies, and f_{0n} are the oscillator strengths. They were found by using the following equations:

$$S_{-3} = \tfrac{2}{3}(\langle\Psi_\parallel'\Psi_\parallel'\rangle + 2\langle\Psi_\perp'\Psi_\perp'\rangle), \tag{7}$$

where Ψ' are first-order perturbation wave functions occurring in the dipole polarizability calculation,

$$S_{-2} = \tfrac{1}{3}(2\alpha_\perp + \alpha_\parallel), \tag{8}$$

$$S_{-1} = \tfrac{2}{3}\langle\Psi_0|r^2|\Psi_0\rangle, \tag{9}$$

$$S_1 = \tfrac{4}{3}T, \tag{10}$$

where T is the electronic kinetic energy, and

$$S_2 = \tfrac{8}{3}\pi\rho, \tag{11}$$

where ρ is the absolute value of the electronic density at the nucleus. The values obtained were often strikingly different from those of Bates (1972), who used Eq. (6) directly. These differences occurred because he failed to include sufficient contributions from the continua in his direct summations.

B. HeH^{2+}

The only other one-electron diatomic molecule that has been extensively studied is HeH^{2+}. Much of the interest is stimulated by the problem of scattering of He^{2+} by H. Murai and Takatsu (1976) have calculated the Born–Oppenheimer electronic energies at R = 0.0, (0.5), 10.0, (1.0), 40.0 bohrs for all those states which in the united atom limit have a principal quantum number of 5 or less—35 states in all. They list ten significant figures. Winter *et al.* (1977) have reported energies and derivatives of the eigenparameters (separation constants) with respect to R for the 20 lowest states at R = 0.1, (0.1), 20.0, (0.2), 80.0 bohrs. Though they give only six significant figures, in fact the energies were calculated to 1 part in 10^{13}. The 2pσ, 4fσ, and 4fπ states were found to be weakly bound. The derivatives are required for treating the He^{2+}–H scattering in the perturbed-stationary-state approximation.

C. H_2 and Its Isotopes

Progress in Born–Oppenheimer calculations for the hydrogen molecule has taken three forms: greater accuracy, larger internuclear separations, and more excited states. Much of it has been an extension of Kołos and Wolniewicz's (KW) pioneering work in the 1960s. Since for H_2, Eq. (1) is solved variationally, increased accuracy is achieved in two ways: (a) larger basis sets (more linear parameters), (b) more flexible basis functions (more nonlinear parameters).

Because of electron indistinguishability the electronic wave function is written as:

$$\Psi_B = \sum_{i=1}^{N} C_i[\Phi_i(1, 2) \pm \Phi_i(2, 1)], \tag{12}$$

where the + and − signs refer to singlet and triplet states, respectively. For Σ states the most flexible form of Φ_i used thus far is

$$\begin{aligned}\Phi_i = {} & \exp(-\alpha_1\xi_1 - \alpha_2\xi_2)[\exp(\beta_1\eta_1 + \beta_2\eta_2) \\ & + (-1)^{k_i+l_i+j}\exp(-\beta_1\eta_1 - \beta_2\eta_2)]\xi_1^{m_i}\xi_2^{n_i}\eta_1^{k_i}\eta_2^{l_i}\rho^{q_i},\end{aligned} \tag{13}$$

where $j = 0$ and 1 for states of g and u symmetry, respectively. In Eq. (13) ξ_i and η_i are the usual elliptical coordinates of the ith electron; $\rho = 2r_{12}/R$ (r_{12} is the interelectronic coordinate); m_i, n_i, k_i, l_i, and q_i are positive integers; α_1, α_2, β_1, and β_2 are four nonlinear parameters which may be optimized. When $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2 = 0$ Eq. (13) takes on the James–Coolidge (1933) form. These wave functions explicitly include the interelectronic coordinate and therefore take into account electron correlation. Wave functions which do this through configuration interaction will be mentioned only briefly when we discuss certain properties later on.

For Π states the most general wave function used thus far is

$$\Psi_B = \sum_{i=1}^{N} C_i[\Phi_i(1, 2)x_1 \pm \Phi_i(2, 1)x_2], \tag{14}$$

where x_i denotes the Cartesian coordinate of the ith electron perpendicular to the nuclear axis and

$$\Phi_i = \exp(-\alpha_1\xi_1 - \alpha_2\xi_2) \cosh(\beta_1\eta_1 + \beta_2\eta_2)\xi_1^{m_i}\xi_2^{n_i}\eta_1^{k_i}\eta_2^{l_i}\rho^{q_i}. \tag{15}$$

1. Ground-State Potential Curve

Near the equilibrium internuclear separation (1.4 bohrs) Bishop and Cheung (1978e) have shown that substantial improvement (a drop of 0.2 cm^{-1}) can be made to the earlier KW (1968b) curve which was based on a 100-term basis set. They achieved this by considering basis sets chosen in a systematic and orderly way; namely, if a basis set forming Ψ_B is denoted as $(a_0a_1a_2 \cdots /b)$, where $a_p = \max(m_i + k_i) = \max(n_i + l_i)$ for $q_i = p$, $b = \max(m_i + n_i + k_i + l_i)$ for any value of q_i (the power of ρ), and all functions within these constraints are included, then raising a_0, $a_1 \cdots$ in concert and separately raising b is an efficient test of energy convergence. It was found that using the basis set (6453/7) energy convergence is less than 0.02 cm^{-1} and $E^0(1.4) = -1.17447565$ hartrees, lower by 0.19 cm^{-1} than the previous "best" value. This basis set contained 249 terms and it was assumed that $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2 = 0$. This drop is due to the *choice* of basis functions as much as to the increase in the number of them: Even for the (5342/5) set (with 107 terms) E^0, at $R = 1.4$ bohrs, is 0.11 cm^{-1} lower than the "best" 100-term value of KW. Optimizing all four nonlinear parameters does not appear to be necessary when one is close to the equilibrium internuclear separation; Kołos and Rychlewski (1978) did optimize them and with a 130-term basis set found $E^0(1.4) = -1.17447540$ hartrees, which is only 4×10^{-8} hartrees lower than Bishop and Cheung's 130-term (6453/5) value. At least in this region of R, the optimization of all the nonlinear parameters does not seem to be worth the effort: It is much easier to expand, in a logical fashion, the basis set.

For larger R values KW (1975b) have improved the ground-state curve in the region $2.4 \leq R \leq 8.0$ bohrs by using the basis functions of Eq. (13) rather than James–Coolidge functions. Their best wave function contained 72 terms. Near 4.4 bohrs the energy decrease is quite substantial (about 5 cm^{-1}) and this goes a long way to remove the previous differences between the theoretical and experimental vibrational quanta; in fact the corrections are almost exactly those which were semiempirically predicted by LeRoy and Bernstein (1968). For $6 \leq R \leq 12$ bohrs a 60-term wave function with the basis functions of Eq. (13) but with $\alpha_1 = \alpha_2$ and $\beta_1 = -\beta_2$ has been used (KW, 1974). Improvement (over KW, 1965) of several cm^{-1} takes place at the short end of the range but it tapers off at the long range to only 0.1 cm^{-1}. These results have been used to investigate the convergence of the perturbation expansion of the polarization energy (Kołos, 1974).

A useful table of the best ground-state potential curve (including adiabatic and relativistic corrections) has been given by Bishop and Shih (1976, 1977).

2. *Excited-State Potential Curves*

The following singlet excited states have been investigated: the two lowest ${}^1\Sigma_g^+$ excited states labeled E,F and G,K (they both have double minima, the first symbol being the label for the inner minimum); the three lowest ${}^1\Sigma_u^+$ states, labeled B, B′, and B″,B̄ (the third having a double minima); the two lowest ${}^1\Pi_u$ states, labeled C and D; and the lowest ${}^1\Pi_g$ state labeled I. In most cases the work involves an improvement on that previously published by an increase in the number of basis functions, so that now, typically, 70–80-term wave functions are used.

For the E,F and G,K states Wolniewicz and Dressler (1977) used basis functions of the form:

$$\Phi_i = \exp(-\alpha_1\xi_1 - \alpha_2\xi_2 + \beta_1\eta_1 + \beta_2\eta_2)\xi_1^{m_i}\xi_2^{n_i}\eta_1^{k_i}\eta_2^{l_i}\rho^{q_i}, \tag{16}$$

which, when symmetrized, are identical to those of Eq. (13). Born–Oppenheimer energies are given for $1 \leq R \leq 15$ bohrs, but the authors caution that these cannot be considered final results and an improvement, in the vicinity of the minima of the GK curve, of several cm^{-1} is quite possible.

The B ${}^1\Sigma_u^+$ curve has been improved in the $1 \leq R \leq 12$ bohrs range by KW (1975a) with a drop of 5.2 cm^{-1} near the equilibrium separation. This was done by extending the basis set, Eq. (13), to 88 terms. It has also now been calculated by Kołos (1975b) to include the $12 \leq R \leq 20$ bohrs range; in addition, he investigated the perturbation expansion of the polarization energy in this range. The B′ curve has been computed to within a few cm^{-1} (Kołos, 1976b) for $1.2 \leq R \leq 20$ bohrs and does not show a previously predicted double minima. On the contrary, however, the next higher ${}^1\Sigma_u^+$ state (B″,B̄) does; this is in agreement with experiment. This curve too should be accurate to a few cm^{-1} (Kołos, 1976b).

For the C and D states (${}^1\Pi_u$), the wave function of Eqs. (14) and (15) was used by Kołos and Rychlewski (1976). Curves are given for $1 \leq R \leq 12$ bohrs and $1 \leq R \leq 25$ bohrs for the C and D states, respectively. An 80-term wave function was used for both states and the energies are anticipated to have converged to 1–2 cm^{-1}.

Only one ${}^1\Pi_g$ state (I) curve has been reported (Kołos and Rychlewski, 1977); it was based on a 75-term wave function with the form of Eq. (14).

In addition to the above singlet states, the following triplet-state curves are available: a ${}^3\Sigma_g^+$ (Kołos, 1975a), b ${}^3\Sigma_u^+$ (KW, 1974), c ${}^3\Pi_u$ and i ${}^3\Pi_g$ (Kołos and Rychlewski, 1977). In general, 60–80 basis functions were used with wave functions having the form of Eqs. (12) and (14) for the Σ and Π states, respectively, and taking the negative sign in these equations.

3. *Other Properties*

Most of the properties of the hydrogen molecule have been evaluated using the adiabatic approximation and these results will be given in Section IV,B. Spectroscopic properties (vibrational quanta, rotational constants, etc.) calculated using the Born–Oppenheimer approximation [i.e., solving Eq. (3) with the $E^0(R)$ of the previous two sections] usually agree poorly with the experimental values and to blame this on the breakdown of the approximation is understandable. In fact, one of the reasons for obtaining BO values is to enable us to make comparison with more accurate values.

Electronic ground-state vibrational quanta $\Delta G(v + \frac{1}{2})$ have been given by Wolniewicz (1966) and Poll and Karl (1966) for H_2, D_2, and HD. They differ by 1–2 cm^{-1} from the experimental values. Vibrational energies and rotational constants for the E,F and G,K states of H_2, D_2, and HD have been listed by Wolniewicz and Dressler (1977), in both cases for values of v up to the dissociation limit. A number of BO properties are known for the B′ and B″,B̄ states (Kołos, 1976b); these include dissociation energies, vibrational and rotational quanta for H_2, D_2, HD. The discrepancies with experiment are quite large (e.g., 40.3 cm^{-1} for D_0 for H_2 in the B′ state) but clearly mass dependent and hence compatible with the breakdown of the BO approximation. The same properties, with similar conclusions have been found for the C, D, I, c, and i states (Kołos and Rychlewski, 1976, 1977).

Expectation values $\langle \Psi_B | O | \Psi_B \rangle$ of the electronic coordinates (r_{12}^{-1}, r_{12},

r_{12}^2, r_a, r_a^{-1}, r_a^2, $r_a r_b$, $x_1 x_2$, etc.) and the quadrupole moment for H_2 with a range of R values have been given by KW (1965) for the electronic ground state.

Ford and Browne (1973) have applied the direct sum-over-states method to calculate the dynamic dipole polarizability and anisotropy of H_2, i.e.,

$$\alpha(\nu) = \sum_n{}' f_{0n}/[(E_n - E_0)^2 - h^2\nu^2]. \tag{17}$$

The energies E_n and oscillator strengths f_{0n} were found from wave functions, which were of the configuration-interaction type but modified so that the $S_0 = 2$ and $S_{-1} = \frac{2}{3}\langle\Psi|r^2|\Psi\rangle$ sum rules were satisfied. This technique had been originated by Dalgarno and Epstein (1969). Agreement with experiment was roughly 1%.

The static higher polarizabilities γ_{zzzz}, $B_{zz:zz}$, and $C_{zz:zz}$, as defined in Eqs. (4) and (5), have been computed by Bishop and Cheung (1979f) at $R = 1.4$ bohrs using the methods previously applied to H_2^+. They turned out to be extremely sensitive to the quality of the unperturbed and perturbed wave functions. In Table III we show a few cases where the basis for the *gerade* component of Ψ is being expressed by larger and larger expansions of the form used by Bishop and Cheung for the energy calculations. It is apparent that even with 194 basis functions convergence is only 1%. Static dipole polarizabilities (α) are also shown; KW (1967) have calculated values of α for a range of internuclear separations using smaller basis sets.

D. HeH^+

Green *et al.* (1974a,b) have carried out configuration-interaction calculations for the ten lowest singlet Σ states of HeH^+. They did so for a range of R values from 0 to 50 bohrs and at the same time calculated the dipole,

TABLE III

STATIC POLARIZABILITIES OF H_2[a,b]

Basis	N	α_{zz}	γ_{zzzz}	$B_{zz:zz}$	$C_{zz:zz}$
(5342/5)	107	6.3847	636	−87.2	5.83
(6453/5)	130	6.3849	661	−88.8	5.94
(6453/6)	194	6.3851	667	−89.2	5.97

[a] For R = 1.4 bohrs.

[b] Values are given in atomic units.

gradient, and radial-coupling matrix elements between the states for use in perturbed-stationary-state collision theory for light atoms.

Kołos (1976a) and Kołos and Peek (1976) have reported a much more accurate calculation for the ground state of HeH^+ for R values of 0.9 to 9.0 bohrs. They used a variational wave function of the form of Eq. (12) with 83 basis functions of the type given in Eq. (16); these were not symmetrized as HeH^+ is heteronuclear. The nonlinear parameters (α_1, α_2, β_1, β_2) were, however, optimized with a smaller (45-term) basis. The resulting potential energy curve has been used to compute the quasi-bound-state spectrum (Kołos and Peek, 1976) and the infrared spectrum (Dabrowski and Herzberg, 1977). The latter workers also, using the dipole moment data of Peyerimhoff (1965) and Michels (1966), found the intensities of the transitions.

In the paper by Kołos (1976a) results are also given for the two excited states of HeH^+: A $^1\Sigma^+$ ($4 \leq R \leq 10$) and a $^3\Sigma^+$ ($3 \leq R \leq 10$) and he has studied the long range part of these curves. He concluded that the $He^+ \ldots H$ interaction energy is mainly determined by the *second-* and higher-order polarization and exchange energies. This is in contradistinction to his findings for H_2, where (a) for two ground-state H atoms the main contributions are *first*-order exchange and second-order polarization energies, (b) for a 1s and 2pσ H atom they are *first*-order polarization and second-order exchange energies.

Bishop and Cheung (1979b) have considerably improved the Kołos–Peek curve. They used the same type of wave function but with 255 basis functions, i.e., a (5321/5) basis set. Some near-linear dependence in these functions was avoided by orthonormalizing with Löwdin's canonical method and excluding those orthonormal functions which correspond to the smallest eigenvalues of the overlap matrix. The lowering of the potential curve ranged from 2.3 to 5.4 cm^{-1} and the extrapolated value at $R = \infty$ (found by putting a zero charge on the proton at $R = 6.0$ bohrs) is only 0.3 cm^{-1} from the exact He value (Kołos' value differed by 2.4 cm^{-1}). The inclusion of basis functions with a r_{12}^3 component is undoubtedly one reason for the improvement.

E. He_2^{2+}

The second state of $^1\Sigma_g^+$ symmetry and the first state of $^1\Sigma_u^+$ symmetry of this molecular ion appear to be capable of supporting six and eight vibrational levels, respectively. This has been confirmed by Bishop and Cheung (1979c), who have calculated the BO curves with wave functions of the form of Eq. (12) and have included 153 basis functions, i.e., the basis set (552/5). In contrast to some earlier calculations these wave functions lead to the correct dissociation products ($He + He^{2+}$). Values of $E_{v,J}$

as well as the usual spectroscopic constants were tabulated. For $(^1\Sigma_g^+)^*$ $R_e = 3.58$ bohrs, $\omega_e = 620\ \text{cm}^{-1}$ and for $^1\Sigma_u^+$ $R_e = 3.26$ bohrs, $\omega_e = 658\ \text{cm}^{-1}$.

III. Relativistic and Radiative Corrections

The relativistic and radiative corrections have only been calculated for the electronic ground states of H_2^+ and H_2. The relativistic correction makes up for the fact that the Schrödinger equation is used in place of the Dirac equation and it has spin–orbit-interaction, mass-variable-effect, and Darwin-effect components. It is of the order of α^2, where α is the fine structure constant. The radiative correction (Lamb shift) is quantum electrodynamic in origin and is largely due to the self-energy of an electron in its field; it is of the order of α^3. Both effects are small compared with those due to the assumption of infinite nuclear mass (see Sections IV and V) but it is convenient to consider them now since they are often included in the adiabatic calculations. As an indication of their importance, for the first vibrational transition in H_2^+, the relativistic and radiative corrections are 0.036 and $-0.009\ \text{cm}^{-1}$, respectively.

A. H_2^+ and Its Isotopes

Luke *et al.* (1969) have shown, through reduction of the Dirac Hamiltonian to nonrelativistic form, that the relativistic correction to the BO potential curve is

$$\Delta E(R) = \langle \Psi_B | H_{\text{rel}} | \Psi_B \rangle, \tag{18}$$

where

$$\begin{aligned} H_{\text{rel}} &= \frac{-\alpha^2}{8}\left[\varepsilon - \frac{8\xi}{R(\xi^2 - \eta^2)}\right]^2 \\ &\quad - 4K_0\alpha^2 \frac{[(\xi^2 + \eta^2)(\xi^2 - 1)(\partial/\partial\xi) - 2\xi\eta(1 - \eta^2)(\partial/\partial\eta)]}{R^3(\xi^2 - \eta^2)^2} \end{aligned} \tag{19}$$

with

$$\varepsilon = -2[E^0(R) - 1/R] \tag{20}$$

and

$$K_0 = [(1 - \alpha^2\varepsilon/4)(\xi^2 - \eta^2) + 8\alpha^2\xi/R]^{-1}. \tag{21}$$

Bishop (1977) has evaluated the integral in Eq. (18) using a 96 × 96 Gauss–Legendre quadrature. The integral had to be subdivided, however, because of the singularity at $\xi = \eta = 1$. He listed values for $R = 0.2$, (0.2), 3.0, (0.5), 10.0 bohrs. These have been used in the H_2^+ calculations discussed in the next two sections.

Gonsalves and Moss (1979) have used a different method to obtain the relativistic-correction operator (H_{rel}) and have shown that all terms to order α^2 in Eq. (19) are included even when K_0 is put equal to unity [using the expression in Eq. (21) brings in some, but not all, terms of order α^4]. With $K_0 = 1$, the second term in Eq. (19) is simply $-\pi\alpha^2\rho(R)$, where $\rho(R)$ is the electronic density at the nucleus. Values of this second term are a mere 0.05% greater when $K_0 = 1$.

To second order in the electromagnetic interaction, the correction to the BO energy due to radiative effects is given in cm^{-1} by

$$\Delta E(R) = 0.2272[9.781 - \ln(k_0/\text{hartree})]\rho(R) \tag{22}$$

(Bethe and Salpeter, 1957; Gersten, 1969). In this equation $\ln(k_0/\text{hartree})$ is the Bethe logarithm and $\rho(R)$ is again the electronic density at the nucleus; both are functions of the internuclear separation. The Bethe logarithm is notoriously difficult to calculate accurately, but it may be approximated (Garcia, 1966) by

$$\ln k_0 = [d(\ln S_k)/dk]_{k=2}, \tag{23}$$

where S_k are the moment functions, Eqs. (7)–(11). The S_0 moment function is equal to the number of electrons (Reich–Thomas–Kuhn sum rule) and $S_{2.5} = \infty$. In order to determine this derivative S_k must be expressed in a functional form. Using the values given by Bishop and Cheung (1978d), S_k has been fitted to the following function of k:

$$S_k(R) = g_k(R) + \sum_{i=0}^{4} a_i(R)k^i + b(R)(2.5 - k)^{-q}, \tag{24}$$

where

$$g_k(R) = f_{01}(E_1 - E_0)^k + f_{02}(E_2 - E_0)^k \tag{25}$$

[these are the first two terms in the sum in Eq. (6)]. The parameter q was established by similar calculations on H and He [where the derivative in Eq. (23) at $k = 0$, 1, and 2 is known exactly] to be in the range $1.5 < q < 1.9$. With $q = 1.5$ Bishop and Cheung found $\ln(k_0/\text{hartree}) = 2.35$ for R near the equilibrium internuclear separation; for $q = 1.9$ it is 2.56. Corrections to the vibrational *spacings* with either of these values of the Bethe logarithm were practically the same. Details are given by Bishop and Cheung (1978d).

B. H_2 and Its Isotopes

The only relativistic corrections for H_2 are those evaluated over 15 years ago by Kołos and Wolniewicz (1964). They were computed using the two-electron relativistic Hamiltonian which can be derived from the

Breit equation in the Pauli approximation. The correction is then a sum of the following four terms:

$$\varepsilon_1 = \alpha^2\{\tfrac{1}{4}\langle\Delta_1\Psi_B|\Delta_2\Psi_B\rangle - \tfrac{1}{2}(E^0)^2 + E^0\langle\Psi_B V\Psi_B\rangle - \tfrac{1}{2}\langle\Psi_B V^2\Psi_B\rangle\}, \tag{26}$$

$$\varepsilon_2 = \tfrac{1}{2}\alpha^2\langle\Psi_B r_{12}^{-1}[\nabla_1\,\nabla_2 + (r_{12}^2)^{-1}\mathbf{r}_{12}(\mathbf{r}_{12}\,\boldsymbol{\nabla}_1)\,\boldsymbol{\nabla}_2]\Psi_B\rangle, \tag{27}$$

$$\varepsilon_4 = \pi\alpha^2\langle\Psi_B[2\delta(\mathbf{r}_{1a}) - \delta(\mathbf{r}_{12})]\Psi_B\rangle, \tag{28}$$

$$\varepsilon_5 = 2\pi\alpha^2\langle\Psi_B\;\delta(\mathbf{r}_{12})\Psi_B\rangle. \tag{29}$$

Values of these terms for $0 \leq R \leq 3.6$ bohrs were found by using a 54-term wave function based on James–Coolidge functions.

The correction to the BO potential curve due to radiative effects is (in cm^{-1})

$$\begin{aligned}\Delta E = 0.2195\{&\varepsilon_4[3.225 - 0.330\ln(k_0/\text{hartree})]\\ &+ \varepsilon_5[0.193 - 0.165\ln(k_0/\text{hartree})]\}\end{aligned} \tag{30}$$

(Garcia, 1966; Bethe and Salpeter, 1957). Using the method already described for H_2^+ Bishop and Cheung (1978c) have shown that, in the region of the equilibrium internuclear separation, the Bethe logarithm is close to 2.5 and practically constant. Using this value and the ε_4 and ε_5 values from KW (1964), ΔE was computed and the effect of the correction on the vibrational levels determined. The $v = 0 \rightarrow v = 1$ transition was changed by 0.022 cm^{-1} and this change declined to 0.011 cm^{-1} for the $v = 9 \rightarrow v = 10$ transition.

IV. Adiabatic Calculations

With respect to rigor, the adiabatic treatment lies between the Born–Oppenheimer and the exact nonadiabatic treatment. It retains the concept of a potential energy curve, while a nonadiabatic treatment does not, and is consequently much more "acceptable" to those who have classical ideas about molecular structure. The theory has recently been thoroughly reviewed by Kołos (1970) and we will therefore simply state the essential equations. Incidentally, Kołos's review contains an interesting discussion of the limiting values of the adiabatic corrections.

Ignoring, for the moment, relativistic and radiative corrections, the total molecular Hamiltonian, after separating out the center of mass coordinates, can be written for an n-electron diatomic molecule as

$$H = H^0 + H', \tag{31}$$

where H^0 is defined by Eq. (2) and

$$H' = -(2\mu)^{-1}\,\nabla_R^2 - (8\mu)^{-1}\left(\sum_{i=1}^{n}\nabla_i\right)^2 - (2\mu_a)^{-1}\,\nabla_R\,\sum_{i=1}^{n}\nabla_i. \tag{32}$$

In Eq. (32) ∇_R^2 is the Laplacian operator for one nucleus relative to the other, ∇_i is the gradient of the ith electron relative to the *geometric* center of the nuclei, μ is the reduced nuclear mass, and $\mu_a = m_a m_b/(m_a - m_b)$, where m_a and m_b are nuclear masses relative to the electron rest mass (m_e). Expressing the total wave function as an expansion of products of nuclear functions and BO electronic wave functions (Born, 1951) and introducing it into the Schrödinger equation with the above Hamiltonian leads to a set of coupled equations. These may be uncoupled by neglecting the so-called off-diagonal correction terms (this is the adiabatic approximation). This leads to the replacement in Eq. (3) of $E^0(R)$ by $U(R)$, where

$$U(R) = \langle\Psi_B|H|\Psi_B\rangle = E^0(R) + \langle\Psi_B|H'|\Psi_B\rangle = E^0(R) + \Delta E(R) \quad (33)$$

and Ψ_B is defined by Eq. (1). $\Delta E(R)$ is the adiabatic correction to the BO potential energy curve. If Ψ_B is symmetric (e.g., H_2^+, HD^+), the last (heteronuclear) term in Eq. (32) does not appear in $\Delta E(R)$.

We will now discuss a few general facts concerning the evaluation of $\Delta E(R)$. First, the correction is not particularly sensitive to the quality of Ψ_B. For example, Bishop and Cheung (1978e) have found that for H_2 near the equilibrium separation the difference in $\Delta E(R)$ using their 107-term (5342/5) wave function and using the KW (1964) 54-term wave function is only 0.02 cm^{-1}; whereas, the difference in $E^0(R)$ for the two cases is about 1.0 cm^{-1}. Consequently, much smaller basis sets can be used to find $\Delta E(R)$ than can be used to find $E^0(R)$.

Kari *et al.* (1973) have suggested an alternative adiabatic approximation, whereby the so-called mass-polarization term, $-(8\mu)^{-1}(\Sigma_{i=1}^{n} \nabla_i)^2$, is added to H^0 before Eq. (1) is solved. This produces a slightly different BO wave function (Ψ_B^*) and $\Delta E(R)$ becomes

$$\Delta E^*(R) = \langle\Psi_B^*|(H^0 + H')|\Psi_B^*\rangle - \langle\Psi_B|H^0|\Psi_B\rangle. \quad (34)$$

Both Pritchard and Wolniewicz (1976) and Bishop and Cheung (1978e) have tried this out for H_2 near the equilibrium separation and concluded that the alternative approximation drops the adiabatic correction by only 0.04 cm^{-1}.

Colbourn (1976) has followed up the proposal of Hunter *et al.* (1966) that in the electronic equation the electron reduced mass $\mu_e = m_e(m_a + m_b)/(m_e + m_a + m_b)$ should be used in place of m_e. She has computed the adiabatic curves for HD^+ with both approaches and finds that they differ by roughly 0.009 cm^{-1}. The differences in the vibrational quanta are negligible—one or two parts in 10^8.

As well as adding $\Delta E(R)$ to $E^0(R)$ one can also add the relativistic and radiative corrections. Use of such a curve in Eq. (3) will then produce adiabatic-relativistic-radiative rovibronic energy levels.

A. H_2^+ and Its Isotopes

Accurate adiabatic corrections for H_2^+ have been tabulated for $R = 0.2, (0.1), 10.0$ bohrs by Bishop and Wetmore (1973a). They were given in their component parts of $-\langle \Psi_B | \nabla_R^2 | \Psi_B \rangle / 2\mu$ and $-\langle \Psi_B | \nabla_1^2 | \Psi_B \rangle / 8\mu$ and the exact Ψ_B was used. With these values, vibrational quanta were calculated for H_2^+, HD^+, and D_2^+ (Bishop and Wetmore, 1973a,b, 1974). As an example of the effect of the adiabatic correction, a few values are shown in Table IV for H_2^+. It is seen that the spacings decrease by several tenths of a reciprocal centimeter.

Beckel *et al.* (1970) have also carried out adiabatic calculations for H_2^+. They used Kolos's adiabatic corrections, which were available for a limited number of R values. By expanding their curves in a Dunham series, they were able to find the spectroscopic constants Y_{ij}, ω_e, and B_e. The values of ω_e are 2323.55 and 2324.38 cm^{-1}, respectively, with and without the adiabatic correction and the values of B_e are 29.9511 and 29.9674 cm^{-1}, respectively.

Two tables of the rovibronic energies for H_2^+ and its isotopes have been published. One is by Hunter *et al.* (1974) for H_2^+, HD^+, and D_2^+ and adiabatic corrections are included. The other is by Bishop (1976a) for HD^+, HT^+, and DT^+ and both adiabatic and relativistic corrections are included.

The effect of the adiabatic corrections on the electronically and vibrationally averaged values of z^2, r^2, r_a, r_a^2, r_a^{-1}, R^{-2}, R^{-1}, R, R^2 has been investigated by Bishop and Cheung (1978a) for H_2^+, HD^+, and D_2^+. It was found that the expectation values of all the coordinates to positive powers increase (attributed to nuclear relaxation) and those to negative powers decrease. The magnitude of the changes was approximately inversely proportional to the reduced nuclear mass. Values for the ground state ($v = J = 0$) of H_2^+ are given in columns 2 and 3 of Table V.

Adiabatic curves have been published for only two of the excited states of H_2^+, namely $2p\pi_u$ and $3d\sigma_g$ (Bishop *et al.*, 1975). For the $2p\pi_u$

TABLE IV

Vibrational Quanta $\Delta G(v + \frac{1}{2})$ for H_2^+ [a]

v	Born–Oppenheimer	Adiabatic
0	2192.0	2191.3
1	2064.7	2064.1
2	1941.6	1941.1
3	1822.1	1821.7

[a] Values are given in cm^{-1}.

TABLE V

ADIABATIC (Δ_a) AND NONADIABATIC (Δ_n) CORRECTIONS TO THE BORN–OPPENHEIMER (BO) EXPECTATION VALUES FOR THE GROUND STATE OF H_2^+ [a]

	BO	Δ_a	Δ_n
$\langle z^2 \rangle$	1.1717	0.0004	−0.0004
$\langle r^2 \rangle$	2.4795	0.0007	0.0003
$\langle r_a \rangle$	1.6925	0.0003	0.0002
$\langle r_a^2 \rangle$	3.5572	0.0012	0.0003
$\langle r_a^{-1} \rangle$	0.84281	−0.00012	−0.00020
$\langle R^{-2} \rangle$	0.24405	−0.00013	0.00000
$\langle R^{-1} \rangle$	0.49084	−0.00013	0.00000
$\langle R \rangle$	2.0634	0.0005	0.0000
$\langle R^2 \rangle$	4.3111	0.0023	0.0000

[a] Values are given in atomic units.

state, corrections are given at $R = 2.0$, (0.5), 5.0, (0.1), 15.0, (0.5), 30.0 (5.0), 90.0 bohrs and for the $3d\sigma_g$ state, at $R = 1.0$, (0.5), 3.0, (0.1), 14.0, (0.5), 30.0, (5.0), 90.0 bohrs. Adiabatic spectroscopic constants and rovibronic energies ($0 \leq v \leq 11$ and $J = 1$ for $2p\pi_u$ and $0 \leq v \leq 30$ and $J = 0$ for $3d\sigma_g$) were determined. The $1s\sigma_g$–$2p\pi_u$ transitions are of importance to astronomers investigating the atmosphere of Jupiter. The Lyman-α satellite associated with the $1s\sigma_g$–$2p\pi_u$ transition, previously predicted to lie at 1240.5 Å, is shifted by adiabatic effects to 1241.3 Å. This change in wavelength makes it now unlikely that the 1240.5 Å absorption in the Zeta Tauri spectrum is due to the $1s\sigma_g$–$2p\pi_u$ satellite.

B. H_2 and Its Isotopes

1. The Electronic Ground State

More than 15 years later, the best values of the adiabatic corrections $\Delta E(R)$ for H_2 remain those of KW (1964). They were determined at $0.4 \leq R \leq 3.7$ bohrs using a 54-term electronic wave function. As mentioned previously, Bishop and Cheung (1978e) have confirmed these results in the region of the equilibrium internuclear separation and found that more accurate wave functions do not substantially change them (only by approximately 0.02 cm^{-1}). In the literature there is only one alternative, but less accurate, calculation: that by Ford *et al.* (1977). They used 14- and 24-term configuration-interaction type wave functions and *do* give some results at R values not considered by KW (i.e., 4, 6, 8 bohrs).

However, their accuracy is probably no better than 2 cm^{-1}. As well as listing the total correction at each R value, they also give its component parts and, in particular, ∇_R^2 is split into $(\partial^2/\partial R^2)$ and $(L_x^2 + L_y^2)$ and the corresponding integrals are calculated separately. It is the integral over $\partial^2/\partial R^2$ which is the most difficult to evaluate. It has also been calculated by a sum-over-states procedure (see Section II,C,3) by Ford (1974) at $R = 1.4$ bohrs.

Bishop and Shih (1976, 1977) have tabulated potential energy curves, for both H_2 and D_2, which incorporate adiabatic and relativistic corrections and are based on the best available data. With these curves they computed $E_{v,J}$ for $0 \leq v \leq 14$ and $0 \leq J \leq 4$ for H_2 and for $0 \leq v \leq 20$ and $0 \leq J \leq 4$ for D_2, as well as the vibrational-kinetic energies and rotational constants. Their results for $E_{v,J}$ differed, on average, by about 0.1 cm^{-1} from those of KW (1975b), since the latter had used outdated values for the nuclear masses.

Expectation values of various coordinates of H_2 (e.g., r^2, $3z^2 - r^2$, r_{12}^{-1}, z_1z_2) using adiabatic vibrational-rotational wave functions, have been calculated by Wolniewicz (1966). They were based on the unaveraged values of KW (1965) and, though somewhat outdated, are still the only ones in the literature. Because of its importance in astrophysics (it allows the determination of H_2 abundances in planetary atmospheres), Poll and Wolniewicz (1978) have recently investigated the quadrupole spectrum of H_2. They calculated the matrix elements of the quadrupole moment between different vibrational-rotational states to an accuracy of approximately 0.1%. The quadrupole moment as a function of R was also given (see Section V,B,1). Similarly, the hexadecapole moment of H_2 has been computed by Karl *et al.* (1974) in order to allow interpretation of the induced spectrum of H_2 at high pressure.

2. *The Electronic Excited States*

Adiabatic corrections for the lowest (double minima) excited state of symmetry $^1\Sigma_g^+$ (E,F) have been calculated (KW, 1969) for H_2 in the region $1.5 \leq R \leq 4.7$ bohrs. They show a sharp peak near the maximum of the BO potential curve. Wolniewicz and Dressler (1977) have computed the corrections for the second excited state of symmetry $^1\Sigma_g^+$ (G,K) for $1.5 \leq R \leq 4.5$ bohrs. They have also tabulated vibrational energies and rotational constants for H_2, HD, and D_2 for both excited states when the adiabatic corrections are added to the BO curve. Though the E,F state can support a number of vibrational levels, the G,K state for H_2 can support only one in each of the minima. The accuracy of their results is difficult to determine since the errors vary unsystematically due to the different interpolations of the adiabatic corrections. In comparison with experiment, the $v = 0$ levels of the E,F state lie about 6 cm^{-1} below the calculated ones for all three species and the near equality of these differences indicates a remaining error in the BO E,F potential curve. As v increases the differences become greater and show a systematic dependence on the reduced nuclear masses. This indicates either errors in the adiabatic corrections or the importance of the interaction with the G,K state (or with even higher $^1\Sigma_g^+$ states), i.e., nonadiabatic contributions, which have been neglected. The lowest inner-minima vibrational level of G,K is within 6 cm^{-1} of the observed level; this good agreement is due to a cancellation of errors (BO error and neglect of nonadiabatic effects).

The adiabatic corrections ($1.5 \leq R \leq 5.0$) for the lowest state of $^1\Sigma_u^+$ symmetry (B) have been worked out by KW (1966) using a 54-term wave function of the correct symmetry. Combining them with an accurate BO curve, KW (1975a) have determined the vibrational levels for H_2, HD and D_2 for this state. They differ from the experimental values by amounts, in places, as large as 12 cm^{-1} (for H_2) or 8 cm^{-1} (for D_2). The discrepancies in vibrational quanta are nonuniform and less than 1 cm^{-1}. The burden for these differences is, once more, placed upon nonadiabatic effects and an inaccurate BO curve. Ford *et al.* (1977) have used their configuration-interaction method to extend the range of the corrections from $R = 5.0$ to 10.0 bohrs, as well as repeating the calculations at KW's points. As was the case for the ground state, the agreement with KW when 26 configurations were used is within 2 cm^{-1}. There appears to be a minimum in the correction near 8 bohrs.

The same workers (Ford *et al.*, 1977) have been the only ones to estimate the corrections for the two lowest $^1\Pi_u$ states. They used a 6- and an 18-term wave function for the C and D states, respectively, and, combining the results with the BO curves of Kołos and Rychlewski (1976), determined the vibrational energies for H_2. They thereby removed most of the mass-dependent part of the previous discrepancies between theory and experiment. The remaining discrepancies are: mass-independent [1.5–3.2 cm^{-1} and 6 cm^{-1} for C and D, respectively (BO in nature)] and mass-dependent [1.0–3.5 cm^{-1} and 4 cm^{-1} for C and D, respectively (nonadiabatic in nature)].

Ford *et al.* (1975) have calculated approximate (10–20 cm^{-1} accuracy) adiabatic corrections for the second lowest $^1\Sigma_u^+$ state (B′) using configuration-interaction wave functions.

Adiabatic corrections have been estimated for only one triplet state: a $^3\Sigma_g^+$ (KW, 1968a). Five values were found in the range $1.5 \leq R \leq 4.0$ bohrs and were used, together with the BO curve, to determine the vibrational energy levels, vibrational quanta, and rotational constants for H_2 in this state. The energy levels differ from the experimental ones from 1 to 5 cm^{-1}.

C. HeH^+

The only accurate adiabatic calculation for HeH^+ was carried out by Bishop and Cheung (1979b). The corrections for $0.9 \leq R \leq 6.0$ bohrs were found by using the same 45-term wave function [with basis functions as in Eq. (16)] that Kołos (1976a) had used to optimize the nonlinear parameters in the calculation of $E^0(R)$. The corrections were added to the "improved" Kołos–Peek BO curve discussed in Section II,D and the rotational-vibrational levels were found for $0 \leq v \leq 2$, $0 \leq J \leq 2$. For these levels the accuracy was considered to be better than 1 cm^{-1} and, due to error cancellation, even better for the vibrational spacings. For higher v values the accuracy declines since the potential curve is defined by a very limited number of points. It was found that adiabatic effects decreased the dissociation energy (if the products are He and H^+) by 7.6 cm^{-1} to 14,873.6 cm^{-1} and that, e.g., $\Delta G(\frac{1}{2})$ and $\Delta G(\frac{3}{2})$ are decreased by 0.9 and 0.8 cm^{-1}, respectively. Expectation values of R^n, $n = -3, (1), 5$ for the three purely vibrational states, were also reported.

An approximate calculation of the corrections for HeH^+ has been given by Price (1978), who has used them to obtain the energies and widths of the quasi-bound (high J value) rotational–vibrational states of several HeH^+ isotopes. The differences between his corrections (adjusted to zero at $R = \infty$) and those of Bishop and Cheung are of the order of a reciprocal centimeter.

V. Nonadiabatic Calculations

The methods we now consider are the most accurate of all and are, *in principle,* exact; they are said to be nonadiabatic. To date, nonadiabatic calculations have only been carried out for two species (and their isotopes): H_2^+ and H_2.

Following KW (1963), two coordinates describing overall rotation can easily be removed from the complete nonrelativistic, nonradiative Schrödinger equation (from which the center of mass coordinates have already been separated out) by introduction of a particular rotating coordinate system. The total wave function is represented by

$$\Psi = \sum_{\Lambda=-K}^{K} \Omega^{K}_{M_K,\Lambda} u^{K}_{\Lambda}, \qquad (35)$$

where $\Omega^K_{M_K,\Lambda}$ are eigenfunctions of $\mathbf{K}^2$ (the operator for the square of the total angular momentum of the molecule in the space-fixed reference system), $K_{z'}$ (the component of $\mathbf{K}$ in the direction of the space-fixed axis z'), and L_z (the z component of the electronic angular momentum). They are, because of the axial symmetry of diatomic molecules, the same as the wave functions of a symmetric top.

The Schrödinger equation then becomes a set of $2K + 1$ coupled equations:

$$\begin{gathered} H_{\Lambda,\Lambda}u^K_\Lambda + H_{\Lambda,\Lambda-1}u^K_{\Lambda-1} + H_{\Lambda,\Lambda+1}u^K_{\Lambda+1} = E_K u^K_\Lambda, \\ \Lambda = -K, \quad -K+1, \quad \ldots, K, \end{gathered} \tag{36}$$

where

$$H_{\Lambda',\Lambda} = \langle \Omega^K_{M_K,\Lambda'} | H | \Omega^K_{M_K,\Lambda} \rangle . \tag{37}$$

The u^K_Λ depend only on the *relative* positions of the particles. Explicit expressions in elliptical coordinates for the terms entering $H_{\Lambda',\Lambda}$ are to be found in the appendix of KW (1963). Only for nonrotational Σ states ($K = 0$) is there a *single* equation, i.e.,

$$H_{0,0}u^0_0 = E_0 u^0_0, \tag{38}$$

which is often written simply as

$$H\Psi = E\Psi \tag{39}$$

with the understanding that H is to be read as $\langle \Omega^0_{0,0} | H | \Omega^0_{0,0} \rangle$.

Differences of approach in nonadiabatic calculations arise when it comes to solving Eqs. (36) or Eq. (39). Bishop and Cheung (1977b, 1978e) have solved Eq. (39) for H_2^+ and H_2 variationally using very large basis sets (e.g., 1070 terms). The basis functions explicitly involved both electronic coordinates and the internuclear separation (R). Because the generalized eigenvalue problem, which had to be solved, involved very large matrices (which were neither sparse nor had off-diagonal elements which were small compared with the diagonal ones) a special method had to be found for its solution. This was achieved through the group-coordinate relaxation method (Cheung and Bishop, 1977), where groups of coefficients in the eigenvectors were varied simultaneously. Iteration was continued until convergence had been achieved.

Wolniewicz and co-workers have used first-order variational perturbation theory to solve a number of nonadiabatic problems: the vibrational energy levels of the electronic ground states of H_2^+ and HD^+; the rotational energies of H_2, HD, and D_2; the dipole transitions in HD.

Much more approximate methods are those involving (a) multichannel quantum defect theory, (b) Rayleigh–Schrödinger perturbation theory coupled with the Unsöld approximation, and (c) semiempirical Hamiltonians.

Another approach to the nonadiabatic problem is that based on the Born (1951) or Born and Huang (1956) expansion. It is assumed that the Born–Oppenheimer electronic equation has been solved to yield a spectrum of electronic wave functions $\Psi_n(r, R)$, where r and R stand for electronic and nuclear coordinates. The full Schrödinger equation is then solved on the assumption that the required solution $\Psi(r, R)$ can be written as

$$\Psi(r, R) = \sum_n \psi_n(r, R)\chi_n(R). \tag{40}$$

Woolley and Sutcliffe (1977) have criticized this procedure, but do admit that it is justified for diatomic molecules, if not for polyatomics. Nonetheless, there are, *a priori,* no grounds for expecting the expansion in Eq. (40) to converge quickly; there are expansions *like* Eq. (40) (see Section V,A,1) which do, but the ψ_n are not then the BO electronic wave functions. Eq. (40), with two terms, has been used by Orlikowski and Wolniewicz (1974) to calculate the nonadiabatic vibrational levels of H_2 and D_2 in their electronic ground state and by Dressler *et al.* (1979), again with two terms, to calculate the nonadiabatic coupling between the E,F and G,K ${}^1\Sigma_g^+$ states of H_2, HD, and D_2.

Very recently it has been suggested that the Generator Coordinator Method (GCM), which has been developed by physicists for constructing nuclear wave functions (Griffin and Wheeler, 1957), could be used in nonadiabatic molecular calculations (Lathouwers *et al.,* 1977; Lathouwers, 1978). This suggestion has yet to be put into practice but we will sketch the basic principles. The complete wave function is written as an integral over generator coordinates (α) and is reminiscent of the integral-transform functions once used by Bishop and Somorjai (1970) for atomic wave functions:

$$\Psi_n(r, R) = \int f(\alpha)\psi_n(r|\alpha)\phi(R|\alpha)\, d\alpha. \tag{41}$$

In this equation r and R represent electronic and nuclear coordinates, respectively; $f(\alpha)$ are weight functions, where a nuclear configuration is denoted by a set of α values; $\psi_n(r|\alpha)$ are eigenfunctions of $H_0(\alpha)$; and $\phi(R|\alpha)$ is a disposable function sharply peaked around α. In essence the finite sum in Eq. (40) has been replaced by an infinite sum of continuously labeled configurations which are weighted by $f(\alpha)$. These latter functions may be determined from integral equations derived from the variational principle. It can be seen that if $\phi(R|\alpha)$ is chosen to be $\delta(R - \alpha)$ then

$\Psi_n(r, R)$ simply becomes $f(R)\psi_n(r|R)$, the BO solution. The GCM has been used as an alternative to the configuration-interaction method in molecular calculations by Laskowski *et al.* (1978); it will be interesting to see how it works out in actual practice for nonadiabatic calculations.

Finally, many of the nonadiabatic wave functions are highly complex and intrinsically avoid the well-loved concepts of structure and potential curves. This is part of the price we must pay for accuracy. However, some attempts have been made to make them more "accessible" and these will be discussed.

A. H_2^+ and Its Isotopes

1. *Exact Calculations*

Following the preliminary investigation by Bishop (1974), Bishop and Cheung (1977b) have calculated certain low-lying infrared transition frequencies for H_2^+, D_2^+, HD^+, HT^+, and DT^+. This work was stimulated by the highly accurate (± 0.002 cm^{-1}) measurement of the vibrational-rotational spectrum of HD^+ by means of an ion-beam laser-resonance method (Wing *et al.*, 1976). Eq. (39) was solved variationally by using as a trial wave function

$$\Psi = \sum_{i=0}^{i_m} \sum_{j=0,2}^{j_m} \sum_{k=0}^{k_m} C_{ijk}\phi_{ijk}(\xi, \eta, R) + \sum_{i=0}^{i'_m} \sum_{j=1,3}^{j'_m} \sum_{k=0}^{k'_m} C_{ijk}\phi_{ijk}(\xi, \eta, R). \tag{42}$$

In this equation the second triple summation only occurs for the heteronuclear molecules. The C_{ijk} are linear variational parameters found by solving the usual secular equation. The basis functions, where ξ and η are electronic elliptical coordinates and R is the internuclear separation, were expressed as

$$\phi_{ijk}(\xi, \eta, R) = \exp(-\alpha\xi)\cosh(\beta\eta)\xi^i\eta^j R^{-3/2}\exp[\tfrac{1}{2}(-x^2)]H_k(x), \tag{43}$$

where $x = \gamma(R - \delta)$; $H_k(x)$ are Hermite polynomials; α, β, γ, and δ are adjustable parameters chosen to minimize the lowest energy level; and i, j, and k are integers. In order to increase the basis-set size and study the resulting energy-convergence logically, the values of i_m, j_m, and k_m were raised in unison, but with the proviso that basis functions with $(i + j + k) > 15$ were excluded. This raising was continued until the energy of the vibrational level under consideration changed by less than 2×10^{-9} hartrees. For the heteronuclear molecules the required number of terms in the second triple summation in Eq. (42) was not very great and i'_m, j'_m, k'_m were raised together, but with no restriction placed on

$(i + j + k)$. As an example of basis-set size, for the third lowest nonrotational energy level of the heteronuclear molecules, the values $i_m = 9$, $j_m = 14$, $k_m = 15$, $i'_m = 4$, $j'_m = 5$, $k'_m = 6$ were used—515 functions in all.

Values of the nonadiabatic as well as the adiabatic energies (the difference is labeled Δ_2) and the relativistic-radiative corrections (Δ_1) were calculated for the low-lying states ($0 \leq v \leq 2$, $0 \leq J \leq 2$; v is the vibrational and J the rotational quantum number) of all five ions. The change in the nonadiabatic correction due to change of rotational quantum number (the nonadiabatic calculations were restricted to rotationless states) was determined by an approximate formula (Colbourn and Bunker, 1976); the effect (Δ_3) is so small that this approximation is not at all serious. Sample results for the ensuing transitions, labeled $(v, J) - (v', J')$, are given in Table VI for HD$^+$. The final frequencies, with estimated accuracy to ± 0.002 cm^{-1}, are in more than satisfactory agreement with the experimental values (final column). It is interesting to speculate that if the theoretical and experimental values were determined to an accuracy greater by one order of magnitude, the accuracy of the masses of the nuclei assumed in the theory would determine the degree of agreement. The dissociation energy of H_2^+ can be found from the results presented to be 21,379.4 cm^{-1}.

An interesting alternative calculation (though not as accurate as the previous one) of the nonadiabatic corrections for the four lowest vibrational levels of H_2^+ and D_2^+ has been performed by Wolniewicz and Poll (1978). They write the total wave function as

$$\Psi = \Psi_{ad} + \Psi', \tag{44}$$

where Ψ_{ad} is the adiabatic wave function, and then solve variationally the first-order perturbation equation

$$(H - E_{ad})\Psi' = -H'\Psi_{ad}, \tag{45}$$

TABLE VI

TRANSITION FREQUENCIES AND CORRECTIONS FOR HD$^+$[a]

Transition	$\nu_{\text{adiabatic}}$	Δ_1	Δ_2	Δ_3	ν_{total}	ν_{expt}
(1,0)–(0,0)	1913.1243	0.0240	−0.1522	0	1912.9961	—
(2,0)–(1,0)	1816.9800	0.0216	−0.1369	0	1816.8647	—
(1,0)–(0,1)	1869.2621	0.0226	−0.1522	0.0026	1869.1351	1869.134
(1,1)–(0,2)	1823.6600	0.0212	−0.1522	0.0052	1823.5342	1823.533
(2,1)–(1,0)	1856.8983	0.0227	−0.1369	−0.0026	1856.7815	1856.778

[a] Values are given in cm^{-1}.

where H' is defined in Eq. (32). The nonadiabatic correction is then given by the second-order energy

$$E'' = \langle \Psi' | H' | \Psi_{\mathrm{ad}} \rangle . \tag{46}$$

Ψ_{ad} was determined in the standard way and Ψ' was expanded as

$$\Psi' = \sum_{nkl} C_{nkl} h_n(R) g_k(\xi) P_l(\eta), \tag{47}$$

where

$$g_i = \exp(-\gamma\xi)(\xi + 1)^{\alpha}[(\xi - 1)/(\xi + 1)]^{i},$$

$$h_n = R^{-3/2} \exp[-\beta(R - R_0)] H_n[\beta(R - R_0)],$$

P_l and H_n are Legendre and Hermite polynomials, respectively, and C_{nkl}, γ, α, β, and R_0 are all variational parameters. Nine Hermite polynomials and 22 $g_k P_l$ combinations were included in the sum in Eq. (47)—198 terms in all. As an example of their results, values for the energies of the lowest three nonrotational states of H_2^+ and HD^+ are shown in Table VII and compared with Bishop and Cheung's (1977b) completely variational results. A reasonable explanation for the latter's better (i.e., lower energy) values is that their calculation involved roughly 50–100% more basis functions.

Nonadiabatic corrections for $\langle z^2 \rangle$, $\langle r^2 \rangle$, $\langle r_{\mathrm{a}} \rangle$, $\langle R^{-2} \rangle$, etc. have been calculated using the best variational wave functions for H_2^+, HD^+, and D_2^+ (Bishop and Cheung, 1978a; see Table V). They are negligible for expectation values involving R (Karl and Poll, 1967, have previously noted that the correction for such values is zero to first-order perturbation theory) and of the same order of magnitude as the adiabatic corrections for the others (making it futile to calculate one correction without the other).

TABLE VII

NONADIABATIC ENERGIES OF H_2^+ AND HD^{+a}

	H_2^+		HD^+	
v	W and P	B and C	W and P	B and C
0	−0.597139050	−0.597139062	−0.597897961	−0.597897967
1	−0.587155616	−0.587155676	−0.589181818	−0.589181825
2	−0.577751795	−0.577751866	−0.580903575	−0.580903679

[a] Values are given in atomic units.

2. *Interpretation of Exact Results*

Here, we would like to briefly discuss two possible ways which have been used to interpret the exact nonadiabatic H_2^+ calculation: One turned out to be a failure and the other a success.

The failure was the belief (Bishop and Hunter, 1975) that if the nonadiabatic wave function for each vibrational state was expressed as

$$\Psi(r, R) = \phi(r, R)f(R), \tag{48}$$

where $\phi(r, R)$, the conditional probability amplitude, is defined such that

$$\langle \phi(r, R)|\phi(r, R)\rangle_r = 1 \tag{49}$$

(integrating over electronic coordinates), then a nonadiabatic pseudopotential

$$U(R) = \langle \phi(r, R)H\phi(r, R)\rangle_r \tag{50}$$

would be both relevant and only slightly different (due to nonadiabatic effects) from one state to the next. It was hoped that $U(R)$, as well as being only slightly different from the adiabatic potential curve, would also be variational so that a trial $\phi(r, R)$ could be optimized and found without recourse to a full nonadiabatic calculation. Unfortunately, Bishop and Cheung (1977a) have shown through a counterexample (the generalized coupled-harmonic-oscillator problem) that $U(R)$ is not variational, and Czub and Wolniewicz (1978) have shown that the $f(R)$ are nodeless and that consequently there are potential barriers in $U(R)$ at the positions of the nodes of the adiabatic vibrational wave function for states $v \neq 0$. The fact that $f(R)$ is nodeless for the coupled-harmonic-oscillator case is verified by inspection of the equations; for H_2^+ a phase change (creating a node) was inserted in the values derived for $f_1(R)$ by Bishop and Hunter at the position where more recent and closer inspection reveals there is only a minimum and not a zero.

More successful and fruitful was a natural-orbital analysis of the H_2^+ nonadiabatic wave function (Bishop and Cheung, 1979a). The theory they developed was independently and almost simultaneously discussed by Goscinski and Palma (1979), though these workers gave no applications. Basically, the idea was to expand the total wave function in natural orbitals (NOs) separately involving electronic and nuclear coordinates, just as two-electron correlated wave functions can be expressed in terms of natural orbitals which involve a single electron. The equation which expresses this is

$$\Psi(r, R) = \sum_{i=1}^{N} n_i^{1/2}\Phi_i(r)X_i(R), \tag{51}$$

TABLE VIII

THE EFFECTS OF TRUNCATION OF THE NATURAL ORBITAL EXPANSION OF THE GROUND-STATE NONADIABATIC WAVE FUNCTION OF H_2^+ ON E, S, $\langle z^2\rangle$, AND $\langle R^2\rangle$[a]

Number of terms	ΔE	ΔS	$\Delta\langle z^2\rangle$	$\Delta\langle R^2\rangle$
1	−0.0053963326	0.0021411826	−0.012252	0.000290
2	−0.0000886488	0.0000167198	−0.000094	0.000006
3	−0.0000014136	0.0000001888	−0.000001	0.000000
4	−0.0000000217	0.0000000023	0.000000	0.000000
5	−0.0000000004	0.0000000000	0.000000	0.000000

[a] The values given are the differences between the exact values and those obtained by using one to five terms in the natural orbital expansion (with renormalization). The exact values are: energy, $E = -0.5971390625$; overlap, $S = 1$; $\langle z^2\rangle = 1.171737$; and $\langle R^2\rangle = 4.313278$. All values are given in atomic units.

where the n_i can be considered occupation numbers, and $\Phi_i(r)$ and $X_i(R)$ are called electronic and nuclear natural orbitals, respectively. The $X_i(R)$ were determined such that the first-order density matrix is diagonal:

$$\rho(R, R') = \langle\Psi(r, R)|\Psi(r, R')\rangle_r = \sum_{i=1}^{N} n_i X_i(R) X_i(R') \tag{52}$$

and the $\Phi_i(R)$ were found from

$$\Phi_i(r) = n_i^{-1/2}\langle\Psi(r, R)|X_i(R)\rangle_R. \tag{53}$$

Eq. (51) bears a resemblance to the Born expansion, but this resemblance is only superficial; plots of the $\Phi_i(r)$ show that though $\Phi_1(r)$ is similar to the lowest BO electronic wave function, the subsequent ones are not. Furthermore, natural-orbital expansions are known to converge very rapidly (there is no such proof for the Born expansion, though it has often been assumed) and this is illustrated by the results in Table VIII. It is apparent that the original 308-term wave function can be satisfactorily reproduced by just five NOs of each type. It might even be possible to determine *a priori* the NOs of H_2^+, and hence the nonadiabatic wave function, by using a method similar to that which Kutzelnigg (1963) used to determine the electronic NOs of He.

3. *Approximate Hamiltonians*

An approximate Hamiltonian to replace the one in Eq. (3) *and* which effectively accounts for nonadiabatic effects has been intuitively deduced by Bishop and Shih (1976, 1977). It was later put on a more rigorous foundation by Bunker and co-workers (1977; Bunker and Moss, 1977). For nonrotational states it takes the form

$$H_{\text{eff}} = -(2\mu)^{-1}\, d^2/dR^2 + U(R) + \alpha(R)\{U(R) - E_{v,0}\}, \tag{54}$$

where $U(R)$ is the adiabatic potential,

$$\alpha(R) = -k(2\mu)^{-1}\langle \psi^0 | d^2/dR^2 | \psi^0 \rangle, \tag{55}$$

k is an adjustable parameter, ψ^0 is the BO electronic wave function and $E_{v,0}$ is the adiabatic vibronic energy. The value of k was chosen so that the lowest energy level coincided with the true nonadiabatic value. It seems to work quite well, since the lowest vibrational spacings for H_2^+ and D_2^+ agree to six figures with the exact ones.

These results, together with some experimental data, have been used by Kuriyan and Pritchard (1977) to model a correction term ΔV such that when it is added to $U(R)$ and Eq. (3) is solved there is agreement with the known nonadiabatic rovibronic energies. By doing this they have been able to tabulate "nonadiabatic" energies for a broad range of rotational and vibrational quantum numbers for H_2^+ and D_2^+. For the $J > 0$ levels these are probably the most reliable values available.

B. H_2 and Its Isotopes

1. The Electronic Ground State

There has only been one entirely successful nonapproximate nonadiabatic calculation of the lowest energy level of H_2 (Bishop and Cheung, 1978e). The method used was completely variational and based on a wave function expanded as

$$\Psi = \sum_{i=1}^{N}\sum_{j=0}^{M} C_{ij}[\Phi_i(1, 2) + \Phi_i(2, 1)]\chi_j, \tag{56}$$

where the C_{ij} are linear variational parameters, the basis functions Φ_i take the form of Eq. (13) with $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2 = 0$, and the basis functions χ_j are written as

$$\chi_j = R^{-3}\exp(-x^2/2)H_j(x), \tag{57}$$

where $x = \gamma(R - \delta)$ and $H_j(x)$ are the usual Hermite polynomials. The nonlinear parameters α_1, γ, and δ were respectively fixed at 1.117 (the optimized value in adiabatic calculations), 4.3 (determined from the force constant of H_2), and 1.4 (the equilibrium internuclear separation). When there are a large number of linear parameters, the results are not overly sensitive to the choice of γ and δ.

By using 107 electronic basis functions ($N = 107$), i.e., the basis set

(5342/5), see Section II,C,1, and ten radial functions ($M = 9$), the lowest energy was found to be -1.16402413 hartrees. One less radial function gave -1.16402411 hartrees, suggesting convergence with respect to M of 0.01 cm^{-1}. Since the BO calculations at $R = 1.4$ bohrs showed that a change of electronic basis from (5342/5) to (6453/7) lowered the energy (E^0) by 0.08 cm^{-1}, the final nonadiabatic energy should also be lowered by this amount. Combining this result with the relativistic and radiative corrections (Bishop and Cheung, 1978c) the final dissociation energy of H_2 is determined to be 36,117.96 cm^{-1} (Table IX), with a possible error due to incomplete convergence of 0.1 cm^{-1}.

This value may be compared with a semiempirical value (36,117.8 ± 0.4 cm^{-1}) found by combining the theoretical dissociation energy of H_2^+ with the experimental ionization potential of H_2 (Herzberg and Jungen, 1972) and the energy of the H atom. Stwalley (1970) has analyzed the experimental vibrational levels of the B $^1\Sigma_u^+$ state and concluded that the ground-state dissociation energy is 36,118.6 ± 0.5 cm^{-1}. Finally, from ultraviolet absorption studies on H_2, Herzberg (1970) finds a value lying between 36,118.3 and 36,116.3 cm^{-1}. Looking at Table IX, it is clear that though there is agreement between the theoretical value and the semiempirical value, there is not with Stwalley's value. The reason for this is not known, though there is a similar disagreement for the dissociation energy of D_2 (LeRoy and Barwell, 1975).

Comparison with adiabatic calculations showed that the lowest level in H_2 is lowered by nonadiabatic effects by 0.42 cm^{-1}. Bunker (1972), by considering the experimental vibrational spacings and the theoretical adiabatic values and assuming that the differences are nonadiabatic in origin, has predicted a value of 0.43 cm^{-1}. This agreement between the two values is very satisfying.

TABLE IX

DISSOCIATION ENERGY OF H_2[a]

Theoretical		Experimental	
Nonadiabatic	36,118.60	Semiempirical (from $D_0(H_2^+)$ and IP)	36,117.8 ± 0.4
Basis-set correction	0.08	Stwalley analysis	36,118.6 ± 0.5
Relativistic correction	−0.54	Herzberg's absorption limits	36,118.3 to 36,116.3
Radiative correction	−0.18		
	36,117.96		

[a] Values are given in cm^{-1}.

Several ground-state nonadiabatic expectation values of H_2 have been evaluated (Bishop and Cheung, 1978a) and comparison with the adiabatic values leads to conclusions similar to those made for the H_2^+ calculations. The ground-state rotational constant (B_0) was found to be 59.337 cm^{-1}, in agreement with the experimental value, 59.336 ± 0.001 cm^{-1}, of Fink *et al.* (1965).

Poll and Wolniewicz (1978) have *estimated* the nonadiabatic contributions to the H_2 quadrupole matrix elements of the form $\langle vJ|Q|v'J'\rangle$ for $v' = J = J' = 0$ and $0 \leq v \leq 5$. They did so by using perturbation theory and the Unsöld approximation, and this led to the expression

$$\Delta Q_v = (2\mu\Delta E)^{-1}\langle \chi_0|d^2Q_e/dR^2|\chi_v\rangle \tag{58}$$

for the contributions. In Eq. (58) ΔE is the Unsöld average excitation energy (approximately 0.61 hartrees), Q_e is the electronic part of the quadrupole moment operator, and χ_v are the adiabatic vibrational wave functions. The values of ΔQ_v were determined to be 0.1–0.2% of the adiabatic totals. For $\langle 00|Q|00\rangle$ they obtained (including the correction) a value of 0.9688 a.u.; using the expectation values of Bishop and Cheung (1978a) one gets 0.9696 a.u. We suspect that most of the difference between these values lies in the adiabatic part. This was calculated by Poll and Wolniewicz on the basis of vibrationally averaging the expectation value $\langle \Psi_B|Q|\Psi_B\rangle$ with 72- and 80-term BO wave functions. This may be inaccurate since Bishop and Cheung (1979f) have shown that the electronic part of this integral at $R = 1.4$ bohrs changes from 0.5225 to 0.5232 to 0.5234 a.u. as one uses 54-, 78-, and 107-term BO wave functions.

A property of fundamental interest to both physicists and chemists alike is the nuclear quadrupole moment of the deuteron (Q). Unfortunately, nuclear-model calculations (McGurk, 1977) of its value differ by a disturbing 2% from those obtained from combining the electric field gradient q in HD or D_2 with the appropriate experimental electric-quadrupole-interaction constant eqQ/h (Reid and Vaida, 1975). Since the evaluation of q had been made using the adiabatic approximation it appeared possible that nonadiabatic effects were responsible for the lack of agreement between the two treatments; Bishop and Cheung (1979d) have shown that this is *not* the case, however, and have confirmed Reid and Vaida's value. They calculated the electric field gradient at the deuteron in the lowest rotationless state of D_2 using a 540-term nonadiabatic wave function of the form of Eq. (56) (with 54 electronic and 10 nuclear basis functions) and found $\langle q'\rangle$ to be 0.16812 a.u. The corresponding adiabatic value is 0.16795, and q is related to $\langle q'\rangle$ by

$$q = 2e\langle q'\rangle, \tag{59}$$

where

$$\langle q' \rangle = \langle \Psi | R^{-3} - \tfrac{1}{2} \sum_{i=1}^{2} (3z_i^2 - r_i^2)/r_i^5 | \Psi \rangle . \tag{60}$$

This difference shows that nonadiabatic effects are much less than the desired 2%.

A calculation, similar to that for H_2, has been reported for HD (Bishop and Cheung, 1978b). They used an 858-term wave function composed of 702 basis functions [(5321/5) with 9 radial functions], with $(k_i + l_i)$ even (Σ_g type) and 156 basis functions [(4210/4) with 6 radial functions] with $(k_i + l_i)$ odd (Σ_u type). The parameters k_i and l_i are powers of η_1 and η_2 and, because of the heteronuclear term in the Hamiltonian $(k_i + l_i)$, can be odd for HD. The relativistic-radiative-nonadiabatic dissociation energy (products = H + D) was estimated to be 36,405.49 cm^{-1}. This agrees only moderately well with the experimental value of 36,406.2 ± 0.4 cm^{-1} (Herzberg, 1970) and the semiempirical value, based on the ionization potential of HD (Takezawa and Tanaka, 1972), of 36,405.8 ± 0.6 cm^{-1}.

The effect of Σ_u terms was to lower the rovibronic ground state energy by 0.046 cm^{-1} and this is in perfect agreement with Wolniewicz's (1975) second-order variation-perturbation value. Several ground-state expectation values were also computed.

Wolniewicz (1976b) has calculated the nonadiabatic corrections to the rotational energies of H_2, HD, and D_2. He solved the first-order perturbation equation variationally and then determined the second-order energy correction due to the nonadiabatic terms in the Hamiltonian. The operator H' of Eq. (32) was split into a part diagonal in the angular momentum representation (H'_d) and the rest (H''), and for H_2 and D_2 the perturbation equation was written as

$$(H - E_0)\Psi' = -H''_{10}\Psi_0, \tag{61}$$

where

$$H = H^0 + H'_d \tag{62}$$

and

$$H''_{10} = [J(J+1)]^{1/2} \langle 1 | L_+ | 0 \rangle / 2\mu R^2, \tag{63}$$

where L_+ has its usual significance and the subscripts 0 and 1 stand for symmetric top wave functions with $\Lambda = 0$ and 1, respectively. The correction was given as

$$E'' = 2\langle \Psi' | H''_{10} | \Psi_0 \rangle \tag{64}$$

and Ψ' was expressed as a linear combination (with variational coefficients) of products of functions based on those defined by Eqs. (13) and (57). The 198 linear coefficients in Ψ' were found by solving the inhomogeneous equations resulting from Eq. (61). The convergence error in the computed corrections was stated to be less than 10^{-3} cm^{-1}. The calculated rotational quanta, after the corrections were included, showed satisfactory agreement with experiment. The remaining discrepancies were assigned to either experimental errors or small inaccuracies in the adiabatic energies.

Using a two-term Born expansion, Eq. (40), containing the two lowest $^1\Sigma_g^+$ BO electronic wave functions (X and E,F), Orlikowski and Wolniewicz (1974) have solved the relevant pair of coupled equations. Because of the very limited expansion they only obtained approximately 20% of the nonadiabatic correction to the vibrational quanta.

Very approximate nonadiabatic corrections to the vibrational quanta have been found by Bunker (1972) and Ford (1974). They used Rayleigh–Schrödinger perturbation theory coupled with the Unsöld approximation. The latter introduces an average excitation energy and, since its value is unknown, it is usually found by inspired guesswork.

2. *The Electronic Excited States*

The number of nonadiabatic calculations for the excited states of H_2, HD, and D_2 is very small indeed. Quite recently, Dressler *et al.* (1979) studied the mutual vibronic coupling between the two lowest double-minima excited states of $^1\Sigma_g^+$ symmetry (E,F and G,K). Essentially, they used a two-term Born expansion:

$$\Psi = \chi_1(R)\Psi_1 + \chi_2(R)\Psi_2, \tag{65}$$

where Ψ_1 and Ψ_2 are the electronic wave functions for the E,F and G,K states, respectively. For these they used 40-term wave functions derived from those of Wolniewicz and Dressler (1977) referred to in Section II,C,2. The functions $\chi_i(R)$ were expressed as

$$\chi_i(R) = (1/R)F_i(R)Y_J^M(\theta, \phi), \tag{66}$$

where the $F_i(R)$ were expressed as combinations of the vibrational bound states in the E,F and G,K adiabatic potentials; the linear coefficients were found by solving the two Born coupled equations variationally. Vibronic energies and rotational constants were tabulated for H_2, D_2, and HD and compared with experiment. The irregularities of the observed vibronic E,F and G,K progressions were well reproduced and most, though not all, of the previous discrepancies between adiabatic theory and experiment were accounted for. The results, however, should not be considered as final since they suffer from two kinds of error: (a) restricted BO electronic wave functions and (b) the neglect of higher states of $^1\Sigma_g^+$ symmetry.

In an extremely important development, Jungen and Atabek (1977) have applied the methods of multichannel quantum defect theory (Seaton, 1966) to the calculation of the rovibronic energy levels of the B $^1\Sigma_u^+$ and C $^1\Pi_u$ states of H_2 and D_2. The theory is essentially an extension of the methods of collision theory into the range of negative electron energies where the scattered electron becomes a bound electron. The only imput data needed were the adiabatic potential energy curve of H_2^+ or D_2^+ in its ground state, the ionization potential, and two BO potential curves of Σ and Π symmetry. The energies, which in principle contain adiabatic and nonadiabatic effects, were found to agree to within a few cm^{-1} with the experimental values for both states. They are substantially better than the BO or adiabatic values.

3. *Dipole Transitions in HD*

The vibrational-rotational transition moments for a $(v, J) \rightarrow (v', J')$ transition in HD may be determined from the equation

$$M_{v,J,v',J'} = \langle \Psi^0_{v,J} | d | \Psi'_{v',J'} \rangle + \langle \Psi'_{v,J} | d | \Psi^0_{v',J'} \rangle, \tag{67}$$

where $d = e(r_1 + r_2)$ and $\Psi'_{v,J}$ is given by the first-order perturbation equation

$$(H^0 - E^0_{v,J})\Psi'_{v,J} = -H'\Psi^0_{v,J}. \tag{68}$$

In Eq. (68) the perturbation is the heteronuclear component of the H' of Eq. (32), i.e.,

$$H' = -\tfrac{1}{2}\mu_a \nabla_R \sum_{i=1}^{2} \nabla_i. \tag{69}$$

Wolniewicz (1975, 1976a) has solved Eq. (68) variationally using the same methods he applied to calculate the nonadiabatic corrections for the rotational quanta of H_2 (see Section V,B,1). His results were separated into the two contributions arising from the Σ_u and Π_u components of Ψ' and were tabulated for the 0–v transitions of the $R(J)$ lines for $0 \leq v \leq 4$, $0 \leq J \leq 3$. The agreement of the moments with experiment was satisfactory (a few percent) except for the 0-0 band, where the theoretical values were 1.4 times larger than the experimental ones.

This discrepancy has also been detected by Ford and Browne (1977) who used Rayleigh–Schrödinger double-perturbation theory (the perturbations are H' and d) to formulate the problem in terms of an instantaneous dipole-moment function. This function was split into components, the largest of which could be evaluated accurately from ground-state expectation values. The remaining components were found by the same sum-over-states method which had been used previously for computing the dynamic polarizability of H_2 (see Section II,C,3). Results were presented for the 0-v vibrational bands of the $P(J)$ and $R(J)$ branches with $0 \leq v \leq 6$ and $J \leq 3$. Again agreement with experiment was good except for the 0-0 band.

Recently (Tipping *et al.*, 1978) this discrepancy has been interpreted as destructive interference between the allowed and the collision-induced dipoles.

4. *Semiempirical Hamiltonians*

The effective Hamiltonian of Eq. (54) has been applied to both H_2 and D_2; the variable parameter k (see Eq. (55)) was chosen in each case so that the rms deviation between experimental and predicted vibronic energies (relative to the lowest) was minimized (Bishop and Shih, 1976, 1977). The quality of this Hamiltonian is reflected by the fact that the rms deviation is only 0.09 cm^{-1} for H_2 and 0.05 cm^{-1} for D_2. The nonadiabatic correction to the dissociation energy of H_2 can be found from this Hamiltonian; it is 0.49 cm^{-1} in comparison to the exact value of 0.42 cm^{-1}.

Kuriyan and Pritchard (1976) have added a nonadiabatic correction $\Delta V(R)$ to the relativistic-adiabatic potential curve of H_2. It was chosen so that it reproduces the experimental rovibronic energies to an accuracy of about ± 0.1 cm^{-1}. The values of $\Delta V(R)$ were tabulated for a finite number of internuclear separations.

VI. The Outlook for the Future

The 1960s saw the foundation of the basic principles and techniques for accurate calculations on one- and two-electron molecules, the 1970s saw the refinement of these methods to give even greater accuracy. What do the 1980s hold for us?

First, we foresee a much closer examination of the excited states of H_2 and its isotopes taking place. For many of these states there are, at present, no adiabatic calculations. Second, the solution of the coupled equations (36) is likely, so that rotational states and Π states can be considered nonadiabatically. Third, the accurate calculation of many properties other than rovibronic energies will be completed, e.g., intensities, polarizabilities. Finally, we are likely to see the extension of the techniques reviewed here into the field of many-electron diatomics. The computational problems are horrific, but in the end, we expect that they will be solved. Clary (1977) has made a start by using limited James–Coolidge wave functions for He_2^+ and He_2.

REFERENCES

Bates, D. R. (1972). *Chem. Phys. Lett.* **17**, 184.

Bates, D. R., Ledsham, K., and Stewart, A. L. (1953). *Philos. Trans. R. Soc. London, Ser. A* **246**, 215.

Beckel, C. L., Hansen, B. D., III, and Peek, J. M. (1970). *J. Chem. Phys.* **53**, 3681.

Beckel, C. L., Shafi, M., and Peek, J. M. (1973). *J. Chem. Phys.* **59**, 5288.

Bethe, H., and Salpeter, E. E. (1957). "Quantum Mechanics of One- and Two-Electron Atoms." Academic Press, New York.

Bishop, D. M. (1974). *Mol. Phys.* **28**, 1397.

Bishop, D. M. (1976a). *At. Data Nucl. Data Tables* **18**, 521.

Bishop, D. M. (1976b). *J. Mol. Spectrosc.* **61**, 474.

Bishop, D. M. (1977). *J. Chem. Phys.* **66**, 3842.

Bishop, D. M., and Cheung, L. M. (1977a). *Chem. Phys. Lett.* **50**, 172.

Bishop, D. M., and Cheung, L. M. (1977b). *Phys. Rev. A* **16**, 640.

Bishop, D. M., and Cheung, L. M. (1978a). *Mol. Phys.* **36**, 501.

Bishop, D. M., and Cheung, L. M. (1978b). *Chem. Phys. Lett.* **55**, 593. Erratum: *Chem. Phys. Lett.* **58**, 320 (1978).

Bishop, D. M., and Cheung, L. M. (1978c). *J. Chem. Phys.* **69**, 1881.

Bishop, D. M., and Cheung, L. M. (1978d). *J. Phys. B* **11**, 3133.

Bishop, D. M., and Cheung, L. M. (1978e). *Phys. Rev. A* **18**, 1846.

Bishop, D. M., and Cheung, L. M. (1979a). *Int. J. Quantum Chem.* **15**, 517.

Bishop, D. M., and Cheung, L. M. (1979b). *J. Mol. Spectrosc.* **75**, 462.

Bishop, D. M., and Cheung, L. M. (1979c). *Mol. Phys.* **38**, 1475.

Bishop, D. M., and Cheung, L. M. (1979d). *Phys. Rev. A* **20**, 381.

Bishop, D. M., and Cheung, L. M. (1979e). *J. Phys. B* **12**, 3135.

Bishop, D. M., and Cheung, L. M. (1979f). *Phys. Rev. A* **20**, 1310.

Bishop, D. M., and Cheung, L. M. (1979g). *Chem. Phys. Lett.* **66**, 467. See also Bishop, D. M., and Cheung, L. M. (1980). *J. Chem. Phys.* **72**, 5125.

Bishop, D. M., and Hunter, G. (1975). *Mol. Phys.* **30**, 1433.

Bishop, D. M., and Shih, S. (1976). *J. Chem. Phys.* **61**, 162.

Bishop, D. M., and Shih, S. (1977). *J. Chem. Phys.* **67**, 4313.

Bishop, D. M., and Somorjai, R. L. (1970). *J. Math. Phys.* **11**, 1150.

Bishop, D. M., and Wetmore, R. W. (1973a). *Mol. Phys.* **26**, 145.

Bishop, D. M., and Wetmore, R. W. (1973b). *J. Mol. Spectrosc.* **46**, 502.

Bishop, D. M., and Wetmore, R. W. (1974). *Mol. Phys.* **27**, 279.

Bishop, D. M., Shih, S., Beckel, C. L., Wu, F., and Peek, J. M. (1975). *J. Chem. Phys.* **63**, 4836.

Bogaard, M. P., and Orr, B. J. (1975). *MTP Review of Molecular Structure, Ser. 2* **2**, 149.

Born, M. (1951). *Nachr. Akad. Wiss. Goettingen* **1**.

Born, M., and Huang, K. (1956). "Dynamical Theory of Crystal Lattices" Oxford Univ. Press, London and New York.

Born, M., and Oppenheimer, R. (1927). *Ann. Phys. (Leipzig)* [4] **84**, 457.

Bunker, P. R. (1972). *J. Mol. Spectrosc.* **42**, 478.

Bunker, P. R., and Moss, R. E. (1977). *Mol. Phys.* **33**, 417.

Bunker, P. R., McLarnon, C. J., and Moss, R. E. (1977). *Mol. Phys.* **33**, 425.

Burrau, O. (1927). *K. Dan. Vidensk. Selsk., Mat.-Fys. Medd.* **7**, 14.

Cheung, L. M., and Bishop, D. M. (1977). *Comput. Phys. Commun.* **13**, 247.

Clary, D. C. (1977). *Mol. Phys.* **34**, 793.

Cohen, E. R., and Taylor, B. N. (1973). *J. Phys. Chem. Ref. Data* **2**, 663.

Colbourn, E. A. (1976). *Chem. Phys. Lett.* **44,** 374.

Colbourn, E. A., and Bunker, P. R. (1976). *J. Mol. Spectrosc.* **63,** 155.

Czub, J., and Wolniewicz, L. (1978). *Mol. Phys.* **36,** 1301.

Dabrowski, I., and Herzberg, G. (1977). *Trans. N.Y. Acad. Sci.* [2] **38,** 14.

Dalgarno, A., and Epstein, S. T. (1969). *J. Chem. Phys.* **50,** 2837.

Dressler, K., Gallusser, R., Quadrelli, P. and Wolniewicz, L. (1979). *J. Mol. Spectrosc.* **75,** 205.

Essén, H. (1977). *Int. J. Quantum Chem.* **12,** 721.

Fink, U., Wiggins, T. A., and Rank, D. H. (1965). *J. Mol. Spectrosc.* **18,** 384.

Ford, A. L. (1974). *J. Mol. Spectrosc.* **52,** 358.

Ford, A. L., and Browne, J. C. (1973). *Phys. Rev. A* **7,** 418.

Ford, A. L., and Browne, J. C. (1977). *Phys. Rev. A* **16,** 1992.

Ford, A. L., Browne, J. C., Shipsey, E. J., and DeVries, P. (1975). *J. Chem. Phys.* **63,** 362.

Ford, A. L., Greenawalt, E. M., and Browne, J. C. (1977). *J. Chem. Phys.* **67,** 983.

Garcia, J. D. (1966). *Phys. Rev.* **147,** 66.

Gersten, J. I. (1969). *J. Chem. Phys.* **51,** 3181.

Gonsalves, J. W., and Moss, R. E. (1979). *Chem. Phys. Lett.* **62,** 534.

Goscinski, O., and Palma, A. (1979). *Int. J. Quantum Chem.* **15,** 197.

Green, T. A., Michels, H. H., Browne, J. C., and Madsen, M. M. (1974a). *J. Chem. Phys.* **61,** 5186.

Green, T. A., Browne, J. C., Michels, H. H., and Madsen, M. M. (1974b). *J. Chem. Phys.* **61,** 5198.

Griffin, J. J., and Wheeler, J. A. (1957). *Phys. Rev.* **108,** 311.

Heitler, W. and London, F. (1927). *Z. Phys.* **44,** 455.

Herzberg, G. (1970). *J. Mol. Spectrosc.* **33,** 147.

Herzberg, G., and Jungen, C. (1972). *J. Mol. Spectrosc.* **41,** 425.

Hunter, G., Gray, B. F., and Pritchard, H. O. (1966). *J. Chem. Phys.* **45,** 3806.

Hunter, G., Yau, A. W., and Pritchard, H. O. (1974). *At. Data Nucl. Data Tables* **14,** 11.

James, H. M., and Coolidge, A. S. (1933). *J. Chem. Phys.* **1,** 825.

Johnson, B. R. (1977). *J. Chem. Phys.* **67,** 4086.

Jungen, C., and Atabek, O. (1977). *J. Chem. Phys.* **66,** 5584.

Kari, R. E., Chan, A. C. H., Hunter, G., and Pritchard, H. O. (1973). *Can. J. Chem.* **51,** 2055.

Karl, G., and Poll, J. D. (1967). *J. Chem. Phys.* **46,** 2944.

Karl, G., Poll, J. D., and Wolniewicz, L. (1975). *Can. J. Phys.* **53,** 1781.

Kołos, W. (1968). *Int. J. Quantum Chem.* **2,** 471.

Kołos, W. (1970). *Adv. Quantum Chem.* **5,** 99.

Kołos, W. (1974). *Int. J. Quantum Chem., Symp.* **8,** 241.

Kołos, W. (1975a). *Chem. Phys. Lett.* **31,** 43.

Kołos, W. (1975b). *Int. J. Quantum Chem.* **9,** 133.

Kołos, W. (1976a). *Int. J. Quantum Chem.* **10,** 217.

Kołos, W. (1976b). *J. Mol. Spectrosc.* **62,** 429.

Kołos, W. (1977). *Acta Phys. Austriaca, Suppl.* **17,** 161.

Kołos, W., and Peek, J. M. (1976). *Chem. Phys.* **12,** 381.

Kołos, W., and Rychlewski, J. (1976). *J. Mol. Spectrosc.* **62,** 109.

Kołos, W., and Rychlewski, J. (1977). *J. Mol. Spectrosc.* **66,** 428.

Kołos, W., and Rychlewski, J. (1978). *Acta Phys. Pol.* **A53,** 281.

Kołos, W., and Wolniewicz, L. (1963). *Rev. Mod. Phys.* **35,** 473.

Kołos, W., and Wolniewicz, L. (1964). *J. Chem. Phys.* **41,** 3663.

Kołos, W., and Wolniewicz, L. (1965). *J. Chem. Phys.* **43,** 2429.
Kołos, W., and Wolniewicz, L. (1966). *J. Chem. Phys.* **45,** 509.
Kołos, W., and Wolniewicz, L. (1967). *J. Chem. Phys.* **46,** 1426.
Kołos, W., and Wolniewicz, L. (1968a). *J. Chem. Phys.* **48,** 3672.
Kołos, W., and Wolniewicz, L. (1968b). *J. Chem. Phys.* **49,** 404.
Kołos, W., and Wolniewicz, L. (1969). *J. Chem. Phys.* **50,** 3228.
Kołos, W., and Wolniewicz, L. (1974). *Chem. Phys. Lett.* **24,** 457.
Kołos, W., and Wolniewicz, L. (1975a). *Can. J. Phys.* **53,** 2189.
Kołos, W., and Wolniewicz, L. (1975b). *J. Mol. Spectrosc.* **54,** 303.
Kuriyan, M., and Pritchard, H. O. (1976). *Can. J. Chem.* **54,** 1865.
Kuriyan, M., and Pritchard, H. O. (1977). *Can. J. Chem.* **55,** 3420.
Kutzelnigg, W. (1963). *Theor. Chim. Acta* **1,** 327.
Laskowski, B., Diamond, J., Waleh, A., and Hudson, B. (1978). *J. Chem. Phys.* **69,** 5222.
Lathouwers, L. (1978). *Phys. Rev. A* **18,** 2150.
Lathouwers, L., Van Leuven, P., and Bouten, M. (1977). *Chem. Phys. Lett.* **52,** 439.
LeRoy, R. J., and Barwell, M. G. (1975). *Can. J. Phys.* **53,** 1983.
LeRoy, R. J., and Bernstein, R. B. (1968). *J. Chem. Phys.* **49,** 4312.
Longuet-Higgins, H. C. (1961). *Adv. Spectrosc.* **2,** 429.
Luke, S. K., Hunter, G., and McEachran, R. P. (1969). *J. Chem. Phys.* **50,** 1644.
McGurk, N. J. (1977). *Phys. Rev. C* **15,** 1924.
Madsen, M. M., and Peek, J. M. (1971). *At. Data* **2,** 171.
Michels, H. H. (1966). *J. Chem. Phys.* **44,** 3834.
Montgomery, H. E., Jr. (1977). *Chem. Phys. Lett.* **50,** 455.
Montgomery, H. E., Jr. (1978). *Chem. Phys. Lett.* **56,** 307.
Montgomery, H. E., Jr., and Rubenstein, T. G. (1978). *Chem. Phys. Lett.* **58,** 295.
Murai, T., and Takatsu, H. (1974). "Contributions from the Research Group on Atoms and Molecules," No. 10, p. 74. Ochanomizu University, Japan.
Murai, T., and Takatsu, H. (1976). "Contributions from the Research Group on Atoms and Molecules," No. 12, p. 99. Ochanomizu University, Japan.
Orlikowski, T., and Wolniewicz, L. (1974). *Chem. Phys. Lett.* **24,** 461.
Peek, J. M. (1965). Research Report SC-RR-65-77. Sandia Corp., Albuquerque, New Mexico.
Peyerimhoff, S. (1965). *J. Chem. Phys.* **43,** 998.
Poll, J. D., and Karl, G. (1966). *Can. J. Phys.* **44,** 1467.
Poll, J. D., and Wolniewicz, L. (1978). *J. Chem. Phys.* **68,** 3053.
Pople, J. A. (1965). *J. Chem. Phys.* **43,** S229.
Power, J. D. (1973). *QCPE* **11,** 233.
Price, R. I. (1978). *Chem. Phys.* **31,** 309.
Pritchard, H. O., and Wolniewicz, L. (1976). *Can. J. Chem.* **54,** 651.
Reid, R. V., Jr., and Vaida, M. L. (1975). *Phys. Rev. Lett.* **34,** 1064.
Seaton, M. J. (1966). *Proc. Phys. Soc., London* **88,** 801, 815.
Shafi, M., and Beckel, C. L. (1973). *J. Chem. Phys.* **59,** 5294.
Stwalley, W. C. (1970). *Chem. Phys. Lett.* **6,** 241.
Takezawa, S., and Tanaka, V. (1972). *J. Chem. Phys.* **56,** 6125.
Tipping, R. H., Poll, J. D., and McKellar, A. R. W. (1978). *Can. J. Phys.* **56,** 75.
Wicke, B. G., and Harris, D. O. (1976). *J. Chem. Phys.* **64,** 5236.
Wing, W. H., Ruff, G. A., Lamb, W. E., Jr., and Spezeski, J. J. (1976). *Phys. Rev. Lett.* **36,** 1488.
Winter, T. G., Duncan, M. D., and Lane, N. F. (1977). *J. Phys. B* **10,** 285.

Wolniewicz, L. (1966). *J. Chem. Phys.* **45,** 515.
Wolniewicz, L. (1975). *Can. J. Phys.* **53,** 1207.
Wolniewicz, L. (1976a). *Can. J. Phys.* **54,** 672.
Wolniewicz, L. (1976b). *J. Mol. Spectrosc.* **63,** 537.
Wolniewicz, L. and Dressler, K. (1977). *J. Mol. Spectrosc.* **67,** 416.
Wolniewicz, L. and Poll, J. D. (1978). *J. Mol. Spectrosc.* **72,** 264.
Woolley, R. G. (1978). *J. Am. Chem. Soc.* **100,** 1073.
Woolley, R. G., and Sutcliffe, B. T. (1977). *Chem. Phys. Lett.* **45,** 393.

Variational Approach to Orthogonality and Normality of Basis Ket-Vectors

Y. J. I'HAYA
Department of Materials Science
The University of Electro-Communications
Chofu, Tokyo, Japan

and

TETSUO MORIKAWA
School of Pharmaceutical Science
Toho University
Funabashi, Chiba, Japan

I. Introduction

In studying the properties of atoms, molecules, and crystals by means of quantum mechanics, one of the most essential characteristics one encounters is the nonorthogonality problem. Löwdin (1947, 1950) first explained the cohesive and elastic properties of alkali halide crystals by using the orthogonalized wave functions. At that time, he proposed a recipe for construction of orthonormal functions, later called the symmetric orthonormalization, expanding a method given by Landshoff (1936). As a method to build up the orthonormal vector set $\{|\phi_i\rangle\} \equiv \boldsymbol{\phi}$ from the nonorthonormal set $\{|\psi_i\rangle\} \equiv \boldsymbol{\psi}$, we recall the well-known Schmidt procedure. This is to construct orthonormal $|\phi_i\rangle$'s step by step by taking out

ISBN 0-12-034812-8

$|\psi_i\rangle$'s one by one from ψ. Since a way of taking out $|\psi_i\rangle$ is arbitrary, one can construct various kinds of ϕ. On the other hand, the symmetric orthonormalization brings forth the only ϕ by taking out all $|\psi_i\rangle$'s equivalently, i.e., symmetrically. In this respect it may safely be said that the symmetric orthonormalization has advantages over the Schmidt method. As will be described in the following sections, our variational method adopts the symmetric orthonormalization as a starting point. Therefore, the method has the same advantages as the symmetric orthonormalization does.

The nonorthogonality problem in molecular orbital theory of atoms and molecules was studied by Pauling (1931), Slater (1931), and others, who developed the so-called maximum overlapping principle from the consideration that the strength of a chemical bond should be measured by the degree of the overlap between the orbitals involved. In this sense, the principle is closely related to the maximum nonorthogonality. Pauling's hybrid orbitals are constructed from an atomic orbital set so as to satisfy, first, the normalization, second, the orthogonality, and third, an implicit condition that hybrid orbitals can be spanned completely by a set of atomic orbitals. The last condition, which may be called a unit orbital contribution, cannot be overlooked in our variation method for ket-vectors.

The degree of orthogonality and normality, or more generally, of overlapping, is measured by the inner product of two ket-vectors $|\psi\rangle$ and $|\theta\rangle$, i.e., the overlap integral $\langle\psi|\theta\rangle$. Here, $\langle\psi|\theta\rangle$ may contain an operator such as $|\theta\rangle = F|\eta\rangle$, $|\eta\rangle$ being another ket-vector. Furthermore, the measure is preferably set in a form such as $\langle\psi|\theta\rangle\langle\theta|\psi\rangle = \|\langle\psi|\theta\rangle\|^2 \equiv \omega$, so as to take a real value. In the first place, a key to the problem of nonorthogonality and nonnormality is therefore in the consideration of a metric matrix, the overlap of two vectors. Next, how to choose the measure ω depends upon the properties of a physical system under consideration. Afterward, one has only to optimize ω variationally.

That one can build up and/or cancel the overlapping just as one wishes is based upon the superposition principle, i.e., the linearity of ket-vectors. From another assumption that the inner product exists, one can consider the complete orthogonality $\langle\psi|\theta\rangle = 0$ and the normality $\langle\psi|\psi\rangle = 1$ as special cases. As is well known, these have not only a mathematical importance but a close connection to quantum mechanical phenomena. The last of the necessary assumptions is that $\langle\psi|\psi\rangle = 0$ if and only if $|\psi\rangle = 0$; this is important in relation to whether the vectors and states under consideration are possible, i.e., whether they are linearly independent.

In 1970 Löwdin published a review article on the nonorthogonality problem in Volume 5 of this series. At present it seems timely to review more closely the same problem from a different point of view. Section II contains a general description of a new variational approach for ket-vectors with the orthonormality constraint (orthonormality-constrained variation, hereafter abbreviated ONCV), a method of characterization of eigenvalues by use of ONCV, comments on the stability of solutions, and derivation of SCF equations with comparison to the usual Lagrange multipliers method, clarifying significance of some unrevealed aspects of Löwdin's orthonormalization method. Section III deals with the maximum and minimum overlap, localized, and hybrid orbitals using a sum of the squared overlap as a measure. Section IV briefly describes a relation of the ONCV method to the Edmiston–Ruedenberg localized orbital method. And finally, Section V expands the ONCV method to a pair of ket-vector sets and proposes a method for constraining pair-orthogonality—orthogonality between two basis sets—in the course of variation.

II. Orthonormality-Constrained Variation (ONCV) for Ket-Vectors

A. Symmetric ONCV

We must first establish a variation method for ket-vectors. Conventional variation methods optimize an assumed functional of vectors (in real form) together with constrained conditions which are functions in real form as well. Different from this, our aim is to treat ket-vectors themselves variationally under constrained conditions.

Let us consider the orthonormality as a most typical metric. A problem to be solved is as follows: Let the orthonormality be fulfilled in every step of arbitrary variations for a given set of orthonormal ket-vectors $|\phi_i\rangle$'s. For the sake of simplicity, we denote a set of $|\phi_i\rangle$'s by a row matrix

$$\boldsymbol{\phi} \equiv (|\phi_1\rangle \quad |\phi_2\rangle \quad \cdots) \tag{1}$$

and its adjoint by a column matrix

$$\boldsymbol{\phi}^\dagger \equiv \begin{pmatrix} \langle\phi_1| \\ \langle\phi_2| \\ \vdots \end{pmatrix}, \tag{2}$$

where $\boldsymbol{\phi}^\dagger\boldsymbol{\phi} = \mathbf{I}$ (unit matrix). We write $\boldsymbol{\phi} \rightarrow \boldsymbol{\phi} + \delta\boldsymbol{\phi} \equiv \boldsymbol{\psi}$ for an arbitrary variation of $\boldsymbol{\phi}$, $\delta\boldsymbol{\phi}$ being an arbitrary and infinitesimal variation vector. It is a matter of course that $\boldsymbol{\psi}$ here is not orthonormal. As is well known, if the shifted vectors $|\psi_i\rangle$ are linearly independent, then $\boldsymbol{\psi}^\dagger\boldsymbol{\psi}$ turns out to be a positive definite matrix. Therefore, the matrix $(\boldsymbol{\psi}^\dagger\boldsymbol{\psi})^{-1/2}$ is uniquely determined. In this case, $\boldsymbol{\psi}(\boldsymbol{\psi}^\dagger\boldsymbol{\psi})^{-1/2}$ corresponds exactly to Löwdin's symmetric orthonormalization (Löwdin, 1947, 1950), a method to build up an orthonormal vector set from a nonorthonormal one. We can make sure of its orthonormality since

$$\{\boldsymbol{\psi}(\boldsymbol{\psi}^\dagger\boldsymbol{\psi})^{-1/2}\}^\dagger\{\boldsymbol{\psi}(\boldsymbol{\psi}^\dagger\boldsymbol{\psi})^{-1/2}\} = (\boldsymbol{\psi}^\dagger\boldsymbol{\psi})^{-1/2}(\boldsymbol{\psi}^\dagger\boldsymbol{\psi})(\boldsymbol{\psi}^\dagger\boldsymbol{\psi})^{-1/2} = \mathbf{I}. \tag{3}$$

If we expand Eq. (3) in terms of the arbitrary variation $\delta\boldsymbol{\phi}$ to first order, we have

$$\bar{\boldsymbol{\phi}} \equiv \boldsymbol{\phi} + \delta\boldsymbol{\phi} - \tfrac{1}{2}\boldsymbol{\phi}(\delta\boldsymbol{\phi}^\dagger\boldsymbol{\phi} + \boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi}), \tag{4}$$

with $\bar{\boldsymbol{\phi}}^\dagger\bar{\boldsymbol{\phi}} = \mathbf{I}$. The process $\boldsymbol{\phi} \to \bar{\boldsymbol{\phi}}$ thus is an answer to the problem that we set up at the outset (Morikawa, 1972). We shall call this the symmetric ONCV method hereafter. If we use the partitioned $\boldsymbol{\phi}$,

$$\boldsymbol{\phi} \equiv (\boldsymbol{\phi}_1|\boldsymbol{\phi}_2|\cdots), \qquad \boldsymbol{\phi}_i \equiv (|\phi_{i1}\rangle \quad |\phi_{i2}\rangle \quad \cdots), \tag{5}$$

the partitioned relationship corresponding to Eq. (4) becomes

$$\boldsymbol{\phi}_i \to \bar{\boldsymbol{\phi}}_i \equiv \boldsymbol{\phi}_i + \delta\boldsymbol{\phi}_i - \tfrac{1}{2}\sum_{k=1}^{m} \boldsymbol{\phi}_k(\delta\boldsymbol{\phi}^\dagger_k\boldsymbol{\phi}_i + \boldsymbol{\phi}^\dagger_k\,\delta\boldsymbol{\phi}_i). \tag{6}$$

Equations (4) and (6) can be written in an operator form as (Morikawa, 1975a)

$$\boldsymbol{\phi} \to (\hat{P}_\phi + \hat{t})\boldsymbol{\phi}, \tag{7}$$

$$|\phi_i\rangle \to (\hat{P}_\phi + \hat{t})|\phi_i\rangle, \tag{8}$$

where

$$\hat{P}_\phi \equiv \boldsymbol{\phi}\boldsymbol{\phi}^\dagger = \sum_{k=1}^{m} \boldsymbol{\phi}_k\boldsymbol{\phi}_k^\dagger, \tag{9}$$

$$\hat{t} \equiv \delta\boldsymbol{\phi}\boldsymbol{\phi}^\dagger - \tfrac{1}{2}\boldsymbol{\phi}(\delta\boldsymbol{\phi}^\dagger\boldsymbol{\phi} + \boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi})\boldsymbol{\phi}^\dagger. \tag{10}$$

What the process $\boldsymbol{\phi} \to \bar{\boldsymbol{\phi}}$ means physically and where it can be applied will be made clearer in the successive sections.

B. Application to Eigenvalue Problems

Most cases when the orthonormality appears are typically eigenvalue problems, and hence investigation of eigenvalue problems is of great importance in theoretical formulations and practical calculations. Generally speaking, eigenvalues cannot be determined accurately, so that one must strive to set up the corresponding variation equations with the restriction of orthonormality.

First, we consider the expectation value of a self-adjoint operator

$F(F^\dagger = F)$ which is bounded from below and does not explicitly depend upon $\boldsymbol{\phi}$:

$$\omega(\boldsymbol{\phi}) = \sum_i \omega_i = \mathrm{tr}(\boldsymbol{\phi}^\dagger F \boldsymbol{\phi}), \tag{11}$$

with

$$\omega_i \equiv \langle \phi_i | F | \phi_i \rangle. \tag{12}$$

Note that ω is a real quantity and $\boldsymbol{\phi}$ is composed of many components. Also note $|\phi_i\rangle\langle\phi_i| \neq 1$. If $\boldsymbol{\phi}$ is composed of only one component and F is a Hamiltonian operator, then the measure ω is the total energy of the system.

An arbitrary variation of ω leads to

$$\delta\omega(\boldsymbol{\phi}) = \mathrm{tr}[\delta\boldsymbol{\phi}^\dagger F \boldsymbol{\phi} + \text{Hermitian conjugate (H.c.)}], \tag{13}$$

which does not meet our present requirement. Instead, we must optimize ω under the orthonormality constraints. Though the Lagrange multipliers method is usually used for this purpose, we here apply the ONCV method to the variation of ω. If we define $\bar{\delta}\omega \equiv \omega(\bar{\boldsymbol{\phi}}) - \omega(\boldsymbol{\phi})$, then a simple calculation using Eq. (4) shows that

$$\bar{\delta}\omega = \mathrm{tr}(\bar{\boldsymbol{\phi}}^\dagger F \bar{\boldsymbol{\phi}}) - \mathrm{tr}(\boldsymbol{\phi}^\dagger F \boldsymbol{\phi}) = \mathrm{tr}\{\delta\boldsymbol{\phi}^\dagger (1 - \boldsymbol{\phi}\boldsymbol{\phi}^\dagger) F \boldsymbol{\phi} + \mathrm{H.c.}\}, \tag{14}$$

retaining terms in first order in $\delta\boldsymbol{\phi}$, so that different from the arbitrary variation $\delta\omega$, $\bar{\delta}\omega$ turns out to be an infinitesimally varied quantity which satisfies the orthonormality conditions. Since $\delta\boldsymbol{\phi}$ is an arbitrary infinitesimal variation vector, a solution to the equation $\bar{\delta}\omega = 0$ is

$$(1 - \boldsymbol{\phi}\boldsymbol{\phi}^\dagger) F \boldsymbol{\phi} = 0, \tag{15}$$

or we can write

$$F\boldsymbol{\phi} = \boldsymbol{\phi}(\boldsymbol{\phi}^\dagger F \boldsymbol{\phi}) = \boldsymbol{\phi}\boldsymbol{\varepsilon}, \qquad \text{with} \quad \boldsymbol{\varepsilon} \equiv \boldsymbol{\phi}^\dagger F \boldsymbol{\phi}. \tag{16}$$

On the other hand, since ω is invariant under a unitary transformation among $|\phi_i\rangle$'s, as is F, the form of $\boldsymbol{\phi}$ is not disturbed by this transformation (Morikawa and I'Haya, 1978b). Therefore, we can choose a specific $\boldsymbol{\phi}$ which diagonalizes the Hermitian matrix $\boldsymbol{\varepsilon}$:

$$F\boldsymbol{\phi} = \boldsymbol{\phi}\boldsymbol{\varepsilon}_{\mathrm{d}}, \qquad \boldsymbol{\varepsilon}_{\mathrm{d}}: \ \text{diagonal}. \tag{17}$$

We thus obtain the eigenvalue equation of the form

$$F|\phi_i\rangle = |\phi_i\rangle \varepsilon_i. \tag{18}$$

Equation (18) includes the Schrödinger equation as a special case. If $\boldsymbol{\phi}$ is a solution to the equation $\bar{\delta}\omega = 0$, i.e., $\boldsymbol{\phi}$ is stable, then $\omega_i = \varepsilon_i$ ($i = 1, 2, \ldots$) are eigenvalues of F and the corresponding vectors become eigenvectors. Thus, a method of variational characterization of eigenvalues is considered to be a basis for the variational treatment of the operator F, on the one hand, and of special importance as the starting point of an argument of upper and lower bounds of eigenvalues, on the other.

Two procedures, recursive characterization and independent (or maximum–minimum) characterization, have been studied in detail (Gould, 1966). In the former, the eigenvalues $\varepsilon_i (i = 1, 2, \ldots)$ are characterized by the recursive use of variation; using the first eigenfunction ϕ_1, which corresponds to the lowest eigenvalue ε_1, the second eigenvalue ε_2 is obtained by the minimization of $\langle \phi_2 | F | \phi_2 \rangle$ subject to the orthogonality condition $\langle \phi_1 | \phi_2 \rangle = 0$ etc. The latter is based on the Rayleigh theorem, which states that a relationship between ω_i and ε_i is given as $\varepsilon_i \leq \omega_i \leq \varepsilon_{i+1}$ under single arbitrary constraint; viz., ϕ which gives an ith eigenvalue of the system is the one that makes $\langle \phi | F | \phi \rangle$ minimum under $i - 1$ constraint conditions. Both procedures are reputed to be always capable of characterizing eigenvalues from a purely theoretical standpoint.

However, it is not clear how one can carry out variation of ket-vectors under given constraints, and what kinds of variational methods one can use. The above-mentioned example of derivation shows that our symmetric ONCV method yields a concrete procedure for characterization of eigenvalues, different from the recursive and independent characterization methods. It may be called a method of symmetric characterization and is discussed in detail in the literature (Morikawa and I'Haya, 1978a).

The equation $\bar{\delta}\omega \to 0$ is thus considered a variational equation including the neighborhood of $\bar{\delta}\omega = 0$. In other words, it is to give an equation of variational expression in which an eigenvalue equation is involved as a specific point. The variational equation of this type will play an important role in what follows.

C. Higher-Order ONCV Method

We now examine the stability of a solution in the neighborhood of the equation $\bar{\delta}\omega = 0$ by use of the ONCV method. From such information it will become clear whether the equation $\bar{\delta}\omega = 0$ is sufficient to obtain a solution that leaves nothing to be desired, and whether the solution is stable. For this purpose, we develop an ONCV method which takes care of second-order terms in $\delta\boldsymbol{\phi}$ as well as first-order terms.

Expansion of $\boldsymbol{\psi}(\boldsymbol{\psi}^\dagger\boldsymbol{\psi})^{-1/2}$ up to second order leads to

$$\begin{aligned}\bar{\bar{\boldsymbol{\phi}}} \equiv \boldsymbol{\phi} + \delta\boldsymbol{\phi} - \tfrac{1}{2}\boldsymbol{\phi}(\delta\boldsymbol{\phi}^\dagger\boldsymbol{\phi} + \boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi}) \\ - \tfrac{1}{2}\,\delta\boldsymbol{\phi}(\delta\boldsymbol{\phi}^\dagger\boldsymbol{\phi} + \boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi}) - \tfrac{1}{2}\boldsymbol{\phi}\,\delta\boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi} + \tfrac{3}{8}\boldsymbol{\phi}(\delta\boldsymbol{\phi}^{\dagger\dagger}\boldsymbol{\phi} + \boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi})^2,\end{aligned} \tag{19}$$

where we denote $\bar{\bar{\boldsymbol{\phi}}}$ in order to distinguish it from $\bar{\boldsymbol{\phi}}$ in Eq. (4). The procedure $\boldsymbol{\phi} \to \bar{\bar{\boldsymbol{\phi}}}$ may be called a second-order ONCV method. Higher-order ONCV methods beyond third order can be developed in a similar manner. The quantity

$$\bar{\bar{\delta}}\omega \equiv \omega(\bar{\bar{\boldsymbol{\phi}}}) - \omega(\boldsymbol{\phi}) \tag{20}$$

corresponds to the variation of the measure ω up to second order, satisfying the orthonormality conditions; i.e.,

$$\begin{aligned} \bar{\bar{\delta}}\omega = {} & \mathrm{tr}\{\delta\boldsymbol{\phi}^{\dagger}(1 - \boldsymbol{\phi}\boldsymbol{\phi}^{\dagger})F\boldsymbol{\phi} + \mathrm{H.c.}\} \\ & + \mathrm{tr}\{\delta\boldsymbol{\phi}^{\dagger}(1 - \boldsymbol{\phi}\boldsymbol{\phi}^{\dagger})F\boldsymbol{\phi}\boldsymbol{\phi}^{\dagger}\,\delta\boldsymbol{\phi} + \mathrm{H.c.}\} + \cdots . \end{aligned} \tag{21}$$

From this, it can be mentioned (refer to Epstein, 1974) that the first two statements are in accord with those described by the well-known variation principle.

(1) If $\boldsymbol{\phi}$ is an eigenvector of F, i.e., $(1 - \boldsymbol{\phi}\boldsymbol{\phi}^{\dagger})F\boldsymbol{\phi} = 0$, then the first term in Eq. (21) vanishes and only the terms of second order in $\delta\boldsymbol{\phi}$ remain. Therefore, ω is stationary at $\boldsymbol{\phi}$ when $\boldsymbol{\phi}$ changes continuously. The error in ω is second order in the error in $\boldsymbol{\phi}$.

(2) Since $\bar{\delta}\omega \to 0$ for an arbitrary variation $\delta\boldsymbol{\phi}$, one can put

$$\delta\boldsymbol{\phi} = -\mu(1 - \boldsymbol{\phi}\boldsymbol{\phi}^{\dagger})F\boldsymbol{\phi} = -\mu(F\boldsymbol{\phi} - \boldsymbol{\phi}\boldsymbol{\varepsilon}_{\mathrm{d}}), \tag{22}$$

where μ is a complex number, and obtain

$$\bar{\delta}\omega = -(\mu + \mu^{*})\,\mathrm{tr}\{(F\boldsymbol{\phi} - \boldsymbol{\phi}\boldsymbol{\varepsilon}_{\mathrm{d}})^{\dagger}(F\boldsymbol{\phi} - \boldsymbol{\phi}\boldsymbol{\varepsilon}_{\mathrm{d}})\} \to 0. \tag{23}$$

This is valid only when $F\boldsymbol{\phi} - \boldsymbol{\phi}\boldsymbol{\varepsilon}_{\mathrm{d}} = 0$. Therefore, if $\boldsymbol{\phi}$ is a stationary point (or vector), $\varepsilon_i \equiv \langle\phi_i|F|\phi_i\rangle$ is an eigenvalue of the operator F and $|\phi_i\rangle$'s are the corresponding eigenvectors. The ω_i has no stationary point other than ε_i.

(3) Suppose that $\delta\boldsymbol{\phi}$ is chosen as a vector that belongs to the same type as $\boldsymbol{\phi}$; the type here implies, e.g., an irreducible representation of point groups. If so, then $\boldsymbol{\psi} = \boldsymbol{\phi} + \delta\boldsymbol{\phi}$ will belong to the same type as well. It is clear that $\bar{\boldsymbol{\phi}}$ also belongs to the same type, since $\boldsymbol{\psi}(\boldsymbol{\psi}^{\dagger}\boldsymbol{\psi})^{-1/2}$ is invariant under a unitary transformation (Slater and Koster, 1954); i.e., the process $\bar{\delta}\omega \to 0$ is used only for obtaining a stationary point in variation space of the same type. This is just what the symmetric ONCV method means physically. One should consider that, in connection with item (1), the variation principle, as well as the ONCV procedure, are to be applied to eigenvalues and eigenvectors of different types individually. For example, one can obtain an optimal state of each irreducible representation of the point group under consideration.

(4) When $\boldsymbol{\phi}$ is composed of just one component, say $|\phi_1\rangle$, we arrange the eigenvalues of F in the order $\varepsilon_1 \le \varepsilon_2 \le \cdots$, with $F - \varepsilon_1 \ge 0$. For a rather large $|\delta\phi_1\rangle$, the relationship $\langle\delta\phi_1|(F - \varepsilon_1)|\delta\phi_1\rangle > 0$ holds, i.e.,

$\varepsilon_1 = \omega$ is a stationary point and an absolute minimum of ω. In other words, $\langle \bar{\phi}_1|F|\bar{\phi}_1\rangle$ is an upper bound to the minimum eigenvalue. The proof is simple. Letting ε_i ($i \neq 1$) be higher eigenvalues and choosing the variation vector as $|\delta\phi_i\rangle = \sum_{j<i}|\phi_j\rangle$, we obtain $\langle\delta\phi_i|F - \varepsilon_i|\phi_i\rangle \leq 0$. The choice of $|\delta\phi_i\rangle = \sum_{j>i}|\phi_j\rangle$ leads to $\langle\delta\phi_i|F - \varepsilon_i|\phi_i\rangle \geq 0$, i.e., ω is a saddle point at ε_i, neither a maximum nor a minimum. Thus ω, the functional of $\boldsymbol{\phi}$, is an upper bound to the minimum eigenvalue ($\varepsilon_i \leq \omega_i$).

D. Derivation of SCF Equations

In this subsection, we consider the case that the operator F depends upon $|\phi_i\rangle$. The most well-known example is the one for the so-called closed-shell model of atoms or molecules, the Hartree–Fock operator defined by (Roothaan, 1951)

$$F = H + \sum_{i}^{\text{occ}} (2J_i - K_i), \tag{24}$$

where H is the core repulsion operator, and J_i and K_i are the Coulomb and exchange operators, respectively. Note that the latter two depend upon $|\phi_i\rangle$.

The total energy of the system is expressed as

$$\omega_c \equiv \sum_{i}^{\text{occ}} \langle\phi_i|F + H|\phi_i\rangle = \text{tr}\{\boldsymbol{\phi}^\dagger(F + H)\boldsymbol{\phi}\}. \tag{25}$$

An arbitrary variation of Eq. (25) becomes

$$\delta\omega_c = 2\,\text{tr}(\delta\boldsymbol{\phi}^\dagger F\boldsymbol{\phi} + \text{H.c.}). \tag{26}$$

Here we have used the relationship

$$\text{tr}(\boldsymbol{\phi}^\dagger\,\delta F\boldsymbol{\phi}) = \text{tr}(\delta\boldsymbol{\phi}^\dagger F\boldsymbol{\phi} + \boldsymbol{\phi}^\dagger F\,\delta\boldsymbol{\phi}), \tag{27}$$

which will be utilized often in later sections. The symmetric ONCV procedure for ω_c leads to

$$\bar{\delta}\omega_c = 2\,\text{tr}\{\delta\boldsymbol{\phi}^\dagger(1 - \boldsymbol{\phi}\boldsymbol{\phi}^\dagger)F\boldsymbol{\phi} + \text{H.c.}\}. \tag{28}$$

Then, from the equation $\bar{\delta}\omega_c = 0$, we obtain the well-known equation

$$F\boldsymbol{\phi} = \boldsymbol{\phi}\boldsymbol{\varepsilon}, \qquad \text{with} \quad \boldsymbol{\varepsilon} \equiv \boldsymbol{\phi}^\dagger F\boldsymbol{\phi}. \tag{29}$$

Since F and ω are invariant under unitary transformation among $|\phi_i\rangle$'s, Eq. (29) is transformed into an eigenvalue equation.

The extended Hartree–Fock and/or unrestricted Hartree–Fock methods are usually used for calculations of the doublet and triplet states of free radicals and molecules (Pople and Nesbet, 1954). The $\delta\omega$ in these methods are $\delta\omega_\alpha = 2\,\text{tr}(\delta\boldsymbol{\phi}_\alpha^\dagger F_\alpha\boldsymbol{\phi}_\alpha + \text{H.c.})$ for the α-spin-orbital set $\boldsymbol{\phi}_\alpha$

and have the same form for the β-spin-orbital set $\boldsymbol{\phi}_\beta$. The ONCV method should be applied separately to $\delta\omega_\alpha$ and $\delta\omega_\beta$, because $\boldsymbol{\phi}_\alpha$ and $\boldsymbol{\phi}_\beta$ can be chosen independently:

$$\begin{aligned}\bar{\delta}\omega_\alpha &= 2\ \mathrm{tr}\{\delta\boldsymbol{\phi}_\alpha^\dagger(1 - \boldsymbol{\phi}_\alpha\boldsymbol{\phi}_\alpha^\dagger)F_\alpha\boldsymbol{\phi}_\alpha + \mathrm{H.c.}\},\\ \bar{\delta}\omega_\beta &= 2\ \mathrm{tr}\{\delta\boldsymbol{\phi}_\beta^\dagger(1 - \boldsymbol{\phi}_\beta\boldsymbol{\phi}_\beta^\dagger)F_\beta\boldsymbol{\phi}_\beta + \mathrm{H.c.}\}.\end{aligned} \tag{30}$$

Thus, we obtain the variation equations $\bar{\delta}\omega_\alpha \to 0$ and $\bar{\delta}\omega_\beta \to 0$ corresponding to the extended Hartree–Fock method. $\boldsymbol{\phi}_\alpha$ and $\boldsymbol{\phi}_\beta$ have the degree of freedom of a unitary transformation. By using the freedom, we will be able to make them orthogonal. The ONCV method effective for a pair of set vectors will be developed in a later section.

Taking a step forward, we consider the case that individual F_i depend upon each $|\phi_i\rangle$:

$$\delta\boldsymbol{\omega}_{\mathrm{SCF}} \equiv 2\left\{\sum_i^{\mathrm{occ}} \langle\delta\phi_i|F_i|\phi_i\rangle + \mathrm{H.c.}\right\}, \tag{31}$$

where

$$F_i = f_i\left[H + \sum_j (2a_{ij}J_j - b_{ij}K_j)\right]. \tag{32}$$

The case has been investigated frequently in many papers (Roothaan, 1960; Birss and Fraga, 1963; Huzinaga, 1969). It is well know that variational treatments of many-electron systems in atoms and molecules lead to an SCF equation. In that case, in order to maintain the orthonormality for orbitals ϕ_i the Lagrange multipliers method is usually applied, introducing an arbitrary parameter θ_{ji} (Roothaan, 1951, 1960). It has been often pointed out (Goddard *et al.*, 1969) that due to an incorrect treatment of the off-diagonal terms of θ_{ji}, errors of certain type have appeared in the literature on this subject. In order to improve this situation, another parameter has been introduced, so that two kinds of arbitrary parameters are involved in the usual SCF operator. Of these, the latter parameter exerts influence upon the form of the so-called coupling operator.

The ONCV method does not contain any arbitrary parameter and gives directly an exact SCF equation when applied to $\delta\omega_{\mathrm{SCF}}$. In what follows, we will generalize the symmetric ONCV method and discuss where those parameters mentioned earlier come from and why unavoidable errors arise in the conventional SCF theory. The method may then be called a generalized ONCV method in which any unknown parameter is not involved (Morikawa, 1975b). We will be able to clarify the physical meaning of the individual vector appearing in the ONCV procedure.

To begin with, we substitute λ_{ik} for the coefficient $-\frac{1}{2}$ before the third term in Eq. (6):

$$|\phi_i\rangle \rightarrow |\bar{\phi}_i\rangle \equiv |\phi_i\rangle + |\delta\phi_i\rangle + \sum_k \lambda_{ik}|\phi_k\rangle(\langle\delta\phi_k|\phi_i\rangle + \langle\phi_k|\delta\phi_i\rangle). \quad (33)$$

This implies that the constraint $\delta\langle\phi_k|\phi_i\rangle = \langle\delta\phi_k|\phi_i\rangle + \langle\phi_k|\delta\phi_i\rangle = 0$ multiplied by the undetermined multiplier λ_{ik}, as well as spanned by the vector $|\phi_k\rangle$, is added to the arbitrary variation of the vector $|\phi_i\rangle$, $|\phi_i\rangle + |\delta\phi_i\rangle$.

What sort of relationship is required among the λ_{ij}'s in order that the varied $|\bar{\phi}_i\rangle$'s maintain the orthonormality? From Eq. (33), we have the following equation to first order in δ:

$$\langle\bar{\phi}_j|\bar{\phi}_i\rangle = \langle\phi_j|\phi_i\rangle + (\langle\delta\phi_j|\phi_i\rangle + \langle\phi_j|\delta\phi_i\rangle)(\lambda_{ij} + \lambda_{ji}^* + 1). \quad (34)$$

Hence the relationship

$$\lambda_{ij} + \lambda_{ji}^* + 1 = 0 \quad (35)$$

or its matrix form

$$\boldsymbol{\lambda} + \boldsymbol{\lambda}^\dagger + \mathbf{u} = 0 \qquad (u_{ij} = 1 \quad \text{for all} \quad i \text{ and } j) \quad (36)$$

is the necessary and sufficient condition for the varied vectors $\{|\bar{\phi}_i\rangle\}$ to satisfy the orthonormality. If the λ_{ij}'s are assumed to be real numbers, then Eq. (35) becomes

$$\lambda_{ij} + \lambda_{ji} + 1 = 0, \quad (37)$$

from which all diagonal elements λ_{ii} are found to be $-\frac{1}{2}$. If $\boldsymbol{\lambda}$ is Hermitian, then from Eq. (36) one finds that $\boldsymbol{\lambda} = -\frac{1}{2}\mathbf{u}$, or $\lambda_{ij} = -\frac{1}{2}$ for all i and j. In a few words, the generalized ONCV method (Eq. (33)) is reduced to the symmetric ONCV method (Eq. (6)) under the specific choice of λ_{ij}'s. This choice corresponds to one degree of freedom for the orthonormal vectors. Thus, it now becomes clear what physical meaning the third and fourth terms in Eq. (6) have; we have shown in what manner the vectors in the third and fourth terms, $\langle\delta\phi_j|\phi_i\rangle$ and $\langle\phi_j|\delta\phi_i\rangle$, should be added to the arbitrarily shifted vector $\delta\boldsymbol{\phi}$ in order to make $|\phi_i\rangle$ and $|\phi_j\rangle$ orthogonal and/or make each of them normal.

In the usual manner of derivation of an SCF equation, a term $-\sum\theta_{ji}\,\delta\langle\phi_j|\phi_i\rangle$ is added to $\delta\omega_{\mathrm{SCF}}$ in order to maintain the orthonormality condition. Instead of doing so, we use the generalized ONCV method and obtain the following equation:

$$\bar{\delta}_{\mathrm{SCF}} = 2\sum_i\left[\langle\delta\phi_i|\left(F_i - \sum_k|\phi_k\rangle\langle\phi_k|G_{ki}\right)|\phi_i\rangle + \text{H.c.}\right]. \quad (38)$$

Here we have defined

$$G_{ki} \equiv -(\lambda_{ik}^* F_i + \lambda_{ki}F_k), \quad (39)$$

and used the relationship

$$\sum_i \sum_k \lambda_{ik} \langle \phi_i | F_i | \phi_k \rangle \langle \delta\phi_k | \phi_i \rangle = \sum_i \sum_k \lambda_{ki} \langle \delta\phi_i | \phi_k \rangle \langle \phi_k | F_k | \phi_i \rangle . \tag{40}$$

If the λ_{ik}'s are real, then G_{ki} becomes a Hermitian operator. Though the operator in the parentheses in Eq. (38) depends upon suffix i, it can be unified into one operator with the use of the so-called coupling operators (Huzinaga, 1969). In defining such an operator, one should pay attention to the condition Eq. (35). The SCF variation equation $\bar{\delta}\omega_{\mathrm{SCF}} \rightarrow 0$ suggests the following methods for solving the SCF equation:

(1) Since $|\delta\phi_i\rangle$ is an arbitrary vector, one obtains from Eq. (38)

$$\Big(F_i - \sum_k |\phi_k\rangle \langle \phi_k | G_{ki}\Big) |\phi_i\rangle = 0, \tag{41}$$

or

$$\Big(F_i - \sum_{k \neq i} |\phi_k\rangle \langle \phi_k | G_{ki}\Big) |\phi_i\rangle = |\phi_i\rangle \langle \phi_i | F_i | \phi_i \rangle , \tag{42}$$

which is an exact SCF equation under the correct constraints of Eq. (35) and becomes a starting point of the coupling operator method. One can obtain a solution in an iterative manner by the diagonalization of the matrix representation.

(2) In Eq. (38), one can put

$$|\delta\phi_i\rangle = -\mu_i \Big(F_i - \sum_k |\phi_k\rangle \langle \phi_k | G_{ki}\Big) |\phi_i\rangle , \tag{43}$$

which is considered to be an extension of McWeeny's steepest descent method (McWeeny, 1956) if μ_i is defined as a positive number.

(3) Equation (38) directly gives a relationship between the SCF energy ω_{SCF} and its gradient. This is because $|\delta\phi_i\rangle$ in Eq. (38) is an arbitrary and infinitesimal variation vector. Note that the necessary orthonormal conditions are involved in grad($\bar{\delta}\omega_{\mathrm{SCF}}$). Therefore, Eq. (38) can be regarded as an extension of Fletcher's conjugate-gradient method (Fletcher, 1970).

III. Maximum and Minimum Orbitals

A. Maximum Overlap Orbitals

In the following, we will show that hybrid orbitals in molecules can be derived from an ONCV treatment using the squared overlap as a measure. As was mentioned in the Introduction, the principle of maximum overlapping serves as one of the most important concepts in obtaining information on bond strengths, the directional property of chemical bonds, hybridization of atomic orbitals, the geometrical structure of molecules, and other various atomic and molecular properties. Some aspects of these old, yet from another vantage new, subjects in quantum chemistry have been reviewed by Randić and Maksić (1972).

Our goal is to throw light upon the theoretical aspects of the maximum and minimum orbitals and to show that the variation equation finally leads to the required solution. Löwdin (1956) proposed a method of canonical orthonormalization as well as that of symmetric orthonormalization. The canonical orthonormalization is obtained by the multiplication of a unitary matrix U which diagonalizes the matrix $\boldsymbol{\theta}^\dagger\boldsymbol{\theta}$:

$$\boldsymbol{\theta}(\boldsymbol{\theta}^\dagger\boldsymbol{\theta})^{-1/2}\mathbf{U} = \mathbf{U}\mathbf{d}^{-1/2}\mathbf{U}^\dagger\mathbf{U} = \mathbf{U}\mathbf{d}^{-1/2} = \boldsymbol{\phi}_\mathrm{c}, \tag{44}$$

where **d** is a diagonal matrix. As Löwdin (1970) pointed out, if $\boldsymbol{\theta}$ is given as a set of atomic orbitals of a molecular or solid-state system, the set $\boldsymbol{\phi}_\mathrm{s} \equiv \boldsymbol{\theta}(\boldsymbol{\theta}^\dagger\boldsymbol{\theta})^{-1/2}$ is rather delocalized compared with $\boldsymbol{\theta}$, though still strongly localized, whereas the set $\boldsymbol{\theta}(\boldsymbol{\theta}^\dagger\boldsymbol{\theta})^{-1/2}\mathbf{U}$ will be much more delocalized over the entire system. It is true that both $\boldsymbol{\phi}_\mathrm{s}$ and $\boldsymbol{\phi}_\mathrm{c}$ are orthonormal vectors, but the latter will often be more advantageous than the former as a basis function for treating phenomena in solid-state systems. In later sections, we will show that solutions in the form of maximum overlap, localized, and hybrid orbitals are nothing but specific species of the Löwdin orthonormalization. In this section, a solution to $\bar{\delta}\omega = 0$, where ω is the measure of squared overlap, is shown to be given by an eigenvector of a certain eigenvalue equation. This is in contrast to the case of the symmetric orthonormalization in which the necessary equation is not an eigenvalue equation but an equation containing unknown vectors.

We start from the sum of the absolute squares of overlap integrals over the basis vector set $\boldsymbol{\chi} \equiv (|\chi_1\rangle\ |\chi_2\rangle \cdots |\chi_l\rangle)$ and an unknown vector $|\phi_i\rangle$:

$$\omega_i = \sum_{\mu=1}^{l} |\langle\chi_\mu|\phi_i\rangle|^2, \tag{45}$$

which was first introduced by Lykos and Schmeising (1961) when they studied maximum overlap orbitals. Using a projection operator $P_i \equiv |\phi_i\rangle\langle\phi_i|$ $(P_i^2 = P_i)$, one can rewrite Eq. (45) compactly as

$$\omega_i = \mathrm{tr}(\boldsymbol{\chi}^\dagger P_i \boldsymbol{\chi}). \tag{46}$$

Instead of Eqs. (45) and/or (46), we consider the measure which collects all ω_i $(i = 1, 2, \ldots, n)$:

$$\omega_\phi = \sum_i \omega_i = \mathrm{tr}(\boldsymbol{\chi}^\dagger P_\phi \boldsymbol{\chi}), \tag{47}$$

with

$$P_\phi = \boldsymbol{\phi}\boldsymbol{\phi}^\dagger = \sum_i |\phi_i\rangle\langle\phi_i|. \tag{48}$$

Using the symmetric ONCV, Eq. (4), one can obtain the following variational expression for the projection operator P_ϕ, provided that terms higher than second order in $\delta\boldsymbol{\phi}$ may be neglected:

$$P_\phi \rightarrow \bar{P}_\phi \equiv \boldsymbol{\phi}\boldsymbol{\phi}^\dagger + \delta\boldsymbol{\phi}^\dagger\boldsymbol{\phi} + \boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi} - \boldsymbol{\phi}(\delta\boldsymbol{\phi}^\dagger\boldsymbol{\phi} + \boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi})\boldsymbol{\phi}^\dagger. \tag{49}$$

A procedure in which Eq. (49) is used may be called an idempotency-constrained variation method for the variation of the projection operator P_ϕ (Morikawa and I'Haya, 1978b). Thus, one can readily write the variation of ω_ϕ as

$$\bar{\delta}\omega_\phi = \mathrm{tr}(\boldsymbol{\chi}^\dagger\bar{P}_\phi\boldsymbol{\chi}) = \mathrm{tr}(\boldsymbol{\chi}^\dagger P_\phi\boldsymbol{\chi}) = \mathrm{tr}\{\delta\boldsymbol{\phi}^\dagger(1 - \boldsymbol{\phi}\boldsymbol{\phi}^\dagger)P_\chi\boldsymbol{\phi} + \mathrm{H.c.}\}, \tag{50}$$

where

$$P_\chi = \boldsymbol{\chi}\boldsymbol{\chi}^\dagger = \sum_\mu |\chi_\mu\rangle\langle\chi_\mu|. \tag{51}$$

From $\bar{\delta}\omega_\phi = 0$, one obtains

$$P_\chi\boldsymbol{\phi} = \boldsymbol{\phi}\boldsymbol{\varepsilon}, \qquad \text{with} \quad \boldsymbol{\varepsilon} = \boldsymbol{\phi}^\dagger P_\chi\boldsymbol{\phi}, \tag{52}$$

which is a special case of the eigenvalue problem, Eq. (16). Similar to the discussion given in Section III,B, the unitary transformation does not disturb the formation of the eigenvectors, so that one has

$$P_\chi\boldsymbol{\phi} = \boldsymbol{\phi}\mathbf{d}_\varepsilon, \qquad \mathbf{d}_\varepsilon\text{: diagonal.} \tag{53}$$

Thus, the eigenvector belonging to the maximum $\mathbf{d}_i$ ($\mathbf{d}_i \equiv \langle\phi_i|P_\chi|\phi_i\rangle$) turns out to be a maximum overlap orbital.

Putting it another way, we consider the matrix representation of Eq. (53) based on the χ-space. If we set $\boldsymbol{\phi} = \boldsymbol{\chi}\mathbf{C}$ just as in linear combination of atomic orbitals, Eq. (53) becomes

$$\mathbf{S}^2\mathbf{C} = \mathbf{S}\mathbf{C}\mathbf{d}_\varepsilon, \tag{54}$$

where $\mathbf{S} = \boldsymbol{\chi}^\dagger\boldsymbol{\chi}$. A solution to this equation then is

$$\mathbf{C} = \mathbf{S}^{-1/2}\mathbf{C}' = \mathbf{C}'\mathbf{d}^{-1/2}, \tag{55}$$

where $\mathbf{C}'$ is a solution to the eigenvalue equation

$$\mathbf{S}\mathbf{C}' = \mathbf{C}'\mathbf{d}', \qquad \mathbf{d}' = \mathbf{C}'^\dagger\mathbf{S}\mathbf{C}' = \mathbf{d}. \tag{56}$$

The maximum overlap orbital can hence be taken as

$$\mathbf{C} = \mathbf{S}^{-1/2}\mathbf{C}' = \mathbf{C}'\mathbf{d}^{-1/2}. \tag{57}$$

$\mathbf{C}$ implies Löwdin's canonical orthonormalization, since $\boldsymbol{\phi}_c = \boldsymbol{\chi}\mathbf{C} = \boldsymbol{\chi}\mathbf{S}^{-1/2}\mathbf{C}' = \boldsymbol{\chi}\mathbf{C}'\mathbf{d}^{-1/2}$. The $|\phi_{ci}\rangle$, which corresponds to the maximum eigenvalue $\mathbf{d}_i$, becomes the maximum overlap orbital. The above-mentioned setting of $\boldsymbol{\phi} = \boldsymbol{\chi}\mathbf{C}$ also implies that we are to obtain a vector which spans over the space with $\boldsymbol{\chi}$ as a basis and is a solution to the equation $P_\chi\boldsymbol{\phi} = \boldsymbol{\phi}(\boldsymbol{\phi}^\dagger P_\chi\boldsymbol{\phi})$. The final result $\boldsymbol{\phi}_c = \boldsymbol{\chi}\mathbf{S}^{-1/2}\mathbf{C}'$ suggests that the transformation based on the e-space, where $\mathbf{e} = \boldsymbol{\chi}(\boldsymbol{\chi}^\dagger\boldsymbol{\chi})^{-1/2} = \boldsymbol{\chi}\mathbf{S}^{-1/2}$ (Löwdin's symmetrically orthonormalized vector), leads to $\boldsymbol{\phi}_c = \mathbf{e}\mathbf{C}'$ and $\mathbf{e}^\dagger P_\chi\mathbf{e} = \mathbf{S}$; i.e., if we express $P_\chi\boldsymbol{\phi} = \boldsymbol{\phi}(\boldsymbol{\phi}^\dagger P_\chi\boldsymbol{\phi})$ in a matrix form with $\mathbf{e}$ as a basis, we get the equation $\mathbf{S}\mathbf{C}' = \mathbf{C}'\mathbf{d}'$, from which we can easily obtain the solution.

On the other hand, the results obtained above can be derived by the direct application of the ONCV method instead of using $P_\phi \to \bar{P}_\phi$ as in Eq. (49). In that case, we rewrite Eq. (47) in terms of operator P_χ instead of P_ϕ:

$$\omega_\chi = \mathrm{tr}(\boldsymbol{\phi}^\dagger P_\chi\boldsymbol{\phi}), \tag{58}$$

and obtain results exactly the same as before from the equation $\bar{\delta}\omega_\chi = 0$ after applying the variation $\boldsymbol{\phi} \to \bar{\boldsymbol{\phi}}$ (Morikawa and I'Haya, 1978b).

B. Maximum Localized and Hybrid Orbitals

Since general aspects of variational treatment for maximum localized and hybrid orbitals are analogous to those for the maximum overlap orbitals mentioned in Section III,A, the former will be discussed here only briefly. Let the partitioned $\boldsymbol{\chi}$ be $(\boldsymbol{\chi}_A|\boldsymbol{\chi}_B|\cdots)$, where $\boldsymbol{\chi}_A$, $\boldsymbol{\chi}_B$, etc. represent a sequence of atomic orbitals centered on atom A, atom B, etc. To what extent an orbital $|\phi_i\rangle$ localizes over atom A will be measured by a real value

$$\omega_{Ai} \equiv \sum_{\mu\in A} |\langle\chi_\mu|\phi_i\rangle|^2 = \langle\phi_i|P_A|\phi_i\rangle, \tag{59}$$

where

$$P_A = \sum_{\mu\in A} |\chi_\mu\rangle\langle\chi_\mu| \tag{60}$$

is the projection operator over atom A.

Therefore, in a manner similar to Section III,A, we can obtain the following equation from the equation $\bar{\delta}\omega_{Ai} = 0$:

$$P_A|\phi_{Ai}\rangle = |\phi_{Ai}\rangle d_{Ai}, \tag{61}$$

which turns out to be a matrix form based on the $\boldsymbol{\chi}$-space,

$$\mathbf{S}_A\mathbf{S}_A{}^\dagger\mathbf{C}_{Ai} = \mathbf{S}\mathbf{C}_{Ai}\mathbf{d}_{Ai}, \tag{62}$$

or based on the e-space,

$$\mathbf{S}'_{\mathrm{AA}}\mathbf{C}'_{\mathrm{A}i} = \mathbf{C}'_{\mathrm{A}i}\mathbf{d}'_{\mathrm{A}i}. \tag{63}$$

Here, $\mathbf{S}_{\mathrm{A}} = \boldsymbol{\chi}^{\dagger}\boldsymbol{\chi}_{\mathrm{A}}$ (rectangular matrix), $|\phi_{\mathrm{A}i}\rangle = \boldsymbol{\chi}\mathbf{C}_{\mathrm{A}i}$, $\mathbf{C}'_{\mathrm{A}i} = \mathbf{S}^{1/2}\mathbf{C}_{\mathrm{A}i}$, and $\mathbf{S}'_{\mathrm{AA}} = \mathbf{S}^{-1/2}\mathbf{S}_{\mathrm{A}}\mathbf{S}^{\dagger}_{\mathrm{A}}\mathbf{S}^{-1/2}$. Thus the maximum localized orbital is given by $|\phi_{\mathrm{A}}\rangle = \boldsymbol{\chi}\mathbf{C}_{\mathrm{A}} = \boldsymbol{\chi}\mathbf{C}'^{\dagger}\mathbf{d}^{-1/2}\mathbf{C}'\mathbf{C}'_{\mathrm{A}}$ ($\mathbf{C}'_{\mathrm{A}}$ is a unitary matrix), which is a special case of Löwdin's orthonormalization. Note that $\mathbf{d}_{\mathrm{A}i}$ represents an extent of localization. If $|\phi_{\mathrm{A}i}\rangle$ is normalized, then an inequality $0 \leq \mathbf{d}_{\mathrm{A}i} \leq 1$ holds since $0 \leq \||P_{\mathrm{A}}|\phi_{\mathrm{A}i}\rangle\| \leq 1$. The $|\phi_{\mathrm{A}i}\rangle$ completely delocalizes on atom A when $\mathbf{d}_{\mathrm{A}i} = 0$, and the localization is maximal when $\mathbf{d}_{\mathrm{A}i} = 1$. The orthogonalized eigenvectors to be obtained are thus made to localize on a specific atom to any extent one wishes.

Next, we describe maximum hybrid orbitals, for which an overlap between the orbitals and the given hybrid orbitals $|\phi_i\rangle$ becomes maximum. Murrell (1960) and Golebiewski (1961) independently constructed approximate hybrid orbitals which maximize $\Sigma\langle\theta_i/\phi_i\rangle$ when the fixed ligand orbitals $|\phi_i\rangle$'s bind to a central atom, such as X in molecule XY_n. The value $\Sigma\langle\theta_i|\phi_i\rangle$, however, is not necessarily real and therefore is not a desirable measure of hybridization. The degree of the overlap is rather measured by

$$\sum_i |\langle\theta_i|\phi_i\rangle|^2, \tag{64}$$

which was suggested by Gilbert and Lykos (1961). Rewriting Eq. (64) as

$$\sum_i \langle\phi_i|P_{\theta i}|\phi_i\rangle \equiv \omega_\theta, \tag{65}$$

where

$$P_{\theta i} \equiv |\theta_i\rangle\langle\theta_i|, \tag{66}$$

one finds that Eq. (65) resembles the measure for the SCF equation mentioned in Section II,D. The only difference from that case is that $P_{\theta i}$ does not depend upon $|\phi_i\rangle$.

Application of the symmetric and/or the generalized ONCV methods to Eq. (65) yields

$$\delta\omega_\theta = \sum_i \left\{ \langle\delta\phi_i| \left(P_{\theta i} - \sum_k \phi_k\rangle\langle\phi_k|Q_{ki} \right) |\phi_i\rangle + \mathrm{H.c.} \right\}, \tag{67}$$

with

$$Q_{ki} \equiv -(\lambda^*_{ik}P_{\theta i} + \lambda_{ki}P_{\theta k}), \tag{68}$$

where λ_{ki}'s are arbitrary parameters among which the relationship Eq. (35) holds. Using the coupling operator in the SCF theory (Huzinaga, 1969), we can rewrite Eq. (68) into an operator that does not depend upon suffix i:

$$R_\theta|\phi_{\theta i}\rangle = |\phi_{\theta i}\rangle\varepsilon_{\theta i}, \qquad \varepsilon_{\theta i} \equiv \langle\phi_{\theta i}|R_\theta|\phi_{\theta i}\rangle. \tag{69}$$

Eq. (69) is a generalized version of the results obtained by Gilbert and Lykos (1961), and is similar in form to the general SCF equation. The maximum hybrid orbital is then obtained as $|\phi_i\rangle$, which corresponds to the maximum eigenvalues.

IV. Unitary Transformation and ONCV

As is well known, a linear transformation with the orthonormality of vectors preserved is a unitary transformation which is often utilized in studies of atoms and molecules. For example, the Hartree–Fock equation for a closed-shell system, Eq. (18), can be obtained, since the Hartree–Fock operator F is invariant under the unitary transformation between occupied molecular orbitals and so is the total electronic energy of the system. Using this degree of freedom, Edmiston and Ruedenberg (1963) maximized the orbital self-repulsion energy

$$\sum_j^{\text{occ}} \langle\phi_j|J_j|\phi_j\rangle \equiv \omega_{\text{L}} \tag{70}$$

and obtained localized orbitals. The usefulness of their localized orbitals was fully discussed by them. In this section, we will derive a method widely applicable to quantum chemical problems, choosing a variation of $\tilde{\boldsymbol{\phi}}$ so as to become a unitary transformation.

In the symmetric ONCV method, we have seen that $\delta\boldsymbol{\phi}$ can move freely and arbitrarily in the variational space under consideration. That the orthonormal $\boldsymbol{\phi}$ receives a variation under a unitary transformation implies that a vector set to be obtained after the variation can be expanded by the original vector set $\boldsymbol{\phi}$. Therefore, we will only have to confine $\delta\boldsymbol{\phi}$ to the $\boldsymbol{\phi}$-manifold, so that we substitute $\boldsymbol{\phi}\boldsymbol{\phi}^\dagger\delta\boldsymbol{\phi}$ instead of $\delta\boldsymbol{\phi}$ into $\tilde{\boldsymbol{\phi}}$ using the projection operator $\boldsymbol{\phi}\boldsymbol{\phi}^\dagger \equiv \sum|\phi_i\rangle\langle\phi_i|$. The result written as $\tilde{\boldsymbol{\phi}}$ is

$$\tilde{\boldsymbol{\phi}} \equiv \boldsymbol{\phi} - \tfrac{1}{2}\boldsymbol{\phi}(\delta\boldsymbol{\phi}^\dagger\boldsymbol{\phi} - \boldsymbol{\phi}^\dagger\,\delta\boldsymbol{\phi}) \tag{71}$$

or in component form:

$$|\tilde{\phi}_i\rangle \equiv |\phi_i\rangle - \tfrac{1}{2}\sum_k |\phi_k\rangle(\langle\delta\phi_k|\phi_i\rangle - \langle\phi_k|\delta\phi_i\rangle). \tag{72}$$

Note that the coefficient $-\frac{1}{2}$ may be replaced by any other real number. The process $\boldsymbol{\phi} \rightarrow \tilde{\boldsymbol{\phi}}$ is the variation method which meets the present purpose and may be called the unitary ONCV method. The process is proved to be a unitary transformation since one can write

$$\tilde{\boldsymbol{\phi}} = \boldsymbol{\phi}\mathbf{U}, \tag{73}$$

$$\mathbf{U} \equiv \mathbf{I} - \tfrac{1}{2}(\delta\boldsymbol{\phi}^{\dagger}\boldsymbol{\phi} - \boldsymbol{\phi}^{\dagger}\,\delta\boldsymbol{\phi}), \tag{74}$$

$$\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I}. \tag{75}$$

Application of Eq. (71) to $\tilde{\delta}\omega_{\mathrm{L}} \equiv \omega_{\mathrm{L}}(\tilde{\boldsymbol{\phi}}) - \omega_{\mathrm{L}}(\boldsymbol{\phi})$ yields

$$\tilde{\delta}\omega_{\mathrm{L}} = 2\sum_{i}\left\{\langle\delta\phi_i|\sum_{j}|\phi_j\rangle\langle\phi_j|\tfrac{1}{2}(J_i - J_j)|\phi_i\rangle\right\} + \mathrm{H.c.} \tag{76}$$

Since $|\delta\phi_i\rangle$ is arbitrary, the following necessary condition is obtained from $\tilde{\delta}\omega_{\mathrm{L}} = 0$:

$$\sum_{j}|\phi_j\rangle\langle\phi_j|\tfrac{1}{2}(J_i - J_j)|\phi_i\rangle = 0 \qquad (i = 1, 2, \ldots), \tag{77}$$

which gives $\langle\phi_j|\frac{1}{2}(J_i - J_j)|\phi_i\rangle = 0$, since ϕ_j's are linearly independent. This is the same as the result obtained by Edmiston and Ruedenberg (1963). The unitary ONCV method has been applied to various atomic and molecular systems (Morikawa and Narita, 1977).

V. Variational Method for a Pair of Basis Sets

A. Construction of Pair-Orthogonal Basis Sets

Instead of treating just one basis set as done previously, we expand our discussion on variation problems to include the treatment of a pair of basis sets. Examples in which such a pair of bases effectively work out are met often: α- and β-spin-orbital sets constructing a single Slater determinant in the unrestricted Hartree–Fock (UHF) theory, a basis and its reciprocal basis called biorthogonal bases and/or dual bases (Narita and I'Haya, 1974; Norbeck and McWeeny, 1975), etc.

Let us denote the two sets by row matrices

$$\boldsymbol{\psi} \equiv (|\psi_1\rangle \quad |\psi_2\rangle \quad \cdots), \qquad \boldsymbol{\xi} \equiv (|\xi_1\rangle \quad |\xi_2\rangle \quad \cdots), \tag{78}$$

and their adjoints by column matrices

$$\boldsymbol{\psi}^{\dagger} \equiv \begin{pmatrix} \langle\psi_1| \\ \langle\psi_2| \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \qquad \boldsymbol{\xi}^{\dagger} \equiv \begin{pmatrix} \langle\xi_1| \\ \langle\xi_2| \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}. \tag{79}$$

Here, we assume that the sets $\{|\psi_i\rangle\}$ and $\{|\xi_j\rangle\}$ are linearly independent, but not necessarily orthonormal. As mentioned in the preceding sections, a mutual relation between the basis vectors can be measured by the degree of overlapping, so that the so-called overlap matrices must be carefully investigated. The most important is a specific case that the overlap is zero, i.e., orthogonality. A unitary transformation among α- and β-spin orbitals in the UHF method leaves the Slater determinant unchanged except for its phase. Using this unitary invariance, one can eliminate off-diagonal terms of the overlap matrix between the orbital sets. This has been called the pairing theorem by Löwdin (1962) and the orbitals obtained in this manner have been called the corresponding orbitals by Amos and Hall (1961). We hereafter refer to orthogonality between two basis sets as *pair-orthogonality* in order to distinguish it from orthogonality between components of a basis set.

The biorthogonal sets $\boldsymbol{\psi}'$ and $\boldsymbol{\xi}'$ have the same dimension and are linked close together by the relations, $\boldsymbol{\psi}' = \boldsymbol{\theta}$, $\boldsymbol{\xi}' = \boldsymbol{\theta}(\boldsymbol{\theta}^{\dagger}\boldsymbol{\theta})^{-1}$, $\boldsymbol{\theta}$ being an arbitrary starting set, e.g., an infinitesimally shifted vector set. Further, they are normalized to unity, $\boldsymbol{\psi}'^{\dagger}\boldsymbol{\xi}' = \mathbf{I}$. On the other hand, two sets constituting pair-orthogonal vectors may have different dimensions and need not necessarily be normalized. In addition, each starting set need not be the same in the two sets; i.e., the pair-orthogonal sets are more generalized than the biorthogonal sets and contain the latter as a special case. Another extension of the biorthogonal sets is possible, where we can take a basis and its reciprocal basis as $\boldsymbol{\psi}' = \boldsymbol{\theta}(\boldsymbol{\theta}^{\dagger}\boldsymbol{\theta})^{-\lambda}$ and $\boldsymbol{\xi}' = \boldsymbol{\theta}(\boldsymbol{\theta}^{\dagger}\boldsymbol{\theta})^{\lambda-1}$, λ being an arbitrary real number (Narita and I'Haya, 1974). These may be called a dual basis and reduce to the biorthogonal basis under the condition $\lambda = 0$ or 1. In what follows, we discuss the properties of pair-orthogonal vectors from a variational point of view, and demonstrate the usefulness of the method of pair-orthogonality-constrained variation (Morikawa, 1977) that will be described.

For the sake of simplicity, we first deal with two orthonormal sets $\boldsymbol{\phi}_0$ and $\boldsymbol{\pi}_0$ as starting vectors: $\boldsymbol{\phi}_0 \equiv (|\phi_{01}\rangle\, |\phi_{02}\rangle \cdots)$, $\boldsymbol{\pi}_0 \equiv (|\pi_{01}\rangle\, |\pi_{02}\rangle \cdots)$, $\boldsymbol{\phi}_0^{\dagger}\boldsymbol{\phi}_0 = \mathbf{I}_{\phi}$, and $\boldsymbol{\pi}_0^{\dagger}\boldsymbol{\pi}_0 = \mathbf{I}_{\pi}$. We wish to find out two vector sets $\boldsymbol{\phi}_{\mathrm{p}}$ and $\boldsymbol{\pi}_{\mathrm{p}}$ which have the properties of the orthonormality, $\boldsymbol{\phi}_{\mathrm{p}}^{\dagger}\boldsymbol{\phi}_{\mathrm{p}} = \mathbf{I}_{\phi}$, $\boldsymbol{\pi}_{\mathrm{p}}^{\dagger}\boldsymbol{\pi}_{\mathrm{p}} = \mathbf{I}_{\pi}$, and of the pair-orthogonality, $\langle \pi_{\mathrm{p}\,i}|\phi_{\mathrm{p}\,j}\rangle = D_{ii}\delta_{ij}$. If $\boldsymbol{\phi}_0$ is orthonormal, then so is $\boldsymbol{\phi}_0\mathbf{A}$, $\mathbf{A}$ being a unitary matrix; $(\boldsymbol{\phi}_0\mathbf{A})^{\dagger}(\boldsymbol{\phi}_0\mathbf{A}) = \mathbf{A}^{\dagger}\mathbf{A} = \mathbf{I}_{\phi}$. Similarly, $(\boldsymbol{\pi}_0\mathbf{B}^{\dagger})(\boldsymbol{\pi}_0\mathbf{B}) = \mathbf{B}^{\dagger}\mathbf{B} = \mathbf{I}_{\pi}$, $\mathbf{B}$ being another unitary matrix; i.e., an orthonormal vector set has the degree of freedom of unitary transformation. Using this freedom, we can calculate the pair-orthogonal sets $\boldsymbol{\phi}_{\mathrm{p}}$ and $\boldsymbol{\pi}_{\mathrm{p}}$ from the starting basis sets $\boldsymbol{\phi}_0$ and $\boldsymbol{\pi}_0$.

Two manifolds spanned by $\boldsymbol{\phi}_0$ and $\boldsymbol{\pi}_0$ are characterized by projection operators $P_{\phi_0} = \boldsymbol{\phi}_0\boldsymbol{\phi}_0^{\dagger}$ and $P_{\pi_0} = \boldsymbol{\pi}_0\boldsymbol{\pi}_0^{\dagger}$, respectively. Expressing one of the unknown sets to be obtained as $\boldsymbol{\phi}$, then $P_{\phi_0}\boldsymbol{\phi}$ turns out to be a projection over the manifold $\boldsymbol{\phi}_0$, and further $P_{\pi_0}P_{\phi_0}\boldsymbol{\phi}$ is a projection of $P_{\phi_0}\boldsymbol{\phi}$ over the manifold $\boldsymbol{\pi}_0$. We now deal variationally with the sum of squared overlaps (squared norms) of $P_{\pi_0}P_{\phi_0}\boldsymbol{\phi}$,

$$\sum_i \|P_{\pi_0}P_{\phi_0}\boldsymbol{\phi}_i\|^2 = \operatorname{tr}(\boldsymbol{\phi}^\dagger P_{\phi_0}P_{\pi_0}P_{\phi_0}\boldsymbol{\phi}) \equiv \omega_{\phi_0}(\boldsymbol{\phi}), \tag{80}$$

under the orthonormality condition $\boldsymbol{\phi}^\dagger\boldsymbol{\phi} = \mathbf{I}$. Application of the symmetric ONCV method to Eq. (80) leads directly to

$$\bar{\delta}\omega_{\phi_0}(\boldsymbol{\phi}) = \operatorname{tr}[\delta\boldsymbol{\phi}^\dagger(1 - \boldsymbol{\phi}\boldsymbol{\phi}^\dagger)P_{\phi_0}P_{\pi_0}P_{\phi_0}\boldsymbol{\phi} + \text{H.c.}]. \tag{81}$$

Solving the equation $\bar{\delta}\omega_{\phi_0}(\boldsymbol{\phi}) = 0$, one obtains

$$P_{\phi_0}P_{\pi_0}P_{\phi_0}\boldsymbol{\phi} = \boldsymbol{\phi}(\boldsymbol{\phi}^\dagger P_{\phi_0}P_{\pi_0}P_{\phi_0}\boldsymbol{\phi}). \tag{82}$$

Substituting $\boldsymbol{\phi} = \boldsymbol{\phi}_0\mathbf{A}$ and multiplying both sides of Eq. (92) by $\boldsymbol{\phi}_0^\dagger$, one has

$$(\boldsymbol{\phi}_0^\dagger\boldsymbol{\pi}_0\boldsymbol{\pi}_0^\dagger\boldsymbol{\phi}_0)\mathbf{A} = \mathbf{A}(\mathbf{D}^\dagger\mathbf{D}), \tag{83}$$

where $\mathbf{D} = \boldsymbol{\pi}_\mathrm{p}^\dagger\boldsymbol{\phi}_\mathrm{p} = \mathbf{B}^\dagger\boldsymbol{\pi}_0^\dagger\boldsymbol{\phi}_0\mathbf{A}$. Equation (82) is nothing else but an eigenvalue equation whose eigenmatrix is $\mathbf{A}$. Similarly, it follows from $\omega_{\pi_0}(\boldsymbol{\pi}) \equiv \operatorname{tr}(\boldsymbol{\pi}^\dagger P_{\pi_0}P_{\phi_0}P_{\pi_0}\boldsymbol{\pi})$ and $\bar{\delta}\omega_{\pi_0}(\boldsymbol{\pi}) = 0$ that

$$P_{\pi_0}P_{\phi_0}P_{\pi_0}\boldsymbol{\pi} = \boldsymbol{\pi}(\boldsymbol{\pi}^\dagger P_{\pi_0}P_{\phi_0}P_{\pi_0}\boldsymbol{\pi}) \tag{84}$$

and the corresponding eigenvalue equation

$$(\boldsymbol{\pi}_0^\dagger\boldsymbol{\phi}_0\boldsymbol{\phi}_0^\dagger\boldsymbol{\pi}_0)\mathbf{B} = \mathbf{B}(\mathbf{D}\mathbf{D}^\dagger). \tag{85}$$

Note that *the order of occurrence of the projection operators is different in* Eqs. (83) and (85).

If the unitary matrices $\mathbf{A}$ and $\mathbf{B}$ diagonalize $\boldsymbol{\phi}_0^\dagger\boldsymbol{\pi}_0\boldsymbol{\pi}_0^\dagger\boldsymbol{\phi}_0$ and $\boldsymbol{\pi}_0^\dagger\boldsymbol{\phi}_0\boldsymbol{\phi}_0^\dagger\boldsymbol{\pi}_0$ respectively, then $\boldsymbol{\phi}_\mathrm{p} = \boldsymbol{\phi}_0\mathbf{A}$ and $\boldsymbol{\pi}_\mathrm{p} = \boldsymbol{\pi}_0\mathbf{B}$ definitely come out to be pair-orthogonal sets, because

$$\mathbf{A}^\dagger\boldsymbol{\phi}_0^\dagger\boldsymbol{\pi}_0\boldsymbol{\pi}_0^\dagger\boldsymbol{\phi}_0\mathbf{A} = \mathbf{D}^\dagger\mathbf{D}(\text{diagonal}) = \boldsymbol{\phi}_\mathrm{p}^\dagger\boldsymbol{\pi}_\mathrm{p}\boldsymbol{\pi}_\mathrm{p}^\dagger\boldsymbol{\phi}_\mathrm{p} \tag{86}$$

$$\mathbf{B}^\dagger\boldsymbol{\pi}_0^\dagger\boldsymbol{\phi}_0\boldsymbol{\phi}_0^\dagger\boldsymbol{\pi}_0\mathbf{B} = \mathbf{D}\mathbf{D}^\dagger(\text{diagonal}) = \boldsymbol{\pi}_\mathrm{p}^\dagger\boldsymbol{\phi}_\mathrm{p}\boldsymbol{\phi}_\mathrm{p}^\dagger\boldsymbol{\pi}_\mathrm{p}. \tag{87}$$

Equations (83) and (85) have been obtained by several ways of derivation which are different from the above (Amos and Hall, 1961; King *et al.*, 1967; Jankowski, 1976). Both equations are quite within the bounds of possibility that one can evaluate pair-orthogonal vectors. The operators $P_{\phi_0}P_{\pi_0}P_{\phi_0}$ and $P_{\pi_0}P_{\phi_0}P_{\pi_0}$ are considered to be a sort of the outer- and inner-projection operators discussed in detail by Löwdin (1965). The pair-orthogonal vectors thus have the properties characterized by these two projections.

We can now construct pair-orthogonal sets $\boldsymbol{\psi}_\mathrm{p}$ and $\boldsymbol{\xi}_\mathrm{p}$ from the original sets $\boldsymbol{\psi}$ and $\boldsymbol{\xi}$. First, with the use of the Löwdin orthonormalization, construct orthonormal vector sets $\psi(\psi^\dagger\psi)^{-1/2}$ and $\xi(\xi^\dagger\xi)^{-1/2}$. Then, the manifolds spanned by ψ and ξ are characterized, respectively, by the projection operators

$$P_\psi = \psi(\psi^\dagger\psi)^{-1/2}\{\psi(\psi^\dagger\psi)^{-1/2}\}^\dagger = \psi(\psi^\dagger\psi)^{-1}\psi^\dagger \tag{88}$$

$$P_\xi = \xi(\xi^\dagger\xi)^{-1/2}\{\xi(\xi^\dagger\xi)^{-1/2}\}^\dagger = \xi(\xi^\dagger\xi)^{-1}\xi^\dagger. \tag{89}$$

Substituting P_ψ and P_ξ for $\phi_0\phi_0^\dagger$ and $\pi_0\pi_0^\dagger$ in Eqs. (83)–(87), one has the corresponding equations for ψ_p and ξ_p. When ψ and ξ are orthonormal, then ψ_p and ξ_p reduce to ϕ_p and π_p respectively.

B. Pair-Orthogonality-Constrained Variation Method

Finally, we propose a method for constraining pair-orthogonality in the course of variation:

$$\{\psi, \xi\} \rightarrow \{\bar{\psi}, \bar{\xi}\}, \tag{90}$$

with

$$\bar{\psi} \equiv \psi + \delta\psi - \tfrac{1}{2}\xi_R(\delta\xi^\dagger\psi + \xi^\dagger\,\delta\psi), \tag{91}$$

$$\bar{\xi} \equiv \xi + \delta\xi - \tfrac{1}{2}\psi_R(\delta\psi^\dagger\xi + \psi^\dagger\,\delta\xi). \tag{92}$$

Here, $\psi_R = \psi(\psi^\dagger\psi)^{-1}$ and $\xi_R = \xi(\xi^\dagger\xi)^{-1}$ are the reciprocals with respect to ψ and ξ, respectively. In this method, direct calculation shows that the equality $\bar{\xi}^\dagger\bar{\psi} = \xi^\dagger\psi$ exists, so that the mixed overlap matrix set up initially is preserved in every step of variation. When one chooses the pair-orthogonal vectors $\{\psi_p, \xi_p\}$ as starting sets, then Eqs. (91) and (92) may be called the pair-orthogonality-constrained variation (POCV) method:

$$\{\psi_p, \xi_p\} \rightarrow \{\bar{\psi}_p, \bar{\xi}_p\}. \tag{93}$$

If one uses the method of constructing ϕ_p and π_p twice, one obtains an alternative form of the POCV method:

$$\{\phi_p, \pi_p\} \rightarrow \{\bar{\phi}_p, \bar{\pi}_p\}, \tag{94}$$

where $\bar{\phi}_p = \bar{\phi}\mathbf{A}$, $\bar{\pi}_p = \bar{\pi}\mathbf{B}$, etc. In contrast to the procedure shown in Eq. (93), this procedure preserves the orthonormality condition. Note, however, that the mixed overlap matrix on the right-hand side is different from that on the left-hand side.

REFERENCES

Amos, A. J., and Hall, G. G. (1961). *Proc. R. Soc. London, Ser. A* **263**, 483.
Birss, F. W., and Fraga, S. (1963). *J. Chem. Phys.* **38**, 2552.
Edmiston, C., and Ruedenberg, K. (1963). *Rev. Mod. Phys.* **35**, 457.

Epstein, S. T. (1974). "The Variation Method in Quantum Chemistry." Academic Press, New York.

Fletcher, R. (1970). *Mol. Phys.* **19**, 55.

Gilbert, T. L., and Lykos, P. G. (1961). *J. Chem. Phys.* **34**, 2199.

Goddard, W. A., III, Dunning, T. H., Jr., and Hunt, W. J. (1969). *Chem. Phys. Lett.* **4**, 231.

Golebiewski, A. (1961). *Trans. Faraday Soc.* **57**, 1849.

Gould, S. H. (1966). "Variational Methods for Eigenvalue Problems," and references cited therein. Univ. of Toronto Press, Toronto.

Huzinaga, S. (1969). *J. Chem. Phys.* **51**, 3971.

Jankowski, K. (1976). *Int. J. Quantum Chem.* **10**, 683.

King, H. F., Stanton, R. E., Kim, H., Wyatt, R. E., and Parr, R. G. (1967). *J. Chem. Phys.* **47**, 1936.

Landshoff, R. (1936). *Z. Phys.* **102**, 201.

Löwdin, P. O. (1947). *Ark. Mat. Astron. Fys.* **35A**, No. 9, 10.

Löwdin, P. O. (1950). *J. Chem. Phys.* **18**, 365.

Löwdin, P. O. (1956). *Adv. Phys.* **5**, 1.

Löwdin, P. O. (1962). *J. Appl. Phys.* **33**, Suppl., 251.

Löwdin, P. O. (1965). *Phys. Rev.* **139**, A357.

Löwdin, P. O. (1970). *Adv. Quantum Chem.* **5**, 185.

Lykos, P. G., and Schmeising, H. N. (1961). *J. Chem. Phys.* **35**, 288.

McWeeny, R. (1956). *Proc. R. Soc. London, Ser. A* **235**, 496.

Morikawa, T. (1972). *Chem. Phys. Lett.* **17**, 297.

Morikawa, T. (1975a). *Int. J. Quantum Chem.* **9**, 969.

Morikawa, T. (1975b). *Chem. Phys. Lett.* **32**, 521.

Morikawa, T. (1977). *J. Chem. Phys.* **67**, 128.

Morikawa, T., and I'Haya, Y. J. (1978a). *J. Chem. Phys.* **51**, 1821.

Morikawa, T., and I'Haya, Y. J. (1978b). *Int. J. Quantum Chem.* **13**, 199.

Morikawa, T., and Narita, S. (1977). *Phys. Lett. A* **62**, 13.

Murrell, J. N. (1960). *J. Chem. Phys.* **32**, 767.

Narita, S., and I'Haya, Y. J. (1974). *Int. J. Quantum Chem.* **8**, 765.

Norbeck, J. M., and McWeeny, R. (1975). *Chem. Phys. Lett.* **34**, 206.

Pauling, L. (1931). *J. Am. Chem. Soc.* **53**, 1367.

Pople, J. A., and Nesbet, R. K. (1954). *J. Chem. Phys.* **22**, 571.

Randić, M., and Maksić, Z. B. (1971). *Chem. Rev.* **72**, 43.

Roothaan, C. C. J. (1951). *Rev. Mod. Phys.* **23**, 69.

Roothaan, C. C. J. (1960). *Rev. Mod. Phys.* **32**, 179.

Slater, J. C. (1931). *Phys. Rev.* **37**, 431.

Slater, J. C., and Koster, G. F. (1954). *Phys. Rev.* **94**, 1498.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 12

Quantum Mechanical Methods for Regular Polymers

JEAN-MARIE ANDRÉ
Laboratoire de Chimie Théorique Appliquée
Facultés Universitaires N.D. de la Paix
Namur, Belgium

I. Introduction

Over the last decade, the quantum theory for polymers and molecular crystals has witnessed a growing interest directly related to the appealing properties (their conductivity, phase transitions, etc.) displayed by one-dimensional unsaturated systems (Keller, 1977; Pal *et al.*, 1977). These are expected to play a major role in the development of organic conductors, e.g., TTF–TCNQ and similar charge complexes, in the discovery of new types of superconductors such as $(SN)_x$ (Street and Greene, 1977) and possible superconductors at room temperature driven by an excitonic mechanism (Little, 1964; Davis *et al.*, 1976), and also in biological processes (Ladik, 1973, 1978). The range of applications of this quantum theory is not at all restricted to such unusual materials; as a matter of fact, many interesting and puzzling questions are still open even in the case of most common stereoregular polymers. For all these materials, and in the same way as for simpler molecules, quantum chemistry is now starting to provide powerful instruments for interpreting and predicting the physical and chemical properties (structure, bonding, reactivity, etc.) for which a rather detailed knowledge of the electronic structure is required.

ISBN 0-12-034812-8

The word "polymer" embraces a large class of complex and diverse compounds among which the quantum theory at present considers only a rather restricted subset. The main limitation comes obviously from the ability of the model to comprise the physics and chemistry of what is actually a "regular" polymer. Fortunately, there are now systems whose structural and physical features allow a reasonable modeling by isolated infinite and regular chains. This legitimizes the search for the full potential of quantum theory in the physics and chemistry of polymers in much the same way as quantum chemistry is applied to gaseous molecules and solid-state physics to three-dimensional crystals. Using stereospecific catalysts, polymer chemists are now able to synthesize chains of high stereoregularity (polyethylene, polypropylene, etc.) whereas nicely organized polymers (polydiacetylene, polyoxymethylene, etc.) are obtained through solid-state polymerization. These few examples do not exhaust the list of possible candidates for quantum investigations but their well-defined structures illustrate well enough the fact that theoretical models will not be far removed from reality and may conceivably bring a valuable contribution to polymer science.

Another reason for developing quantum methods for polymers is the increasing use of refined techniques (XPS, IR, NMR, etc.) to solve fundamental polymer problems; e.g., XPS (or ESCA—electron spectroscopy for chemical analysis) spectra obtained with monochromatic X-ray radiation turn out to be powerful tools with which to study the electronic structure of polymers, both by providing precise information on core level binding energies and line shapes and by allowing the recording of the distribution of the valence electronic levels that might constitute a real and unique "fingerprint" of the polymer.

This proclaims the need of theoretical methods for polymers in order to supply the necessary background for interpreting more and more complex experimental data.

From a conceptual point of view, it appears that polymer quantum chemistry is an ideal field for cooperation between solid-state physicists and molecular quantum chemists. There exists a common interpretation in discussions concerning orbital energies, orbital symmetry, and gross charges by chemists and solid-state physicists who use less familiar terms like first Brillouin zone, dependence of "wave function" (in fact the monoelectronic wave function is called an orbital by the chemist) with respect to wave vector k, Fermi surfaces, Fermi contours, and density of states.

Actually, both scientists use the same quantum mechanical background, the same approximations, and even the same degree of sophistication in order to understand one another. Thus an interesting feature of polymer theoretical chemistry is that its scientific language is partly that of quantum chemistry and partly that of solid-state physics. For example, model polymers are described by three-dimensional LCAO orbitals with only one direction of periodicity so that usual concepts of solid-state physics such as Bloch functions, Brillouin zones, and energy bands can be applied. Consequently, this field could have a considerable impact in teaching solid-state terms (such as first Brillouin zones, energy bands, or density of states) to chemists with a certain knowledge of basic quantum chemistry. However, polymer quantum chemistry is not, strictly speaking, one-dimensional physics; the three-dimensionality of the basic Bloch orbitals renders important theorems for a strictly one-dimensional system invalid (McCubbin and Teemull, 1972). A now classical example is the linear zigzag polyethylene whose energy bands exhibit extrema at points of the first Brillouin zone which are not high symmetry ones.

In this chapter, we summarize the general LCAO theory of polymers while emphasizing some practical aspects which are very important for a meaningful comparison with experiment. These are band ordering and accurate calculations of long-range electrostatic effects. In the last part, we show by specific examples the actual ability of polymer quantum chemistry to interpret and predict experimental results.

II. LCAO Methods for Band Structure Calculations of Polymers

A. General Formalism

The aim of this part is to present the general LCAO theory, as presently used, for studying the electronic properties of regular macromolecules. Since bonds in a polymer are not different from usual bonds in organic molecules, the methods already proved successful in the investigation of the electronic structure of molecules can be used for describing the electronic properties of polymers. Therefore, the one-electron energy band model has been and is still one of the most used approaches to the description of the electronic states in solids and stereoregular polymers.

If one uses a molecular approach, electronic states of large chains are represented by drawing the orbital energies (filled and empty) associated with chains of increasing length. In a series of homogeneous molecules, the number of energy levels increases with the molecular size and, correspondingly, the distance between these energy levels vanishes.

For the infinite regular chain, another description can be given. Due to the one-dimensional periodicity, the monoelectronic levels (orbital energies in the molecular description) may be represented as a multivalued function of a reciprocal wavenumber defined in the inverse space dimension. The set of all those branches (energy bands) plotted versus the reciprocal wave number (k-point) in a well-defined region of the reciprocal space (first Brillouin zone) is the band structure of the polymers. In the usual terminology, we note the analogy between the occupied levels and the valence bands, the unoccupied levels and the conduction band.

The deduction of the restricted Hartree–Fock equations for infinite regular polymers is well described in the literature (Ladik, 1965, 1975, 1978; Del Re *et al.*, 1967; André *et al.*, 1967a,b, 1971a; André, 1969, 1970, 1975; André and Delhalle, 1978). However, for sake of clarity, we find it appropriate to begin with a brief summary of the key LCAO–CO (CO—crystalline orbital) relations. Moreover, a recent appraisal of long-range contributions (Delhalle *et al.*, 1978a; Piela and Delhalle, 1978a,b; André *et al.*, 1978), typical of systems built of infinitely repeating patterns and generally not included in actual polymer calculations, makes this survey opportune.

Let us consider a one-dimensional periodic lattice consisting of a large number, $2\mathcal{N} + 1$ ($\mathcal{N} \rightarrow \infty$), of unit cells of length a, each containing Ω nuclei at positions $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\Omega$ relative to the cells origins $ja\hat{z}$, $j = 0, \pm 1, \pm 2, \ldots, \pm\mathcal{N}$, and $2n_0$ electrons distributed along the chain. A strict electroneutrality of the cells ($2n_0 = Z_1 + Z_2 + \cdots + Z_\Omega$) is necessary. The electrons are assumed to doubly occupy a set of one-electron orbitals, $\phi_n(k, \mathbf{r})$, of Bloch type written as periodic combinations of ω basis functions, $\chi_p(\mathbf{r})$.

$$\phi_n(k, \mathbf{r}) = (2\mathcal{N} + 1)^{-1/2} \sum_{j=-\mathcal{N}}^{+\mathcal{N}} \sum_{p=1}^{\omega} C_{np}(k) e^{ikja} \chi_p(\mathbf{r} - \mathbf{P} - ja\hat{z}), \qquad (1)$$

where $\mathbf{r}$ is the position vector measured from an arbitrary but fixed origin, $\hat{z}$ is a unit vector in the direction of lattice periodicity, $\mathbf{P}$ is the center of the basis function $\chi_p(\mathbf{r})$, and k is a point in the first Brillouin zone (BZ), $[-\pi/a, +\pi/a]$, of the polymer; L_{RC}, the length of this Brillouin zone, is equal to $2\pi/a$; ω is the size of the basis set.

The $\phi_n(k, \mathbf{r})$'s represent the wave function of a single electron in the periodic potential created by the nuclei and the other electrons. The optimal set of those polymer orbitals for a given atomic basis set is obtained in the usual way by solving the Hartree–Fock equations. The self-consistent field (SCF) monoelectronic operator has the explicit form

$$F(\mathbf{r}_i) = -\tfrac{1}{2}\,\nabla^2(\mathbf{r}_i) - \sum_{h=-\mathcal{N}}^{+\mathcal{N}} \sum_{u=1}^{\Omega} Z_u |\mathbf{r}_i - \mathbf{A}_u - ha\hat{z}|^{-1}$$

$$+ \sum_{n'=1}^{n_0} \sum_{k'} \{2J_{n'k'}(\mathbf{r}_i) - K_{n'k'}(\mathbf{r}_i)\}, \qquad (2)$$

where n_0 is the number of doubly occupied bands. The Coulombic and exchange terms are detailed as

$$J_{n'k'}(\mathbf{r}_i)\phi_n(k, \mathbf{r}_i) = \int \phi^*_{n'}(k', \mathbf{r}_j) r_{ij}^{-1} \phi_{n'}(k, \mathbf{r}_j)\, dv_j \phi_n(k', \mathbf{r}_i), \tag{3}$$

$$K_{n'k'}(\mathbf{r}_i)\phi_n(k, \mathbf{r}_i) = \int \phi^*_{n'}(k', \mathbf{r}_j) r_{ij}^{-1} \phi_n(k, \mathbf{r}_j)\, dv_j \phi_{n'}(k', \mathbf{r}_i). \tag{4}$$

The terms included in Eq. (2) are, respectively, the kinetic operator, the attraction of a single electron with all nuclei centered in all cells, the averaged electrostatic potential of all electrons, and the averaged exchange interaction.

As is the case for complicated problems where expansions into known functions are required, the final computational expressions are cast in matrix form. By forming the expectation value of the monoelectronic Hartree–Fock operator and by applying the variational procedure for the LCAO coefficients, $C_{np}(k)$, the following system of equations, of size ω, is obtained:

$$\sum_p C_{np}(k) \left(\sum_j e^{ikja}(F^j_{pq} - E_n(k) S^j_{pq}) \right) = 0. \tag{5}$$

The compatibility condition of this system

$$\left| \sum_j e^{ikja}(F^j_{pq} - E(k) S^j_{pq}) \right| \equiv 0 \tag{6}$$

gives the band structure, $\{E_n(k)\}$, as a multivalued function of k in the reduced zone scheme. We note that F^j_{pq} is a matrix element of the monoelectronic operator, $F(\mathbf{r}_i)$, between the atomic orbital χ_p centered in the origin cell and the atomic orbital χ_q centered in cell j. S^j_{pq}, an overlap integral, has the same meaning for the unit operator. Both matrix elements decrease exponentially with the distance between the orbitals, giving rise to a natural convergence of the summation over cells appearing in the secular systems (5) and determinants (6).

In the Hartree–Fock self-consistent procedure, the matrix elements F^j_{pq} have the general form:

$$F^j_{pq} = T^j_{pq} - \sum_h \sum_u Z_u V^j_{pq}(h, \mathbf{A}_u) + \sum_h \sum_l \sum_r \sum_s D^{hl}_{rs}\left[\left(2\,{}^{oj}_{pq}\middle|{}^{hl}_{rs}\right) - \left({}^{oh}_{pr}\middle|{}^{jl}_{qs}\right)\right] \tag{7}$$

when defining:

the kinetic integrals

$$T^j_{pq} = -\tfrac{1}{2}\int \chi_p(\mathbf{r} - \mathbf{P})\; \nabla^2(\mathbf{r}) \chi_q(\mathbf{r} - \mathbf{Q} - ja\hat{z})\, dv, \tag{8}$$

the nuclear attraction integrals

$$V^j_{pq}(h, \mathbf{A}_u) = \int \chi_p(\mathbf{r} - \mathbf{P}) |\mathbf{r} - \mathbf{A}_u - ha\hat{z}|^{-1} \chi_q(\mathbf{r} - \mathbf{Q} - ja\hat{z})\, dv, \tag{9}$$

the electron repulsion integrals

$$\left(\begin{smallmatrix}0j\\ pq\end{smallmatrix}\middle|\begin{smallmatrix}hl\\ rs\end{smallmatrix}\right) = \iint \chi_p(\mathbf{r}_1 - \mathbf{P})\chi_q(\mathbf{r}_1 - \mathbf{Q} - ja\hat{z})\mathbf{r}_{12}^{-1} \times \chi_4(\mathbf{r}_2 - \mathbf{R} - ha\hat{z})\chi_s(\mathbf{r}_2 - \mathbf{S} - la\hat{z})\, dv_1\, dv_2, \tag{10}$$

In these expressions and from now on (j, h, l), (u, v), and (p, q, r, s) indices will refer, respectively, to summations over cells, nuclei, and basis functions.

The iterative part of the computation involves elements of the so-called density matrix, D^j_{pq}, which are computed at each iteration by numerical integration over the first Brillouin zone of the polymer (Delhalle, 1972, 1974a; Delhalle and Delhalle, 1975):

$$D^j_{pq} = \sum_{k,n} C^*_{np}(k)C_{nq}(k)e^{ikja} = L^{-1}_{\mathrm{RC}} \int_{n=1}^{n_0} \sum C^*_{np}(k)C_{nq}(k)e^{ikja}\, dk \tag{11}$$

$$= L^{-1}_{\mathrm{RC}} \int D_{pq}(k)e^{ikja}\, dk. \tag{12}$$

The total energy per unit cell of the polymer, electronic plus nuclear, is obtained in the following way:

$$E_{\mathrm{T}} = \sum_j \sum_p \sum_q D^j_{pq}(F^j_{pq} + I^j_{pq}) + \sum_h{}' \sum_v \sum_u Z_v Z_u |\mathbf{A}_v - (\mathbf{A}_u + ha\hat{z})|^{-1}. \tag{13}$$

I^j_{pq} collects the one-electron terms

$$I^j_{pq} = T^j_{pq} - \sum_h \sum_u Z_u V^j_{pq}(h, \mathbf{A}_u). \tag{14}$$

The prime on summations excludes terms whose denominator vanishes. The above expressions are general.

In every LCAO *ab initio* procedure of the single determinant quality, Fock matrix elements thus consist of three physically different terms: a kinetic contribution, an electrostatic attraction and repulsion contribution, and an exchange part.

The kinetic term

$$T^j_{pq} = \int \chi_p(\mathbf{r} - \mathbf{P})\{-\tfrac{1}{2}\, \nabla^2(\mathbf{r})\}\chi_q(\mathbf{r} - \mathbf{Q} - ja\hat{z})\, dv \tag{15}$$

decays almost exponentially with distance between orbital $\chi_p(\mathbf{r} - \mathbf{P})$, centered in cell o, and orbital $\chi_q(\mathbf{r} - \mathbf{Q} - ja\hat{z})$, centered in cell j. Recalling the probability density interpretation of the wave function in quantum theory, the orbital product

$$P^j_{pq}(\mathbf{r}) = \chi_p(\mathbf{r} - \mathbf{P})\chi_q(\mathbf{r} - \mathbf{Q} - ja\hat{z}) \tag{16}$$

is the probability density of an electron shared by the orbitals $\chi_p(\mathbf{r} - \mathbf{P})$ and $\chi_q(\mathbf{r} - \mathbf{Q} - ja\hat{z})$. The kinetic integral can be interpreted as the mean kinetic energy of the charge distribution $P^j_{pq}(\mathbf{r})$. The explicit dependance, however, is not straightforward because of the particular nature of the operator ∇^2. The kinetic energy integrals are usually small when the overlap involving the same orbitals is small, thus producing a fast decrease with respect to the distance between orbital centers.

The exchange contributions

$$X^j_{pq} = -\sum_h \sum_l \sum_r \sum_s D^{hl}_{rs} \left({}^{oh}_{pq} \middle| {}^{jl}_{rs} \right) \tag{17}$$

correct for the electron self-repulsion included in the Coulomb electrostatic potential and also include the effect of the Pauli principle on the independent electron model. Clearly, if the centers of two functions χ_p and χ_q are widely separated, the exponential form of the basis orbitals ensures that the product $P^j_{pq}(\mathbf{r})$ will be small everywhere so that the leading terms in the exchange contributions to a matrix element are:

$$X^j_{pq} \simeq -\sum_r \sum_s D^j_{rs} \left({}^{oo}_{pr} \middle| {}^{jj}_{qs} \right) \simeq -D^j_{pq} \left({}^{oo}_{pp} \middle| {}^{jj}_{qq} \right). \tag{18}$$

It is to be pointed out, however, that exchange contributions can be important for those matrix elements between widely separated orbitals since the decay of both two-center repulsion integrals $\left({}^{oo}_{pp} \middle| {}^{jj}_{qq} \right)$ and elements of density matrices is very slow ($\sim r^{-1}$) with distance and not exponential as for other contributions. This results in a global r^{-2} dependence of exchange quantities.

The electrostatic attraction and repulsion terms are very important; they have the explicit form:

$$C^j_{pq} = \sum_h \left(-\sum_u Z_u V^j_{pq}(\mathbf{h}, \mathbf{A}_u) + \sum_l \sum_r \sum_s 2D^{hl}_{rs} \left({}^{oj}_{pq} \middle| {}^{hl}_{rs} \right) \right). \tag{19}$$

They are simply the Coulomb energy of a negative charge distribution $P^j_{pq}(\mathbf{r})$ in the field of the positive framework of all the nuclei and of the negative electron distribution $P^{hl}_{rs}(\mathbf{r})$. Due to cell electroneutrality, the total charge is zero so that electrostatic contributions of cells far apart from the center of the charge distribution $P^j_{pq}(\mathbf{r})$ goes to zero. In this sense, even if individual nuclear attraction and electron repulsion integrals decrease very slowly (in fact they diverge individually) there is a strict cancellation at medium or large distances between the diverging nuclear attraction and repulsion terms since positive nuclear charges compensate the negative electron density of an equal number of electrons. This is easily illustrated by analyzing the contributions of electrostatic interactions from cell h with charge distribution $P^j_{pq}(\mathbf{r})$:

$$P^j_{pq}(\mathbf{r}_1)\left(-\sum_u \frac{Z_u}{|\mathbf{r}_1 - \mathbf{A}_u - ha\hat{z}|} + \sum_l \sum_r \sum_s D^{hl}_{rs} \int \frac{2P^{hl}_{rs}(\mathbf{r}_2)}{r_{12}}\, dv_2\right) dv_1. \tag{20}$$

Let us denote the center of charge distribution $P^j_{pq}(\mathbf{r})$ by $\mathbf{U}(j)$ and the center of charge distribution $P^{hl}_{rs}(\mathbf{r})$ by $ha\hat{z} + \mathbf{V}(h - l)$. When the h index becomes large, one can approximate Eq. (20) as follows:

$$S^j_{pq}\left(-\sum_u \frac{Z_u}{|ha\hat{z} + \mathbf{A}_u - \mathbf{U}(j)|} + \sum_l \sum_r \sum_s \frac{2D^{hl}_{rs}S^{hl}_{rs}}{|ha\hat{z} + \mathbf{V}(h - l) - \mathbf{U}(j)|}\right). \tag{21}$$

For very large values of h, quantities $\mathbf{A} - \mathbf{U}(j)$ and $\mathbf{V}(h - l) - \mathbf{U}(j)$ become negligible with respect to $ha\hat{z}$ and at infinity an absolute cancellation takes place due to electroneutrality requirements:

$$-\sum_u Z_u + 2\sum_l \sum_r \sum_s D^{hl}_{rs}S^{hl}_{rs} = 0. \tag{22}$$

In a practical way, the summation over h in equation (19) is of poor convergence, especially in cases where the centers of positive (nuclei) and negative (electron cloud) charges within the unit cell are quite distant; this corresponds mathematically to the following series:

$$-\sum_{n=1}^{\infty} n^{-1} + \sum_{n=1}^{\infty} (n + \delta)^{-1}, \tag{23}$$

well known for its slow convergence. The number of terms required for a desired accuracy depends on the value of δ. Too strong a limitation will give erroneous results.

After completion of the integral calculations, energy terms and band structure are iteratively obtained from the previously defined matrices. The complex eigenvalue problem is solved for the k-points requested by the numerical integration (11). Diagonalizations are generally achieved with the very efficient and self-contained routine CBORIS (Zupan, 1975) which uses successively the Cholesky decomposition, the Houselholder similarity transformation, and the QL algorithm.

It is to be noted that the dimensions of the matrix equations to be solved are equal to the number of atomic orbitals, ω, per unit cell, the effect of the infinite lattice being included in the naturally converging but formally infinite sums.

The equations show that in a general way a polymer can be considered as a large molecule, making the usual methods of quantum chemistry appropriate to investigate the electronic structure of polymers. As a consequence, we are allowed to introduce, in practice, the well-known approximations of molecular quantum mechanics into the formalism of polymeric orbitals. For instance, we can either restrict the number of electrons to be considered or approximate several electronic integrals or groups of integrals. In this way, we will speak about *ab initio* or "nonempirical" techniques when considering all electrons and calculating all the necessary integrals. On the other hand, we will use semiempirical methods if a reduced number of electrons (e.g., valence electrons) are taken into account and, hence, the necessary integrals are approximated from experimental data.

B. Practical Procedures

In this section, we summarize some of the specific procedures which have been reported in the literature for calculating band structures of polymers.

1. The Parametric and Extended Hückel Methods for Polymers

In the simple Hückel method, only π-electrons of purely conjugated molecules are taken into account and all interactions, excepting those of nearest neighbors, are neglected (Hückel, 1931, 1932). This method is similar to the original formulation of the "tight binding approximation" in solid-state physics (Bloch, 1928). It obtained a considerable success because of its simplicity and the good validity of the approximations made for conjugated hydrocarbons. Due to the neglect of nonnearest interactions, only the first translation is retained in the secular system and determinant. As a consequence simple k-dependence of the π-energy bands are produced. Practical calculations are even easier since the atomic Hückel basis is assumed to be orthogonal; the overlap Bloch sums in those conditions reduce to the δ functions. The method has been applied with success to polyene chains, to graphite and to simple conjugated polymers (Ladik and Biczó, 1965; André *et al.*, 1967a,c).

Around 1950, a very simple type of Hückel parametrization, easily extended to σ-bonded systems, was suggested (Mulliken, 1949; Wolfsberg and Helmholtz, 1952). Ten years later, this method was taken up and applied to a large variety of saturated and conjugated molecules (Hoffmann, 1963).

This is now well known as the extended Hückel theory (EHT). In this procedure, only valence orbitals are considered. One-center interactions are estimated from ionization potentials and electroaffinities. Two-center integrals are computed from one-center ones by the Mulliken approximation and calibrated by a scale factor. Those integrals vary exponentially with distance so that Bloch sums are generally extended over 10- or 15-unit cells in secular systems. Due to the nonorthogonality, diagonalization procedures are more involved and sophisticated band structures can be generated. Since the procedure does not take into account the explicit electron–electron interactions, the method is noniterative and directly produces a stable band structure. Modifications to introduce explicit dependence on electronic charges have been attempted, giving rise to more complex iterative extended Hückel methods (IEHT) (Viste and Gray, 1964; Berthier *et al.*, 1965; Cusachs and Reynolds, 1965; Rein *et al.*, 1966).

A large variety of polymers have been studied by the standard EHT techniques including polyethylene and its fluoro-derivatives, polydiacetylenes, and substituted compounds, biopolymers. A detailed survey of the literature is given in Section III of this article.

2. *The Semiempirical NDO Methods for Polymers*

In 1965, more sophisticated methods were proposed which no longer have their justification on a global evaluation of matrix elements between basic orbitals (as in EHT) but rather on a closer analysis of energy terms (Pople *et al.*, 1965; Pople and Segal, 1965; Cook *et al.*, 1967). The matrix elements are split into their kinetic, attraction, and repulsion parts and some integrals are either neglected or approximated by careful procedures. Fundamentally, those methods attempt to reproduce by means of one-electron models the one-electron results of the many-electron Hartree–Fock theory. Sometimes, the fit to experimental data results in an implicit inclusion of a part of the correlation effects, although in a very indirect way. In the quantum theory of polymers, those relatively cheap empirical techniques could, in principle, play an interesting role but, from the start, they suffer from crude approximations and arbitrary simplifications. This makes the interpretative work ticklish, especially when at the point of discriminating between an interesting chemical effect and a failure of the method.

3. *The ab Initio LCAO Methods for Polymers*

The Hartree–Fock theory obtained its considerable success with the *ab initio* calculations on molecules of "chemical" sizes. In these calculations, all the electrons are taken into account and all the necessary integrals are explicitly computed for a given basis of atomic orbitals. There is actually a considerable development of *ab initio* type calculations on molecules.

However, it must be stressed that sophisticated *ab initio* methods, already very time consuming for middle- or large-sized molecules, become very costly when applied to polymers of interest. It thus makes sense to find compromises between the necessary computer investment and the quality of the results. This has prompted some groups to develop techniques, working at a strict minimal basis set level and retaining the full coherence of *ab initio* approaches, in order to perform affordable quantum mechanical studies of regular polymers.

An interesting approach is a generalization of the floating spherical Gaussian orbital (FSGO) technique (Frost, 1967) to calculate electronic band structures of regular saturated polymers within a simple and fast computing scheme (André *et al.*, 1976b). The iterative cycles of the usual SCF scheme are completely avoided by using a number of basic orbitals equal to the number of electron pairs in the unit cell. In that case, the density matrices are unequivocally fixed and computed by an inversion procedure on overlap integrals between Bloch functions. The order of one-electron energies of various alkanes computed with this FSGO technique relates well with the results obtained when using more extended basis sets and confirms this procedure as a valuable approach for interpreting XPS spectra (André *et al.*, 1976a). Similarly, variations in bond lengths and structures are in good agreement with experimental data. Such trends hold equally well for hydrocarbon polymers where X-ray scattering factors and properties directly related to electron distribution are well estimated (André and Brédas, 1977) while successful comparisons of theoretical electronic density of states with XPS spectra have been recently obtained for polyethylene and isotactic polypropylene.

However, in its initial form, the FSGO–CO method is not at all adapted for applications to unsaturated systems or more generally to those polymers for which an adequate description of the electron density within the unit cell requires a number of basis functions larger than the number of electron pairs. A recent and more elaborated FSGO–CO technique avoids this difficulty and keeps the full versatility and applicability of more sophisticated *ab initio* procedures (Brédas *et al.*, 1978). This technique has been found to be conceptually very simple and very well designed for testing numerical methods of the quantum theory of polymers. In practice, the procedure can be implemented either by the use of localized fragment orbitals or by atomic Gaussian lobe functions.

In the first approach generally known as "Christoffersen's molecular fragment approach" (Christoffersen, 1971, 1972), only 1s Gaussian-type orbitals (GTO) are used. Part of these orbitals are placed at points of space other than on nuclei, e.g., in bonding regions. In conjugated systems, the σ-framework is adequately described by somewhat localized FSGO's while the π-density is reproduced by antisymmetrical combination of s-Gaussian orbitals which are symmetrically placed above and below the atoms involved in the π-bonding.

In the second approach, we no longer make use of localized orbitals but instead keep the notion of atomic orbitals constructed as representations of Slater-type orbitals (STO). This type of basis set (Whitten, 1963, 1966) differs from the ordinary ones by the choice of 2p orbitals which, as in the previous case, are formed by non-nuclei-centered antisymmetrical combination of two s-GTO's.

4. *The Simulated ab Initio LCAO Methods for Polymers*

As previously noted, it is highly desirable to dispose of techniques which offer the advantages of both semiempirical (computationally fast) and *ab initio* (more reliable) schemes. Methods based on the simulated *ab initio* molecular orbital techniques (SAMO) (Duke and O'Leary, 1973; Duke *et al.*, 1975; O'Leary *et al.*, 1975) or on the application of the linear combination of localized orbitals (LCLO) (Delhalle *et al.*, 1977a) have been proposed. The principle of those methods is to build the Fock matrix elements of the polymer from transferable *ab initio* data obtained on large molecules. The *a priori* advantages are a very negligible cost (typically of the order of magnitude of an extended Hückel calculation) and the *ab initio* character of the approach. However in addition to the rather tedious generation of a huge number of matrix elements, it is still impossible to interpret persistent and significant discrepancies (shape and widths of electronic energy bands, etc.) with respect to reference LCAO–CO results. Moreover the total energy is out of reach by this method, thus severely restricting its interest.

Another promising and tempting procedure is based on the use of effective Hamiltonians (Barthelat and Durand, 1978). In such an approach, it is supposed that the valence Hamiltonian that reproduces the Fock operator is the sum of a kinetic term and of various effective atomic potentials for atoms within their characteristic chemical environment. For computational reasons, these effective potentials are chosen as nonlocal spherical or nonspherical Gaussian projectors. Parametrization of these potentials is performed by least-square fitting on corresponding valence Fock operators for small pattern molecules. This model has been successfully checked on small alkanes and on polyethylene and polypropylene (P. Durand, G. Nicolas, J. M. André, L. A. Burke, and J. Delhalle, unpublished results). It is also able to correctly reproduce interatomic distances and force constants.

At present time, all the previously cited methods (empirical, semiempirical or *ab initio*) are used to calculate energy band structure of polymers. Nonempirical methods are dependent on the fourth power of the number of orbitals considered, whereas empirical methods are more dependent on the square power. If N is the number of cells for which all the overlap and Fock elements are greater than a given threshold (e.g., 10^{-6}) and ω is the size of the basis set in a unit cell, we have to calculate

$N\omega(\omega + 1)/2$ Fock integrals and also $N\omega(\omega + 1)/2$ overlap terms. The repulsion operator by its two-electron character gives rise to much more complex matrices. Those matrices (of number N^3) contain $\omega^2(\omega + 1)^2/8$ integrals as a first approximation. As a consequence, the time-consuming part of an *"ab initio"* program is of a general $N^3\omega^4/8$ time dependence. Extended Hückel or CNDO band calculations have a much easier $N\omega^2$ dependence for the integral calculations. The time-consuming process is then the diagonalization and, as a consequence, those types of calculations are feasible on medium-sized computers in realistic computation times. Since extended Hückel calculations avoid SCF iterative process, one might think that the computing time ought to be less than the time needed for the CNDO method. However, the former involves heavy matrix transformations in order to treat the nonorthogonal atomic and Bloch basis, while the atomic and, consequently, the Bloch basis, are assumed to be orthogonal in the CNDO-type methods. As a result, computing time is very similar for both methods.

C. Numerical Aspects

In this part, we would like to show some numerical aspects of the actual quantum theory of polymers. Two main problems will be discussed; namely, the accurate computation of long-range electrostatic effects in linear chains and the correct indexing of energy bands.

1. Accurate Calculation of Long-Range Interactions

As pointed out in Section II,A, the Fock matrix elements are built from two divergent series representing the nuclear attraction and electron repulsion effects. Indeed, each term of the nuclear attraction as well as each term of the electron repulsion decays like r^{-1} and diverges individually. Only an appropriate combination will produce a finite result provided the neutrality of the cell is satisfied. For a system as simple as the linear chain of hydrogen atoms, it was found that going from 6 to 16 nearest neighbor interactions, the total energy is affected by an amount of 5×10^{-4} a.u. and the Fermi level changes by an amount of 10^{-2} a.u. (Kertesz, 1976).

Polar systems will even be more difficult to deal with since a large separation of positive and negative charges will result in an extended summation over unit cells. In a linear model of LiH molecules, 100 cells have to be considered on each side of the origin unit cell for getting matrix elements accurate up to 10^{-6} a.u. (Delhalle *et al.*, 1978a).

In an FSGO study of polyethylene, including 50 cells on each side of the origin unit cell does not give numerical values more accurate than 10^{-5} a.u. Moreover, too strong limitations produce band structures with incorrect band crossings (André *et al.*, 1978). This difficulty is not restricted to the computation of matrix elements but is also apparent in the calculation of total energies. This problem has been outlined for a long time (Löwdin, 1956a,b).

Total energies of large systems are proportional to the number of electrons, i.e., to the length of the polymer or to the number of unit cells it contains. For sufficiently large systems, the contribution to the total energy per unit cell is thus a constant. A detailed analysis shows that three divergent terms are involved in this contribution: the electron–nuclei attraction, the electronic repulsion, and the nuclear repulsion. A finite and stable numerical value can only be found by combining those three divergent terms in a proper way to electroneutral unit cells.

It must be noted that accurate results of total energies are requested if a conformational analysis or a vibrational analysis must be performed. In those cases, an accuracy of about 0.5 kcal/mole must be obtained on total energies (i.e., 10^{-3} a.u.). Such an accuracy on total energy needs a precision of about 10^{-5} to 10^{-6} a.u. in the evaluation of individual matrix elements.

In current applications, it is obviously impossible to deal with arbitrary large values of N (the number of unit cells between which we accurately compute all elementary integrals) since the two electron part of an LCAO calculation involves a large number of repulsion integrals going like $N^3\omega^4$.

A first possibility can be obtained by using accurate approximations of the F_0 functions involved in the integral evaluation over spherical Gaussian-type orbitals. It has been shown (Schaad and Morell, 1971) that for large arguments ($x \geq 16$), the F_0 function is accurately computed (error less than 10^{-8}) by the expression:

$$F_0(x) = \tfrac{1}{2}(\pi/x)^{1/2}. \tag{24}$$

Such an argument is rapidly obtained in molecular and polymer calculations. For hydrocarbons, in a FSGO scheme, this corresponds to distances of the order of 3.6 Å for C—C bonds and C—H bonds, of 1 Å for 1s carbon cores. The use of the previous approximation results in a significant gain in computation time but does not conceptually change the numerical procedure (Delhalle *et al.*, 1978a).

Much more elegant is the use of the Fourier transformation technique or the multipole expansion.

The first technique (Piela and Delhalle, 1978a,b) is based on the multipole expansion of the interaction operator r_{ij}^{-1}, within the Fock operator. This procedure is valid as soon as there is no longer overlap between the charge distributions whose interactions are to be calculated. This technique has the advantage that it takes into account the long-range effects on the Fock matrix elements directly within the framework of the exact LCAO–CO formalism. It is thus possible to perform the iterations to relax the charge distribution of the system in its own field. In this way, the induced multipole interactions are automatically taken into account. Exact summations up to infinity are carried out in such a way as to minimize the round-off errors. Selective control on the convergence and, more specifically, on the quantities one can afford to compute (electric moments) is possible. The method is implemented straightforwardly into computer programs since the necessary quantities are one-electron integrals and are easy to evaluate. The number of such integrals grows as the second power of the basis length and is proportional to the number of interacting cells, i.e., $\omega^2 N$. The computing time is thus negligible with respect to the usual technique where a number of $\omega^4 N^3$ is to be computed. The method has the supplementary advantage of being physically appealing through the multipole moments which constitute a traditional framework for interpreting electrostatic interactions.

The second technique (André *et al.*, 1978) makes use of the approximation, Eq. (24), of the F_0 function. It is thus only valid for the formalism where those functions are found, i.e., mainly for Gaussian-type integrals. If this approximation is used in the infinite summations of the LCAO–CO formalism, we isolate quantities which are immediately recognized as traditional Madelung expressions. Those Madelung summations, characteristic of the system and of the basis set used, can be evaluated by a direct application of the Fourier representation method (Harris and Monkhorst, 1970; Harris, 1972, 1975).

As in the multipole expansion, the long-range effects are included within the SCF cycles. There is a formal advantage since no approximation of the r_{ij}^{-1} operator is involved as is the case in the multipole expansion. In its application to the determination of long-range effects in polymers, the technique is presently restricted to Gaussian bases only.

2. *The Band Indexing Difficulty*

In the reduced zone scheme, the band structure $E_n(k)$ is a multivalued function of k and the successive values of n for any given k correspond to increasing levels: first to the inner core electrons, then to valence bands, and finally to unoccupied bands. Due to prohibitive computing time associated with the complex diagonalization:

$$F(k)C_n(k) = S(k)C_n(k)E_n(k), \tag{25}$$

the energy bands are known at discrete wave number values only, $E_n(k_i)$. Bands are usually given an index n, increasing in order of increasing energy, but difficulties arise because the rule breaks down at degeneracy

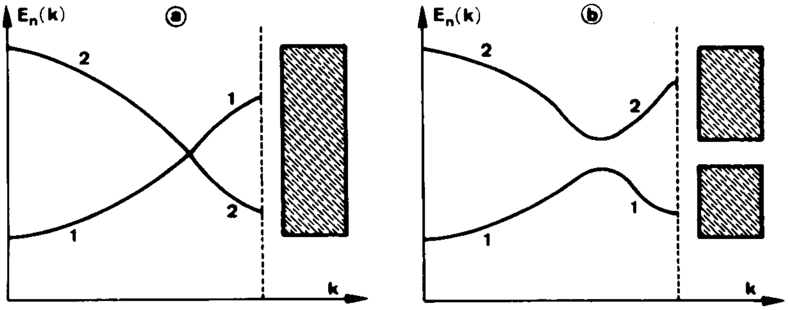


Fig. 1. Crossing (a) or noncrossing (b) of energy bands resulting in the absence (a) or the existence (b) of an energy gap. (Reproduced with permission from Delhalle, 1975.)

points. Figure 1 illustrates this situation in the case of two energy bands. Two possibilities exist: either band crossing or noncrossing. To these two alternatives correspond entirely different physical realities; in the first case, there is no band gap, whereas such an energy gap does exist in the second case. If point symmetry exists within the unit cell the problem can be solved by group theory (McCubbin, 1971), but when there is a low symmetry and a large number of bands this becomes practically untractable. For the present time, the band indexing problem is generally solved by visual inspection or in a more based way using a graphic interactive program with the knowledge of the local first derivatives, $E_n'(k_i)$ of the energy eigenvalues. A general technique, easily applicable, for determining locally the k-derivatives of all bands in each k-point (André *et al.*, 1972; Delhalle, 1972) has been obtained by considering the scalar product:

$$E_n(k) = C_n^\dagger(k)F(k)C_n(k) \tag{26}$$

with

$$C_n^\dagger(k)S(k)C_n(k) = 1. \tag{27}$$

Hence,

$$\frac{\partial E_n(k)}{\partial k} = \frac{\partial C_n^\dagger(k)}{\partial k} F(k)C_n(k) + C_n^\dagger(k) \frac{\partial F(k)}{\partial k} C_n(k) + C_n^\dagger(k)F(k) \frac{\partial C_n(k)}{\partial k}. \tag{28}$$

After straightforward matrix manipulation, we get the working formula:

$$\frac{\partial E_n(k)}{\partial k} = C_n^\dagger(k) \frac{\partial F(k)}{\partial k} C_n(k) - E_n(k)C_n^\dagger(k) \frac{\partial S(k)}{\partial k} C_n(k), \tag{29}$$

which simplifies in an orthogonal atomic or Bloch basis to:

$$\frac{\partial E_n(k)}{\partial k} = C_n^{\dagger}(k)\,\frac{\partial F(k)}{\partial k}\,C_n(k). \tag{30}$$

Since the k-dependence of $F(k)$ and $S(k)$ matrices is very simple:

$$\frac{\partial F_{pq}(k)}{\partial k} = \frac{\partial}{\partial k}\sum_j \exp(ikja)F_{pq}^j = ia\sum_j j\exp(ikja)F_{pq}^j, \tag{31}$$

$$\frac{\partial S_{pq}(k)}{\partial k} = \frac{\partial}{\partial k}\sum_j \exp(ikja)S_{pq}^j = ia\sum_j j\exp(ikja)S_{pq}^j. \tag{32}$$

It has been possible to implement this procedure into an algorithm of general use for polymer consideration (Delhalle *et al.*, 1978c).

In a practical way, the band structure calculation program is linked to a graphical device (plotter or graphical display) and the variations of $E_n(k)$ are graphically represented by a rotating fixed length vector centered at $(E_n(k), k)$ points as represented in Fig. 2. When two or more bands are not separated enough in the graph, it is possible to enlarge the graphic. Experience has shown that nearly all kinds of bands can be labeled by this procedure and the time and the difficulties of data manipulation are drastically reduced.

D. Band Structure Calculations and Electronic Properties of Polymers

As soon as the final eigenvalue equation (5) has been set up according to one of the various methods described above, the problem of getting information to compare with experiment has to be considered. Indeed, making guesses or interpretations on properties is the most fruitful and rewarding aspect of theoretical studies. Numerous experimental quantities can be correlated with terms deduced from electronic charge distribution and from band structure calculations. However, band structure is not measured directly and one has to apply transformations to bring calculated data to a form readily comparable to experiment. This is typical of the electronic density of states distributions. Electronic charge distribution is also very attractive to consider when discussing chemical bonds.

In polymeric materials and in three-dimensional crystals, the number of atoms is very large and the eigenvalues are very dense and bounded; it is therefore more convenient to deal with energy levels distribution functions than with individual values. Density of states is defined as the number of allowed energy levels per unit energy:

$$\mathcal{D}(E) = (a/\pi)\sum_{n=1}^{n_0} |dk/dE_n(k)|_{E_n(k)=E}. \tag{33}$$

Such a function has the advantage of providing a synthetic description of one-electron levels and of being closely related to experiment.

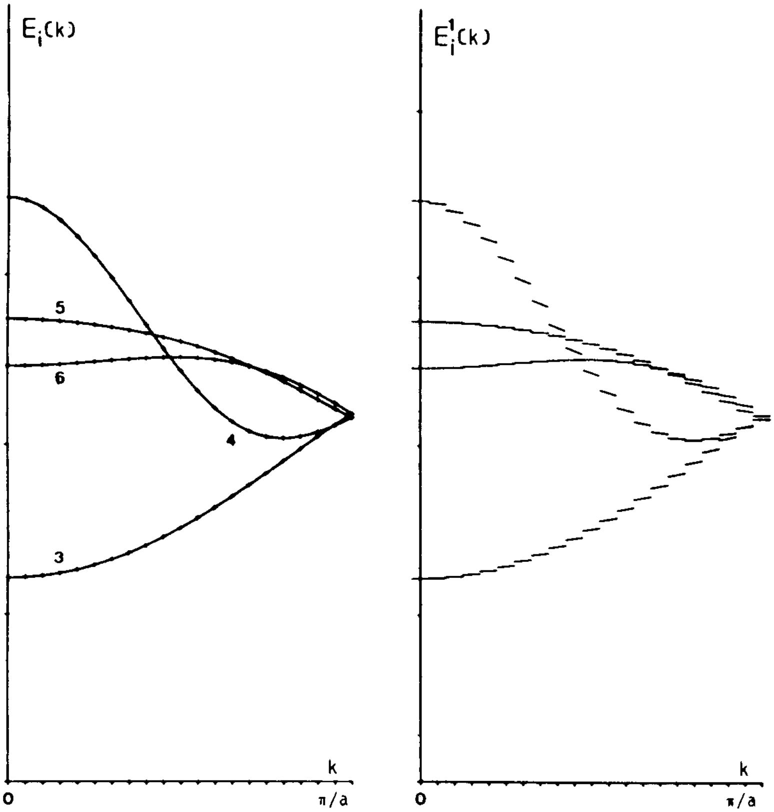


Fig. 2. Energy bands $E_i(k)$ plotted from their discrete values at 21 k-points and their first derivatives $E_i^1(k)$. The example represents the four last-occupied valence bands of polyethylene (3, 4, 5, 6).

Though conceptually simple, electronic density of states distributions are not straightforward to evaluate numerically but the problem has been well analyzed and described in the literature (Delhalle, 1974a,b; Delhalle and Delhalle, 1975). Whenever the first derivative of $E_n(k), E'_n(k)$, goes to zero, there will be a large contribution to the density of states, giving rise to singularities in the spectra. We know in advance there are at least three values of the wave number k where $E'_n(k)$ vanishes identically, i.e., $k = 0$ and $k = \pm\pi/a$. However other extrema can occur but their number and location cannot be predicted *a priori*. When implementation by computer is to be considered, one has to turn to a new function $D(E_i)$, the density of electronic states histogram. It exhibits the same behavior as $\mathcal{D}(E)$ but fortunately gives rise to finite peak height only. It is an average of $\mathscr{D}(E)$ over some energy interval ΔE

$$D(E_i) = (\Delta E)^{-1} \int_{E_i-\Delta E/2}^{E_i+\Delta E/2} \mathscr{D}(E)dE,$$
$$E_i = E_0 + i\,\Delta E, \quad i = 0, 1, 2, \ldots, \quad \Delta E > 0. \tag{34}$$

The value of ΔE is chosen according to a particular experiment to interpret. $\mathscr{D}(E_i)$ is obviously the limiting function of $D(E_i)$ when ΔE goes to zero. $D(E_i)$ is the height of the ith rectangular box centered at E_i; it represents an element of the histogram of the density of states.

Various approaches to compute these $D(E_i)$ have been proposed. The histogram method involves generating the eigenvalues $E_n(k)$ of the system at a large number of points in the Brillouin zone and then counting the results into boxes that subdivide the energy range. Another approach, more effective in the sense of saving computing time and numerical stability, is currently preferred. Basically, it consists of an analytical or continuous integration over small intervals $[k_i - \Delta k/2, k_i + \Delta k/2]$ located at wavenumber values for which a diagonalization has been requested.

Density of electronic states distributions can be used to test results on energy band structure obtained from various methods. It is known that derivatives enhance variations of a function; therefore, it is no wonder a density of states distribution is a more sensitive test on energy bands since it involves the inverse of first derivatives $E_n'(k)$.

Most of the comparisons between theory and experiment cited in the next section are based on photoelectron spectroscopy mainly by XPS (or ESCA), which has been proven to be a valuable means for investigating ground-state properties of polymers (Delhalle *et al.*, 1974b; Pireaux *et al.*, 1974, 1977).

Experimental XPS spectra are not directly comparable to density of states, but two additional effects must be introduced in the calculations to produce curves with almost the same type of content as in the experiment. Firstly, one has to take the photoionization cross section into account. It induces an appreciable modification of the general shape of the density of states. Until now the calculations of the XPS intensities have been made mainly according to the empirical model of Gelius (1972):

$$I_n(k) \sim \sum_p C_{np}^*(k) \left(\sum_q S_{pq}(k) C_{nq}(k) \right) \sigma_p, \tag{35}$$

where $I_n(k)$ represents the intensity of the nth energy band at the point k of the first Brillouin zone whose LCAO coefficients are $\{C_{np}(k)\}$. $S_{pq}(k)$ is an overlap integral between Bloch functions p and q at point k. The σ_p's are the relative photoionization cross sections referring to a particular atomic subshell; approximate values were obtained by Gelius by applying the intensity model to small molecules with known orbital assignments. This produces the density of electronic states distribution modulated by the XPS cross sections, $\bar{D}(E_i)$, where each energy eigenvalue contribution is weighted by $I_n(k)$:

$$\bar{D}(E_i) \sim (\Delta E)^{-1} \sum_{n=1}^{n_0} \int_{E_i-\Delta E/2}^{E_i+\Delta E/2} \{|E'_n(k)|^{-1} I_n(k)\}\, dE. \tag{36}$$

Secondly, XPS spectrometers have an inherent limitation in their resolution. For example, the HP.5950 A ESCA spectrometer shows a resolution ranging from 0.5 to 0.7 eV. However, it is found that the experimental resolution function is reasonably well fitted by a Gaussian function of full width, Γ, at half maximum of 0.7 eV. The final theoretical simulation $D_s(E_i)$ of the experimental spectrum is obtained after the convolution of $\bar{D}(E_i)$ by the Gaussian fitting curve has been made:

$$D_s(E_i) \sim \int_{-\infty}^{+\infty} \bar{D}(E) \exp[-(E - E_i)^2/2\sigma^2]\, dE, \tag{37}$$

where $\Gamma = 2.345\sigma$.

III. Applications of LCAO Methods to Stereoregular Polymers

Historically, polymeric orbital theory has followed the general scheme which in turn was followed by molecular orbital theory. In this sense, early applications were mainly concerned with the general understanding of π-electrons of aromatic or conjugated one- or two-dimensional polymers. The pioneering works in this field were the independent descriptions of the electronic structure of graphite in a Hückel-type procedure (Bradburn *et al.*, 1947; Wallace, 1947; Barriol and Metzger, 1950). Similar work was already published for the infinite polyene chain (Lennard-Jones, 1937; Kuhn, 1949). After the work of these pioneers, both graphite and polyene were studied by approximate theories taking account of the details of electron interactions. The aim of these works was primarily to establish or to demonstrate the existing link between molecular orbital theory and solid-state physics. The generalization of LCAO methods to all electrons or at least to all valence electrons in a polymer is quite recent, and the first all valence calculations on a significant polymer is the extended Hückel band structure of polyethylene (McCubbin and Manne, 1968). More recently, advances in mathematical techniques and programming have reached the point at which moderate-sized polymers can be studied by sophisticated methods. This allows the quantum theory of polymers to advance in a clearly defined way. The purpose of this section is to present some selected problems which can be accurately investigated by band structure methods. The first question is concerned with polyethylene. Polyethylene is to the quantum theory of polymers what the hydrogen molecule is to molecular physics: a traditional test case for each newly developed technique. Then some results for polypropylene will be presented. Chemically polypropylene is the methyl derivative of polyethylene. This substitution has the effect of producing different tacticities. The methyl groups can be located on the same side of the main chain (isotactic polypropylene), can alternate on each side (syndiotactic polypropylene), or can be aleatory (atactic polypropylene). Furthermore, while polyethylene in its stable conformation is a perfect zigzag chain, polypropylene in its different tacticities is a helix. This raises the questions of the influence of conformational effects on observable electronic properties and the possible inference by quantum methods. Finally, the series of fluoropolymers provides an example of the effects of systematic substitutions.

Most of the compounds studied in this part are available as well-crystallized samples. Experimental ESCA spectra have been determined (Pireaux *et al.*, 1974, 1977) and thus a homogeneous series of experimental results can be compared to theoretical results of band structure calculations.

A. Polyethylene

Polyethylene provides an example of the recent proliferation of band structure calculations for polymers. Its band structures is well described in the literature as obtained from extended Hückel (McCubbin and Manne, 1968; Imamura, 1970; André *et al.*, 1971b; André and Delhalle, 1972; Falk and Fleming, 1973), CNDO/2 (Morokuma, 1970, 1971; Fujita and Imamura, 1970; André *et al.*, 1971b, 1972; Morosi and Simonetta, 1971; André and Delhalle, 1972; McAloon and Perkins, 1972; Falk and Fleming, 1973; Wood *et al.*, 1975), INDO (Beveridge *et al.*, 1972), MINDO/2 (Beveridge *et al.*, 1972; Beveridge and Wun, 1973), MINDO/3 (Dewar *et al.*, 1974), SAMO and localized orbitals (Duke and O'Leary, 1973; Delhalle *et al.*, 1977a), MNDO (Dewar *et al.*, 1977), and *ab initio* (André, 1970a; Clementi, 1971; André *et al.*, 1973; Delhalle *et al.*, 1974b) methods. All calculations show a large gap between the filled and the vacant band in agreement with experimental facts and data. Figure 3 reproduces band structures and their electronic density of states distributions of polyethylene calculated with *ab initio* (André and Leroy, 1970b), FSGO (André *et al.*, 1976b), LCLO (Delhalle *et al.*, 1977a), CNDO/2 and CNDO/W (Delhalle *et al.*, 1974b), and extended Hückel (André *et al.*, 1971b) methods. As expected, the general structure of valence bands as

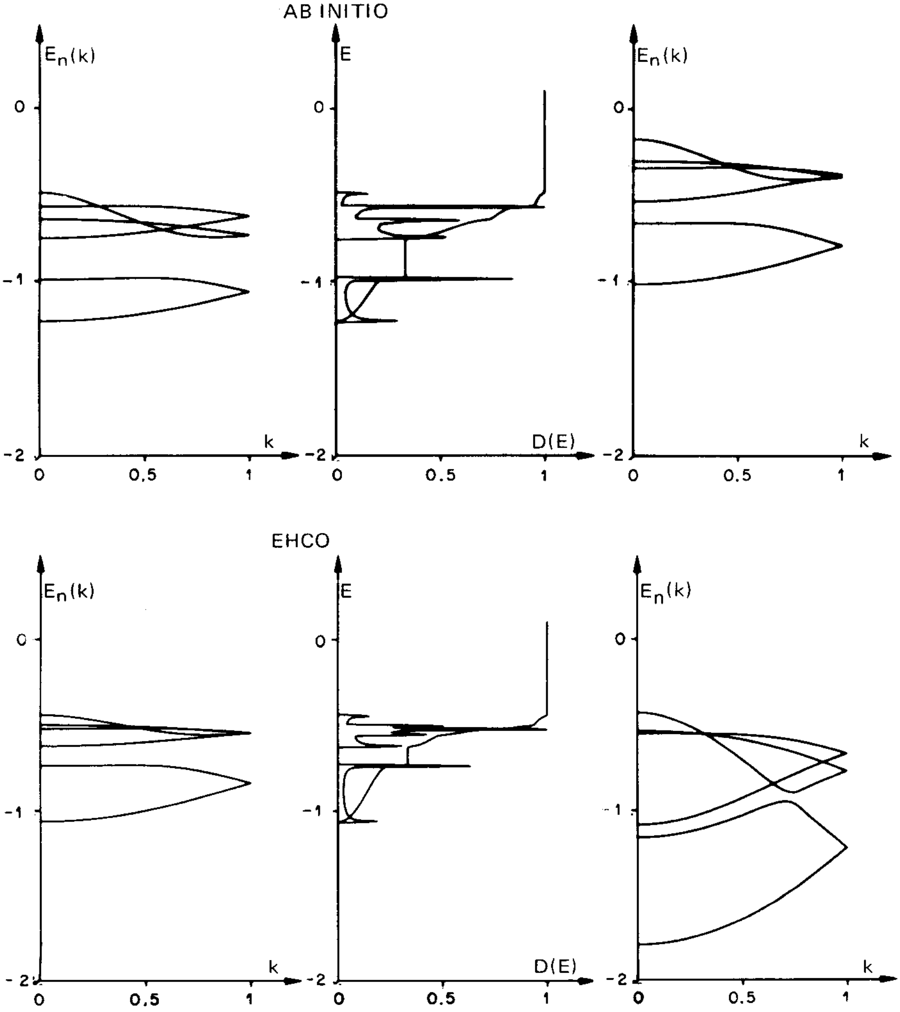
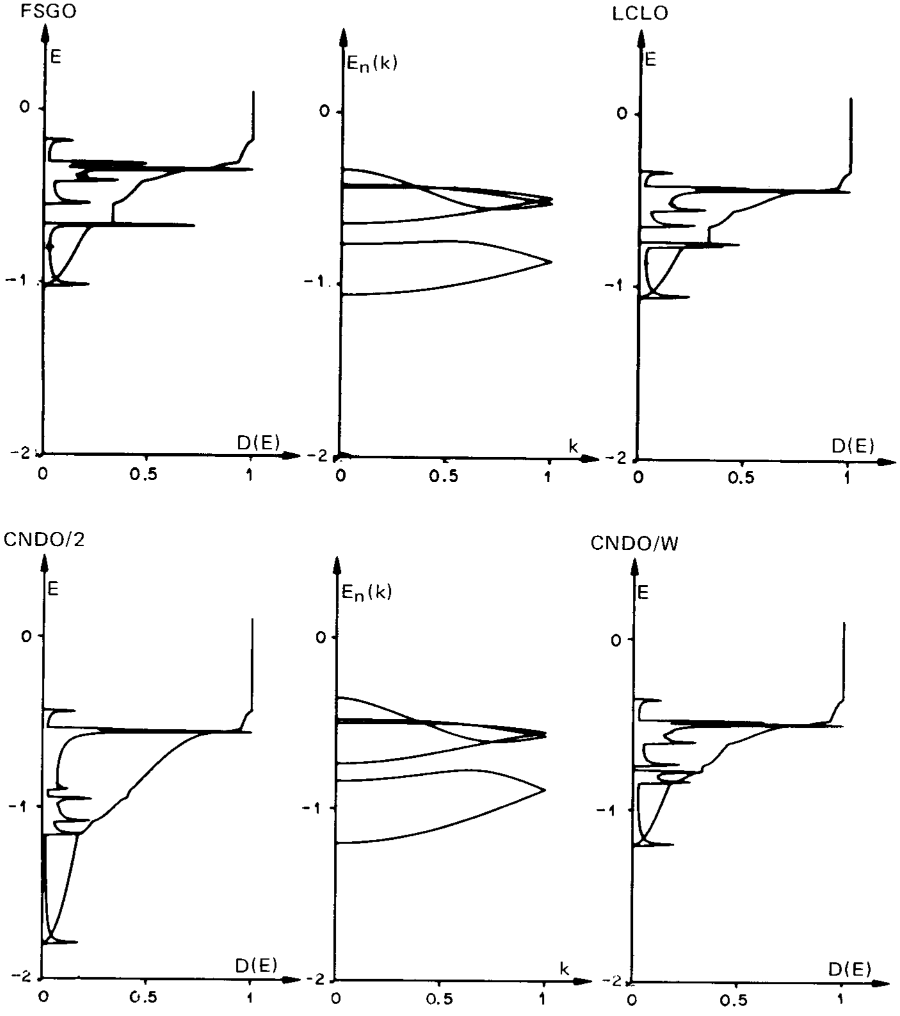


Fig. 3. Valence energy bands of polyethylene in various approximations: EHCO = extended Hückel calculation in the original Hoffman's parametrization, CNDO/2 = CNDO calculation in the original Pople's parametrization, CNDO/W = CNDO calculation in a modified parametrization aimed to reproduce orbital energies of alkanes, AB INITIO = *ab*

obtained from the six methods appears to be very similar. However, CNDO/2 bands in the original parametrization are much too large and band centers of gravity lie far below *ab initio* ones. At first sight, this



inito calculation using Clementi's 7s–3p Gaussian basis, FSGO = *ab initio* calculation using original Frost's floating spherical Gaussian orbitals, LCLO = simulated *ab initio* calculation from localized orbitals obtained from localization on ethane and propane. (Reproduced with permission from André and Delhalle, 1978.)

defect can be corrected by the parametrization CNDO/W especially designed for reproducing molecular orbital levels of hydrocarbons (Sichel and Whitehead, 1968).

On the other hand, while the band structures show a very similar overlap shape, the density of states which are directly related to energy band curves are very dependent on the approximation used. Two sets can be distinguished: a first group containing CNDO/2 and CNDO/W calculations and a second one including extended Hückel, LCLO, *ab initio* and FSGO results. From experimental comparisons, it turns out that the methods of the second group are better designed for correctly reproducing experimental density of states than CNDO-type ones. When taking into account the cross-section effects, the comparison between a theoretical density of states and experimental valence spectra is excellent (J. Delhalle *et al.,* 1975). Such a comparison is given in Fig. 4 in the case of FSGO results. The positions of both theoretical and experimental peaks agree surprisingly well and both fine structures are directly comparable. It should be stressed at this point that, in author's opinion, band structure plots do not offer the best representation of valence band properties, particularly in those systems where a large number of bands lies in a narrow energetic region. The density-of-states (possibly convoluted for simulating experimental resolution and taking into account cross-section effects) histograms appear to be a more apt picture of the distribution of valence levels when the comparison with experiment is within an energetic space.

A natural development of this work is to study the effects of structural perturbations, like conformational changes on the valence bands of stereoisomers. Though conformations are practically impossible to trap rigidly in pure and highly crystalline polymers, such an investigation could, however, be applied to stereoisomers. We consider in this section the four conformations relevant to the interpretation of the vibrational spectrum of polyethylene (Morosi and Simonetta, 1971; Snyder, 1967), i.e., $(T)_\infty$, $(G)_\infty$, $(TG)_\infty$, $(TGTG')_\infty$, whose structures are given in Fig. 5. The conformation $(T)_\infty$ is found to be the most stable experimentally. The variations of the spectrum of monoelectronic levels due to geometry changes are important and can, in principle, be observed by experimental techniques like XPS, giving a new and appealing dimension to both theoretical and experimental studies.

The occupied valence electronic energy bands and the density of states of the four relevant polymers are reproduced in Fig. 6 as obtained by an LCLO method. Similar results are given in the literature in the framework of the extended Hückel method (S. Delhalle *et al.,* 1975). Significant differences are theoretically observed. The C—C bands, e.g., are affected by the conformational changes. Their bottom energy values are almost constant whereas a modification takes place at the top of the bands, giving rise to important changes in the bandwidths (9.45 eV for T, 11.27 eV for

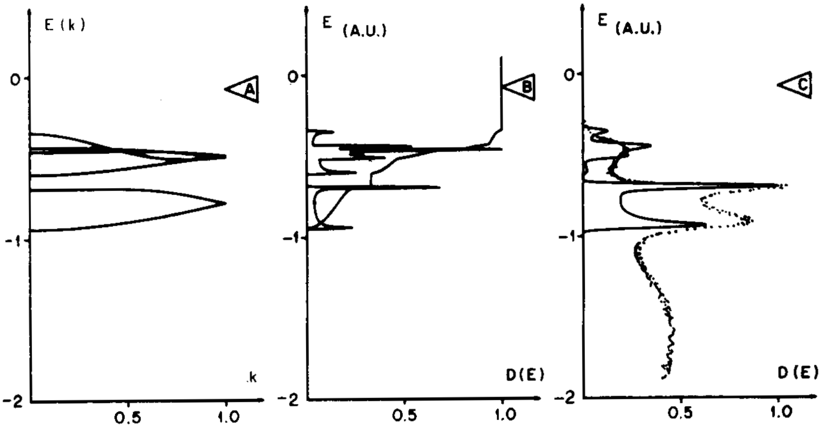


Fig. 4. Polyethylene zigzag studied by an *ab initio* technique using FSGO's, from left to right: valence energy bands, valence density of states, theoretical convoluted density of states with cross-section effects (full line), experimental ESCA spectrum (dashed line).

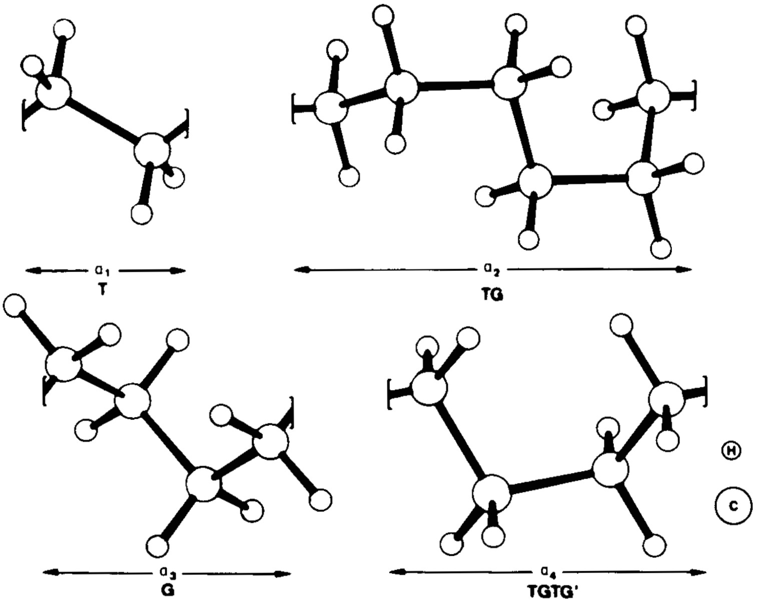


Fig. 5. Four possible conformations of polyethylene. T is the conformation experimentally observed in crystal form (a_1 = 2.51 Å, a_2 = 6.16 Å, a_3 = 3.56 Å, a_4 = 4.35 Å). (Reproduced with permission from Delhalle *et al.*, 1977a.)

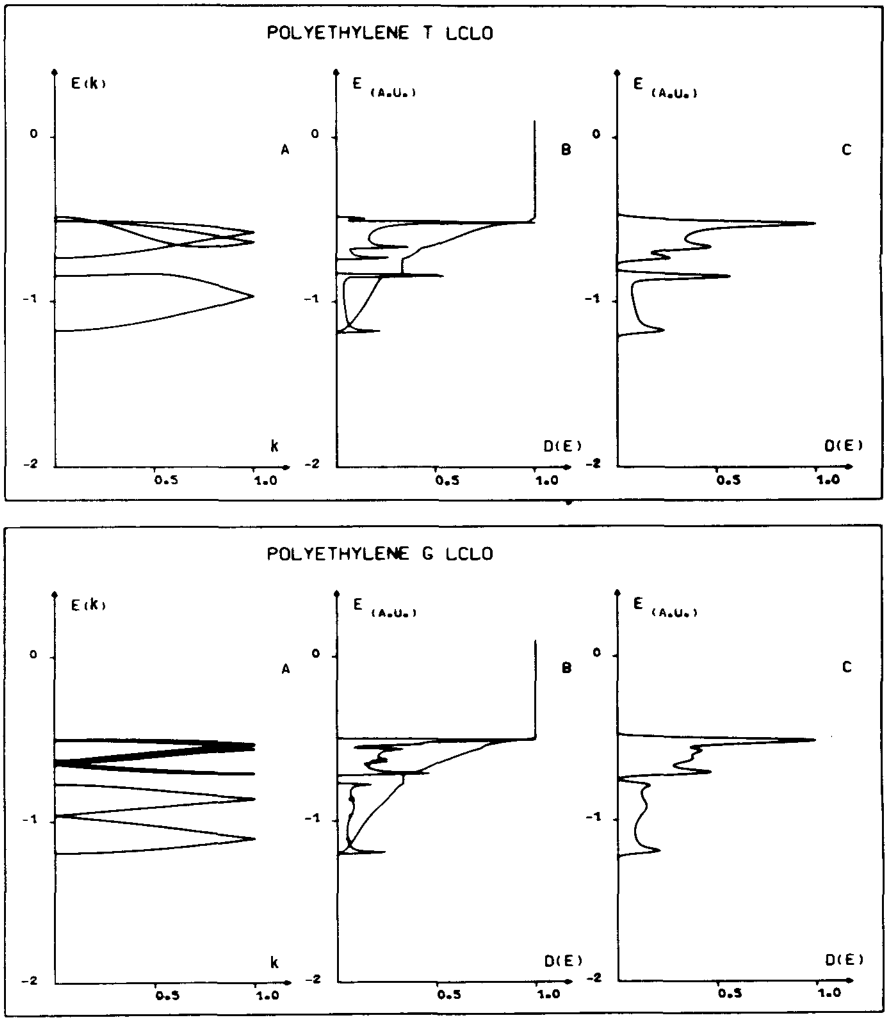
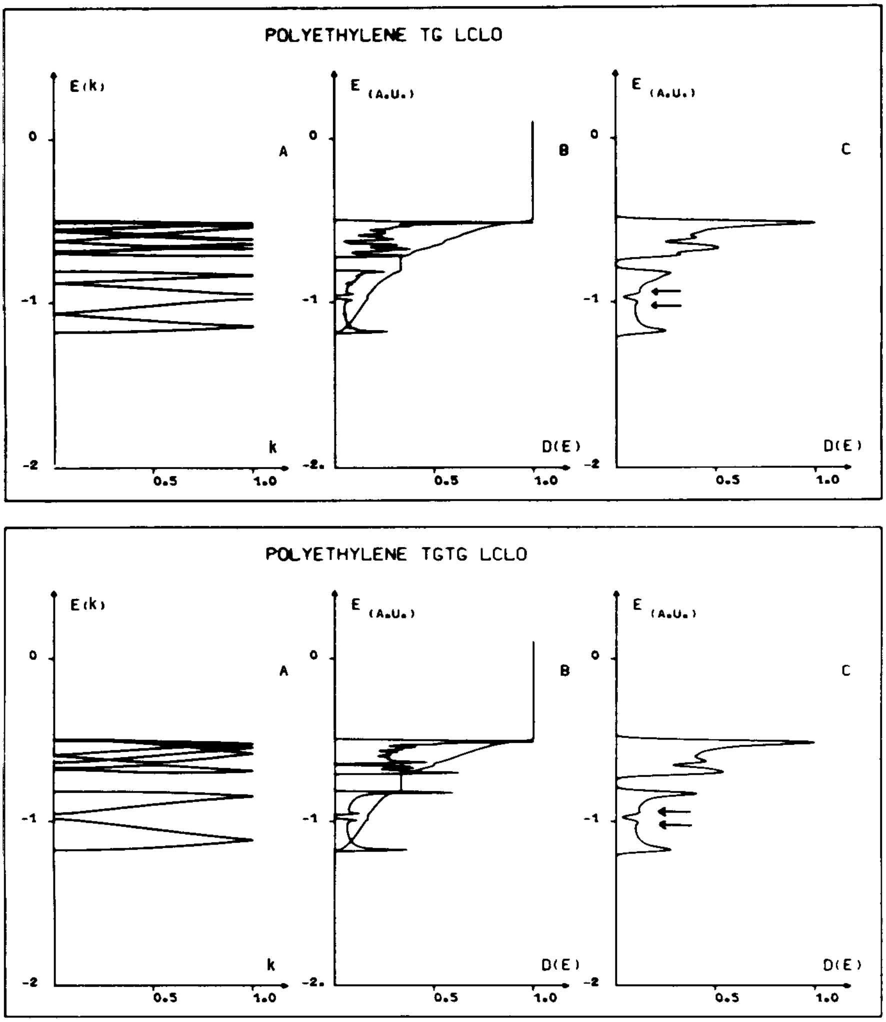


Fig. 6. Valence energy bands, density of states and convoluted density of states (without cross-section effects) of T, TG, G, and TGTG′ conformations of polyethylene obtained from a simulated *ab initio* technique using localized orbitals. Note the existence of an energy gap

G, 9.69 eV for TGTG′, 10.15 eV for TG). The experimental resolution of recent XPS spectrometers (0.5–0.7 eV on HP 5950 A, e.g.) and the relative high XPS cross sections of $2s_C$ and $2p_C$ atomic orbitals lead to the



(noted by an arrow) in the lowest part of the valence bands of TG and TGTG′ conformations. (Reproduced with permission from Delhalle *et al.*, 1977a.)

expectation of detectable variations. It is interesting to note that the G conformation with the unit cell presenting the lowest symmetry among the four compounds exhibits the largest width. Another interesting aspect is that an energy gap (noted by arrows in Fig. 6) appears in the C—C bands of compounds for which it is possible to find —CH_2—CH_2— groups with the same relative orientations of the CH_2's as in polyethylene T. The same relative orientation means that the two hydrogens of one CH_2-group are related to the ones of the second CH_2 by an inversion operation. As a consequence, TGTG′ and TG conformations show such an energy gap while T and G conformations do not. Again, this could probably be experimentally observed if well-characterized samples of those four polyethylenes were available. Though conformations like the ones we discussed previously are practically untractable for experimental works, it turns out from these theoretical computer experiments that similar and actual effects could, in principle, be observed by an actual XPS experiment. A real example is described in Section II,B.

B. Polypropylene

The previous computer experiment on various conformations of polyethylene predicts definite and actually observable influence of the conformation on the electronic density of states distributions. It is the

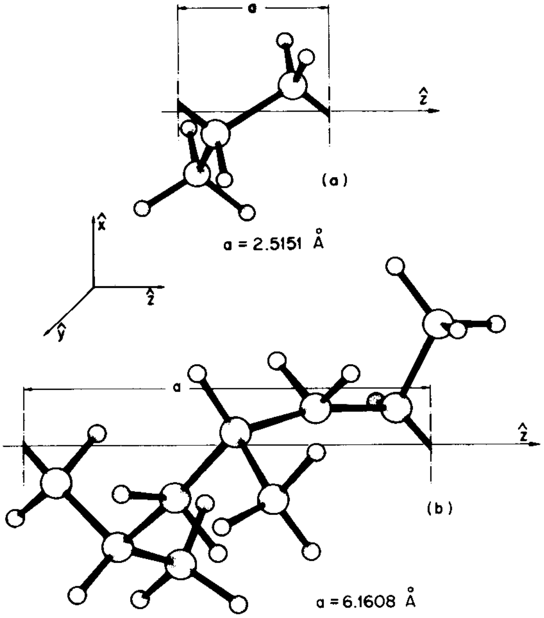


Fig. 7. Zigzag planar (top) and helical 2 × 3/1 (bottom) conformations of isotactic polypropylene. (Reproduced with permission from Delhalle *et al.*, 1979b.)

purpose of this section to investigate a real case and prove the existence of observable conformational effects on the electronic density of states of an isotactic polypropylene.

Most of the calculations so far performed on real polymers assume a zigzag planar conformation of the main chain. In the case of an isotactic polypropylene, the geometrical disposition is that of an isoclined 2 × 3/1 helix (Natta *et al.*, 1960). An ideal linear situation and the actual conformation are represented in Fig. 7. Figure 8 plots the theoretical valence spectrum for an isotactic polypropylene of assumed zigzag conformation (top) and of the helix form (bottom) and the experimental XPS spectrum. (Delhalle *et al.*, 1979b). A tentative interpretation of the experimental XPS spectrum by assuming a zigzag planar conformation has been reported (Pireaux *et al.*, 1977). No satisfactory agreement as to the peak structure was obtained. For example, the intense and sharp peak (C in Fig. 8) emerging in the middle of the experimental C—C clump does not appear in the theoretical zigzag calculation. Moreover, only four distinct peaks are observed in the calculated spectrum instead of five as in the experiment. The previous and inspiring computer experiment could suggest the role of conformational effects. Thus, the problem has been reinvestigated by computing the properties of an isotactic polypropylene in its actual helix form. As a consequence of the sole conformation inclusion the right structure now appears in the C—C clump: the missing intense and sharp peaks (C) is now present and the outermost subbands split into two components. Peaks intensities also follow the right order. The same type of agreement has been obtained by three different methods: extended Hückel, FSGO, and pseudopotentials crystalline theories.

C. Fluoropolymers

Following the path initiated in the previous sections, the effects of a systematic substitution on the electronic structure of model polymers can be carried out. The series of polymers obtained from fluoro-substituted ethylene is ideally designed for this purpose, as most compounds are available to the experimentalists and fluorine is likely to induce some of the strongest electronic effects on the band structures. The unit cells of the actual polymers are presented in Fig. 9.

Foremost attention is paid in the literature to measurements and calculations of core-level binding energies and of chemical shifts associated with changes in chemical environment (Clark and Kilcast, 1971a,b; Ginnard and Riggs, 1972; André and Delhalle, 1972; Clark *et al.*, 1973, 1974; Delhalle *et al.*, 1974a). A direct estimation can be obtained by making use of Koopman's theorem and correlating the mean value of core bands

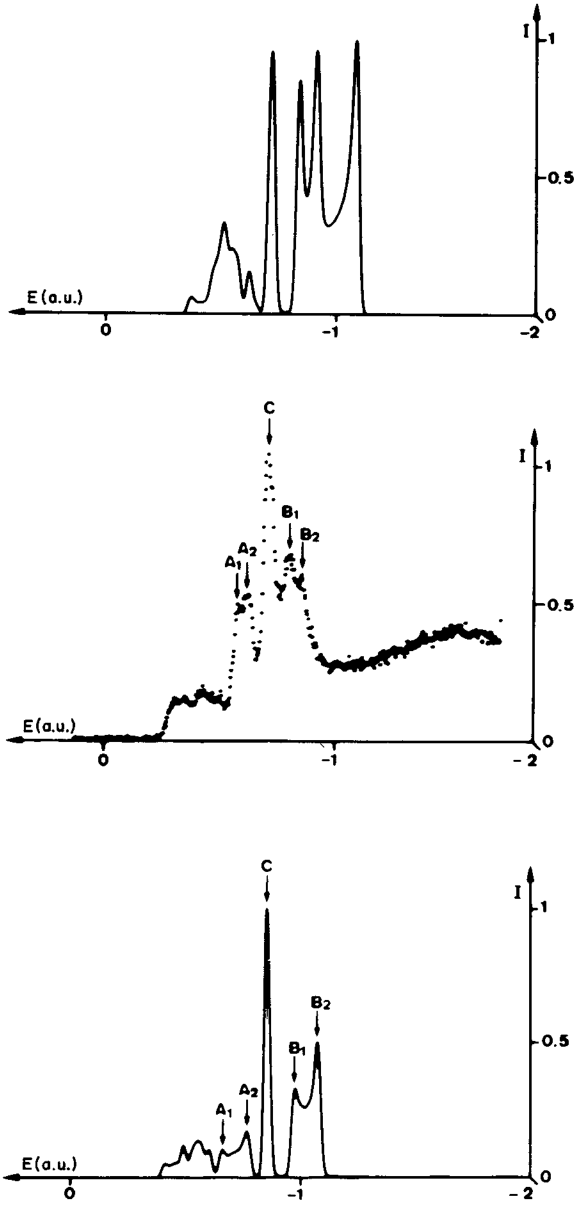


Fig. 8. Calculated spectra (full lines) of the valence bands of an isotactic polypropylene in a zigzag conformation (top) and in a helical 2 × 3/1 conformation (bottom) as obtained from an extended Hückel calculation including cross-section effects; the experimental XPS spectrum is given in dashed line. (Reproduced with permission from Delhalle *et al.*, 1979b.)

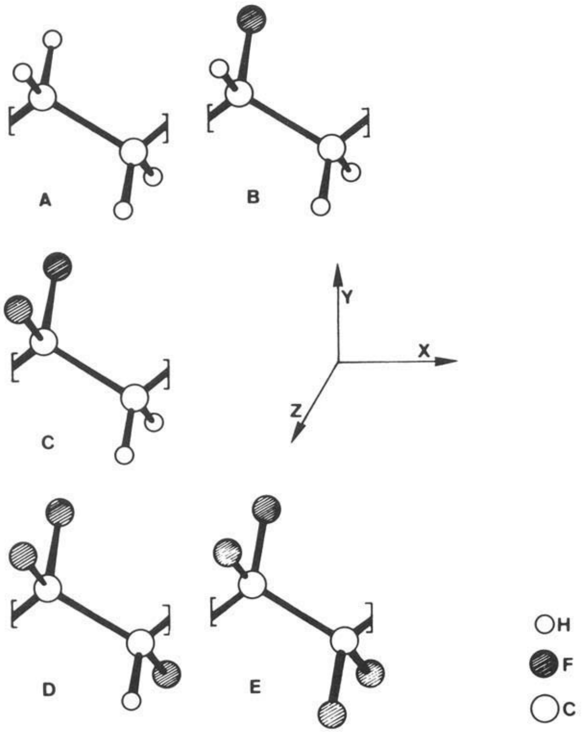


Fig. 9. Unit cells of five possible fluoro-derivatives of polyethylene; A is polyethylene, E is polytetrafluoroethylene (known as Teflon). (Reproduced with permission from Delhalle *et al.*, 1977b.)

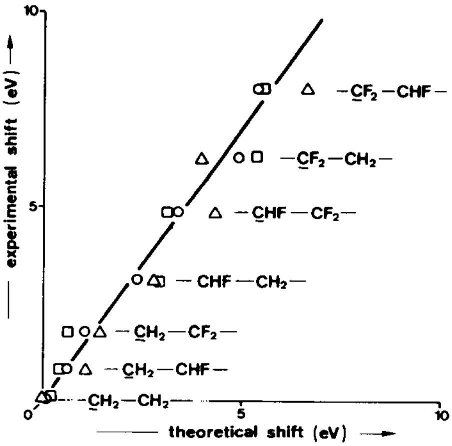


Fig. 10. Calculated carbon 1s-core shifts in the series of fluoropolymers: ○ from the charge potential model using CNDO/2 gross charges, △ from a simulated *ab initio* technique using localized orbitals, □ from an *ab initio* FSGO technique.

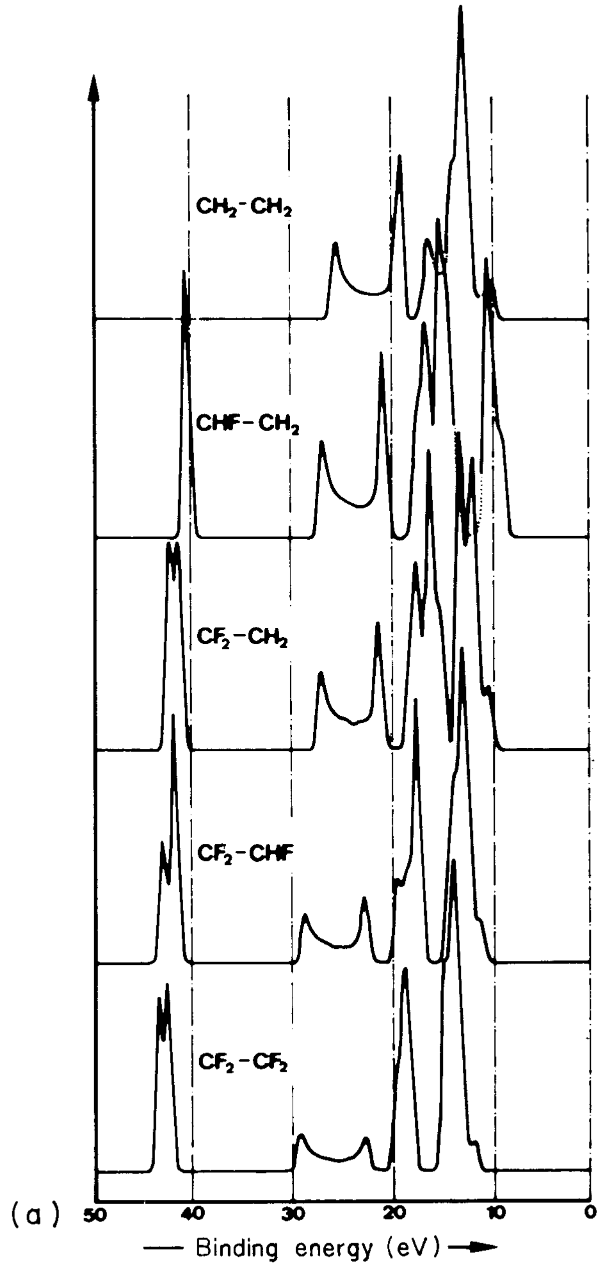


Fig. 11. (a) Theoretical convoluted density of states obtained from an *ab initio* FSGO technique (cross-section effects not included).

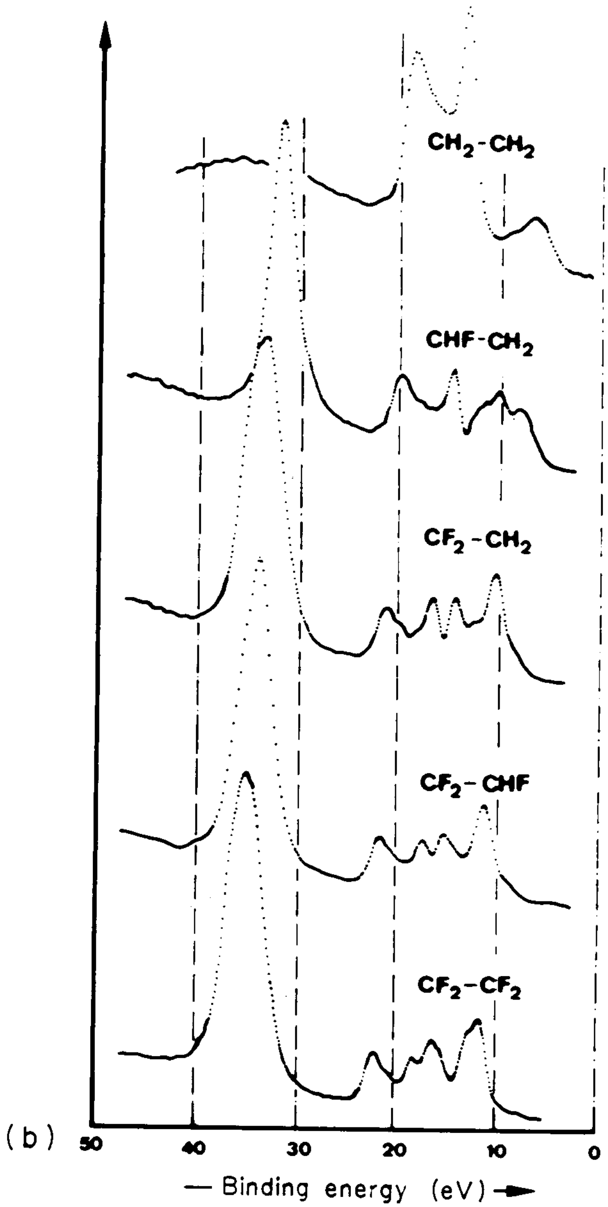


Fig. 11. (b) Experimental XPS spectra (right) of the valence bands of fluoropolymers.

(always very narrow, less than 0.01 eV) to experimental core binding energy. For obvious reasons, this technique can only be used in the methods where an explicit consideration of core orbitals is included (*ab initio,* localized orbitals). For the other techniques (extended Hückel, CNDO-type, pseudopotentials), only valence bands are available. In such cases it has become traditional to use the now classical charge potential model (Sieghbahn *et al.*, 1969). This model relates core binding energies to charge distributions by means of the formula:

$$E_i = E^0 + kQ_i + \sum_{i=j} Q_j/R_{ij} = E^0 + kQ_i + V_i \qquad (38)$$

where E_i is the core binding energy of atom i; E^0, a reference level; Q_i, the net charge on atom i; and V_i, the Madelung potential on atom i. Several improvements have been proposed based mainly on the splitting of net charges into s- and p-components (Ellison and Larcom, 1971; André and Delhalle, 1972). The k-parameter can be theoretically calculated (André and Delhalle, 1972) or semiempirically determined (Clark and Kilcast, 1971a,b; Clark *et al.*, 1972). The correlations obtained by the direct evaluation (Koopmans theorem) and by the model potential are illustrated in Fig. 10.

As it can be seen, the theoretical methods correctly reproduce the trends of experimentally observed $1s_F$ and $1s_C$ peaks. The effects are approximatively additive; each substitution of a hydrogen atom attached to an adjacent atom produces a positive contribution (about 3.0 ± 0.5 eV) whereas the substitution of the atom under consideration gives a negative contribution of the same order of magnitude.

A detailed study of the effects of substitution on valence bands can be carried out. Valence bands have been rarely studied by experimental techniques. The recent proliferation of theoretical valence bands has proved, however, that those bands constitute a real and unique "fingerprint" of the polymer. Indeed, they provide information complementary to and sometimes less ambiguous than that of core levels. Examples of theoretically determined valence bands are given in Fig. 11 as compared to experimental data. We can note that there exists a good correlation between the two. Valuable information for future applications to other fluoropolymers or copolymers is obtained from those results. For example, the energy location of CF and bottom of CC peaks is found to be constant in each spectrum (40 eV for CF, 20 eV for bottom of CC). Practically, those peaks may act directly as internal reference points for experimentalists, especially as the CF band can always be unambiguously assigned.

IV. Conclusion

The main conclusion of this chapter is that quantum mechanical methods are now able to provide significant understanding and insight into the electronic properties of regular polymers. It is hoped that the present development of refined theoretical techniques will soon succeed in correctly interpreting electronic properties of disordered polymers. This would contribute to a better knowledge of large chain systems, which is now urgently needed in physics, chemistry, biology, pharmacology, and even medicine.

ACKNOWLEDGMENT

I should like to express my warm appreciation to Profs. E. Clementi, G. Leroy, P. O. Löwdin and M. Van Meerssche for their inspiring and continuous interest in this work and the orientation they gave to my scientific life. I am further indebted to Prof. J. Delhalle for his friendly cooperation without which much of the work reported here could not have been done; I am very grateful to all my co-workers who contributed substantially to the material of this chapter. Dr. L. A. Burke IV made this manuscript more readable and the help of Miss P. Lonnoy was invaluable in speeding up the typing of the manuscript. I am very grateful to the publishers (D. Reidel, Springer-Verlag and Elsevier) who granted permission to reproduce some figures from my own work.

REFERENCES

André, J. M. (1969). *J. Chem. Phys.* **50**, 1536.

André, J. M. (1970). *Comput. Phys. Commun.* **1**, 391.

André, J. M. (1975). *In* "Electronic Structure of Polymers and Molecular Crystals" (J. M. André and J. Ladik, eds.), p. 1. Plenum, New York.

André, J. M., and Brédas, J. L. (1977). *Chem. Phys.* **10**, 367.

André, J. M., and Delhalle, J. (1972). *Chem. Phys. Lett.* **17**, 145.

André, J. M., and Delhalle, J. (1978). *In* "Quantum Theory of Polymers" (J. M. André, J. Dehalle, and J. Ladik, eds.), p. 1. Reidel Publ., Dordrecht, Netherlands.

André, J. M., and Leroy, G. (1970a). *Chem. Phys. Lett.* **5**, 71.

André, J. M., and Leroy, G. (1970b). *Ann. Soc. Sci. Bruxelles, Ser 1* **84**, 133.

André, J. M., Gouverneur, L., and Leroy, G. (1967a). *Int. J. Quantum Chem.* **1**, 427.

André, J. M., Gouverneur, L., and Leroy, G. (1967b). *Int. J. Quantum Chem.* **1**, 451.

André, J. M., Gouverneur, L., and Leroy, G. (1967c). *Bull. Soc. Chim. Belg.* **76**, 661.

André, J. M., Delhalle, J., Fripiat, J. G., and Leroy, G. (1971a). *Int. J. Quantum Chem.* **5**, 67.

André, J. M., Kapsomenos, G. S., and Leroy, G. (1971b). *Chem. Phys. Lett.* **8**, 195.

André, J. M., Delhalle, J., Kapsomenos, G. S., and Leroy, G. (1972). *Chem. Phys. Lett.* **14**, 485.

André, J. M., Delhalle, J., Delhalle, S., Caudano, R., Pireaux, J. J., and Verbist, J. J. (1973). *Chem. Phys. Lett.* **23**, 206.

André, J. M., Lambert-Gerard, M. E., and Lamotte, C. (1976a). *Bull. Soc. Chim. Belg.* **85**, 745.

André, J. M., Delhalle, J., Demanet, C., and Lambert-Gerard, M. E. (1976b). *Int. J. Quantum Chem.* **S10**, 99.

André, J. M., Fripiat, J. G., Demanet, C., Brédas, J. L., and Delhalle, J. (1978). *Int. J. Quantum Chem.* **S12**, 233.

Barriol, J., and Metzger, J. (1950). *J. Chim. Phys.* **47**, 432.

Barthelat, J. C., and Durand, P. (1978). *Gazz. Chim. Ital.* **108**, 225.

Berthier, G., Lemaire, H., Rassat, A., and Veillard, A. (1965). *Theor. Chim. Acta* **3**, 213.

Beveridge, D. L., and Wun, W. (1973). *Chem. Phys. Lett.* **18**, 570.

Beveridge, D. L., Jano, I., and Ladik, J. (1972). *J. Chem. Phys.* **56**, 4744.

Bloch, F. (1928). *Z. Phys.* **52**, 555.

Bradburn, M., Coulson, C. A., and Rushbrooke, G. S. (1947). *Proc. R. Soc. Edinburgh, Sect A* **62**, 336.

Brédas, J. L., André, J. M., Fripiat, J. G., and Delhalle, J. (1978). *Gazz. Chim. Ital.* **108**, 307.

Christoffersen, R. E. (1971). *J. Am. Chem. Soc.* **93**, 4104.

Christoffersen, R. E. (1972). *Adv. Quantum Chem.* **6**, 333.

Clark, D. T., and Kilcast, D. (1971a). *J. Chem. Soc. B* p. 2273.

Clark, D. T., and Kilcast, D. (1971b). *Nature (London)* **77**, 233.

Clark, D. T., Adams, D. B., and Kilcast, D. (1972). *Chem. Phys. Lett.* **13**, 439.

Clark, D. T., Feast, W. J., Kilcast, D., and Musgrave, W. K. R. (1973). *J. Polym. Sci., Part A-1* **11**, 389.

Clark, D. T., Feast, W. J., Kilcast, D., Musgrave, W. K. R., Modena, M., and Ragazzini, M. (1974). *J. Polym. Sci., Part A-1* **12**, 1049.

Clementi, E. (1971). *J. Chem. Phys.* **54**, 2492.

Cook, D. B., Hollis, P. C., and McWeeny, R. (1967). *Mol. Phys.* **13**, 553.

Cusachs, L. C., and Reynolds, J. W. (1965). *J. Chem. Phys.* **43**, S160.

Davis, D., Gutfreund, H., and Little, W. A. (1976). *Phys. Rev. B* **13**, 4766.

Delhalle, J. (1972). *Ann. Soc. Sci. Bruxelles, Ser. 1* **86**, 277.

Delhalle, J. (1974a). *Int. J. Quantum Chem.* **8**, 201.

Delhalle, J. (1974b). *Chem. Phys.* **5**, 306.

Delhalle, J. (1975). *In* "Electronic Structure of Polymers and Molecular Crystals" (J. M. André and J. Ladik, eds.), p. 53. Plenum, New York.

Delhalle, J., and Delhalle, S. (1975). *Ann. Soc. Sci. Bruxelles, Ser. 1* **89**, 403.

Delhalle, J., Delhalle, S., and André, J. M. (1974a). *Bull. Soc. Chim. Belg.* **83**, 107.

Delhalle, J., André, J. M., Delhalle, S., Pireaux, J. J., Caudano, R., and Verbist, J. J. (1974b). *J. Chem. Phys.* **60**, 595.

Delhalle, J., Delhalle, S., and André, J. M. (1975). *Chem. Phys. Lett.* **34**, 430.

Delhalle, J., André, J. M., Delhalle, S., Pivont-Malherbe, C., Clarisse, F., Leroy, G., and Peeters, D. (1977a). *Theor. Chim. Acta* **43**, 215.

Delhalle, J., Delhalle, S., André, J. M., Pireaux, J. J., Riga, J., Caudano, R., and Verbist, J. J. (1977b). *J. Electron Spectrosc. Relat. Phenom.* **12**, 293.

Delhalle, J., André, J. M., Demanet, C., and Brédas, J. L. (1978a). *Chem. Phys. Lett.* **54**, 186.

Delhalle, J., Thelen, D., and André, J. M. (1979a). *Comput. Chem.* **3**, 1.

Delhalle, J., Montigny, R., Demanet, C., and André, J. M. (1979b). *Theor. Chim. Acta* **50**, 343.

Delhalle, S., Delhalle, J., Demanet, C., and André, J. M. (1975). *Bull. Soc. Chim. Belg.* **84**, 1071.

Del Re, G., Ladik, J., and Biczo, G. (1967). *Phys. Rev.* **155,** 997.

Dewar, M. J. S., Suck, S. H., and Weiner, P. K. (1974). *Chem. Phys. Lett.* **29,** 220.

Dewar, M. J. S., Yamaguchi, Y., and Suck, S. H. (1977). *Chem. Phys. Lett.* **50,** 259.

Duke, B. J., and O'Leary, B. (1973). *Chem. Phys. Lett.* **20,** 459.

Duke, B. J., Eilers, J. E., and O'Leary, B. (1975). *Chem. Phys. Lett.* **32,** 602.

Durand, Ph., Nicolas, G., André, J. M., Burke, L. A., Delhalle, J., unpublished results (1978).

Ellison, F. O., and Larcom, L. L. (1971). *Chem. Phys. Lett.* **10,** 580.

Falk, J. E., and Fleming, R. J. (1973). *J. Phys. C* **6,** 2954.

Falk, J. E., and Fleming, R. J. (1975). *J. Phys. C* **8,** 627.

Frost, A. A. (1967). *J. Chem. Phys.* **47,** 3707.

Fujita, H., and Imamura, A. (1970). *J. Chem. Phys.* **53,** 4555.

Gelius, U. (1972). *In* "Electron Spectroscopy" (D. A. Shirley, ed.), p. 311. North-Holland Publ., Amsterdam.

Ginnard, C. R., and Riggs, W. M. (1972). *Anal. Chem.* **44,** 1310.

Harris, F. E. (1972). *J. Chem. Phys.* **56,** 4422.

Harris, F. E. (1975). *In* "Theoretical Chemistry: Advances and Perspectives" (H. Eyring and D. Henderson, eds.), Vol. 1, p. 147. Academic Press, New York.

Harris, F. E., and Monkhorst, H. (1970). *Phys. Rev., B* **2,** 4400.

Hoffmann, R. (1963). *J. Chem. Phys.* **39,** 1397.

Hückel, E. (1931). *Z. Phys.* **70,** 204.

Hückel, E. (1932). *Z. Phys.* **76,** 628.

Imamura, A. (1970). *J. Chem. Phys.* **52,** 3168.

Keller, H. J. (1977). "Chemistry and Physics of One-dimensional Models." Plenum, New York.

Kertesz, M. (1976). *Acta Phys. Acad. Sci. Hung.* **41,** 107.

Kuhn, H. (1949). *J. Chem. Phys.* **17,** 1298.

Ladik, J. (1965). *Acta Phys. Acad. Sci. Hung.* **18,** 185.

Ladik, J. (1973). *Adv. Quantum Chem.* **7,** 397.

Ladik, J. (1975). *In* "Electronic Structure of Polymers and Molecular Crystals" (J. M. André and J. Ladik, eds.), p. 23. Plenum, New York.

Ladik, J. (1978). *In* "Quantum Theory of Polymers" (J. M. André, J. Delhalle, and J. Ladik, eds.), p. 257. Reidel Publ., Dordrecht, Netherlands.

Ladik, J., and Biczó, G. (1965). *J. Chem. Phys.* **42,** 1658.

Lennard-Jones, J. E. (1937). *Proc. R. Soc. London, Ser. A* **158,** 280.

Little, W. A. (1964). *Phys. Rev.* **134,** A1416.

Löwdin, P. O. (1956a). *Philos. Mag.* [8] **5,** Suppl., 1.

Löwdin, P. O. (1956b). *Adv. Phys.* **5,** 1.

McAloon, B. J., and Perkins, P. G. (1972). *J. Chem. Soc., Faraday Discuss. 2* **N7,** 1121.

McCubbin, W. L. (1971). *Chem. Phys. Lett.* **8,** 507.

McCubbin, W. L., and Manne, R. (1968). *Chem. Phys. Lett.* **2,** 230.

McCubbin, W. L., and Teemull, F. A. (1972). *Phys. Rev., A* **6,** 2478.

Morokuma, K. (1970). *Chem. Phys. Lett.* **6,** 186.

Morokuma, K. (1971). *J. Chem. Phys.* **54,** 962.

Morosi, G., and Simonetta, M. (1971). *Chem. Phys. Lett.* **8,** 358.

Mulliken, R. S. (1965). *J. Chim. Phys.* **46,** 497.

Natta, G., Corradini, P., and Ganis, P. (1960). *Makromol. Chem.* **39,** 238.

O'Leary, B., Duke, B. J., and Eilers, J. E. (1975). *Adv. Quantum Chem.* **9,** 1.

Pal, L., Gruner, G., Janossy, A., and Salyom, J. (1977). "Organic Conductors and Semiconductors." Springer-Verlag, Berlin and New York.

Piela, L., and Delhalle, J. (1978a). *Ann. Soc. Sci. Bruxelles, Ser. 1* **92,** 42.

Piela, L., and Delhalle, J. (1978b). *Int. J. Quantum Chem.,* **13,** 605.

Pireaux, J. J., Riga, J., Caudano, R., Verbist, J. J., André, J. M., Delhalle, J.,and Delhalle, S. (1974). *J. Electron Spectrosc. Relat. Phenom.* **5,** 531.

Pireaux, J. J., Riga, J., Caudano, R., Verbist, J. J., Delhalle, J., Delhalle, S., André, J. M., and Gobillon, Y. (1977). *Phys. Scr.* **16,** 329.

Pople, J. A., Segal, G. A. (1965). *J. Chem. Phys.* **43,** S136.

Pople, J. A., Santry, D. P., and Segal, G. A. (1965). *J. Chem. Phys.* **43,** S129.

Rein, R., Fukuda, A., Win, H., Clarke, G. A., and Harris, F. E. (1966). *J. Chem. Phys.* **45,** 4743.

Schaad, J., and Morell, G. O. (1971). *J. Chem. Phys.* **54,** 1965.

Sichel, J. M., and Whitehead, M. A. (1968). *Theor. Chim. Acta* **11,** 220.

Sieghbahn, K., Nordling, C., Johanson, G., Hedman, J., Hecken, P. F., Hamrin, K., Gelius, U., Bergmark, T., Werme, L. O., Manne, R., and Baer, Y. (1969). "ESCA Applied to Free Molecules," p. 104. North-Holland Publ., Amsterdam.

Snyder, R. G. (1967). *J. Chem. Phys.* **47,** 1316.

Street, G. B., and Greene, R. L. (1977). *IBM J. Res. Dev.* **21,** 99.

Viste, A., and Gray, H. B. (1964). *Inorg. Chem.* **3,** 1113.

Wallace, P. R. (1947). *Phys. Rev.* **71,** 622.

Whitten, J. L. (1963). *J. Chem. Phys.* **39,** 349.

Whitten, J. L. (1966). *J. Chem. Phys.* **44,** 359.

Wolfsberg, M., and Helmholtz, L. (1952). *J. Chem. Phys.* **20,** 837.

Wood, M. H., Barber, M., Hillier, I. H., and Thomas, J. M. (1972). *J. Chem. Phys.* **56,** 1788.

Wood, M. H., Lavery, R., Hillier, I. H., and Barber, M. (1975). *Faraday Discuss. Chem. Soc.* **n60,** 173.

Zupan, J. (1975). *Ann. Soc. Sci. Bruxelles, Ser. 1* **89,** 337.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 12

The Cluster Approach in Theoretical Study of Chemisorption

M. SIMONETTA and A. GAVEZZOTTI
Istituto di Chimica Fisica e Centro CNR
University of Milan
Milan, Italy

I. Introduction

Heterogeneous catalysis, one of the key fields of applied chemical research, has been the subject of great experimental effort. For years chemical intuition and educated guesswork have guided the development and improvement of solid catalysts of different nature. Thermodynamic and kinetic concepts have been used in deriving quantitative laws that account, sometimes satisfactorily, for the energetics and mechanisms of surface processes.

In the large majority of cases, however, molecular detail of the transformations implied in a catalytic reaction is missing. The concept of active site is practically useful, but somewhat ill-defined in terms of structure. As a consequence, the important characteristics of a catalyst, such as activity and selectivity, often correspond to something that can be measured but not predicted. In short, the need for an improved chemical theory of metals, and for a better detailed understanding of chemical

ISBN 0-12-034812-8

interactions between metals and organic fragments, is strongly felt [for recent review articles on the interplay between practical and theoretical aspects of heterogeneous catalysis, see Boudart (1977) and Bonzel (1977); for a preliminary discussion of the favorable conditions for catalysis, see Salem and Leforestier (1978)].

Band theory is the traditional tool by which physicists approach the description of the electronic properties of crystalline metals. The problem of the presence of an adsorbate on a metal surface can therefore be treated, from the same standpoint, as a perturbation of the metal band structure due to the ligand orbitals or, at a further level of sophistication, to a ligand–cluster complex embedded in the metal surface. For a number of reasons, this approach is not suitable to the study of the chemist's crucial concerns, especially ligand–catalyst bond energy, structure dependence of the catalytic activity and selectivity, and the role of fragmentation and particle size; a good example of a study of the influence of cluster size on the catalytic activity can be found in a paper by Gallezot *et al,* (1978). Moreover, catalytic activity is thought to be a very localized effect, depending on the interactions between adsorbate and a few neighboring metal atoms. Therefore, an alternative approach is that of considering the "molecular" properties of the aggregate formed by an organic ligand and a cluster containing a finite number of metal atoms. This way of looking at the matter at once establishes a link between surface chemistry and organometallic chemistry, two fields connected by a so-called "fuzzy interface" (Schaefer, 1977). The natural tool by which these "molecular" properties can be calculated is quantum chemistry as has been applied to the study of free molecules.

At the same time, a wealth of chemical information on the metal–ligand interactions is being provided by a number of relatively new experimental techniques, such as LEED for surface crystallography, UPS for level shifts on going from the free to the adsorbed molecule, IR for adsorbate stretching frequencies (among the many review papers on these subjects, one that stresses the connections with catalysis is by Yates, 1974). Elucidation of the full structure of ordered adsorbate monolayers by such methods seems a goal to be reached in the very near future (for critical consideration of UPS and LEED, see, respectively, Brundle, 1975 and Somorjai, 1977). Also, these techniques are well on their way in the study of stepped and less regular surfaces, opening exciting possibilities by stressing the relationships between different atomic environments and catalytic properties (Chesters and Somorjai, 1975; Buchholz and Somorjai, 1976).

The flexible cluster–ligand model seems to be the most apt to lend theoretical support in the interpretation of experimental evidence; therefore, this article has been focused on this type of approach, and consists essentially of three parts. In Section II a survey of computational methods will be presented, more wiith the aim of outlining the basic principles than of describing the details, that are often well known or can be found easily in standard reference sources. The main section (Section III) reviews the results obtained in the study of the structures of metal surfaces and adsorbates, as well as in predicting or reproducing bond energies. In this section, the performance of the cluster model in mimicking band theory results will be discussed. Section V will deal with the few studies of surface reactivity that have been presented to date, using trajectories on potential surfaces obtained in various ways.

A substantial fraction of the works published so far in the rapidly developing field of the calculation of the properties of chemisorption is still at an exploratory level. The relative merits and limitations of the various approaches to, and levels of simplification of the full problem are still being discussed, as exemplified by the uncertainties about the influence of parametrization for the heavy metal atoms in semiempirical methods, or the size of the clusters to be used (see, e.g., Messmer, 1977). It therefore seems appropriate, at this point, to draw an outline of the most attractive and promising procedures, and to critically compare the results that can be expected from each of them. In this perspective, the literature search has sometimes been carried out more with the aim of representativeness than of completeness.

II. Methods of Calculation

A. Outlook

In what concerns molecular orbital methods, the semiempirical schemes that have been used so far in chemisorption studies are extended Hückel theory (the original paper is by Hoffmann, 1963; a suggested reading for an updated account is in Elian and Hoffmann, 1975) and CNDO (Pople *et al.*, 1965; Pople and Beveridge, 1970). Recently, Schaefer and his co-workers have introduced large-scale *ab initio* computation of the properties of metal clusters and hydrogen adsorption, together with a number of experiences with small metal-containing molecules (see a review in Schaefer, 1977). Goddard and co-workers are exploring the applicability of valence bond to the calculation of surface reactivity (Goddard *et al.*, 1977).

A crucial problem in the application of quantum mechanical formalisms to chemisorption is that, since large numbers of basis orbitals are usually needed, including the d orbitals of the many metal atoms in the clusters, the computation costs may rise very sharply. Extended Hückel theory (henceforth EHT) requires only overlap integrals and one or two matrix diagonalizations. CNDO, while not substantially increasing the number of integrals to be calculated, requires a number (10–20) of matrix diagonalizations to reach self-consistency. Of course, all these problems are more and more important in *ab initio* calculations, in which the number of integrals increases very rapidly with increasing size of the system under consideration.

The Xα-scattered wave formalism (Slater, 1965; Johnson, 1966), originally devised for small molecules, seems to be a convenient compromise between the need for rigorous calculations and the problem of computing time. It is claimed that this method has better performance than semiempirical MO methods, without substantial increase of computing requirements.

Among the non-quantum mechanical methods, the bond energy-bond order (BEBO) method may be mentioned (Johnston and Parr, 1963). It is based on empirical correlations between the quantities of which its name consists; although some encouraging results have been reported (Weinberg and Merrill, 1972, 1973; Weinberg *et al.*, 1974), its use in chemisorption studies suffers from the lack of suitable thermochemical information on metal–metal and metal–adsorbate bonds. Besides, the fundamental assumptions have been the object of criticism for the simplest cases (Jordan and Kaufman, 1975).

B. EHT, CNDO, and Modifications

These methods use molecular orbitals written as (neglecting spin)

$$\phi_i(r) = \sum_{\lambda} c_{i\lambda}\chi_\lambda(r) \tag{1}$$

in which the χ_λ are Slater AOs and the $c_{i\lambda}$ are to be found by a variational procedure. Then one is confronted with solution of

$$\mathrm{Det}(\mathbf{F} - E\mathbf{S}) = 0, \tag{2}$$

where

$$S_{\lambda\sigma} = \langle\lambda|\sigma\rangle, \tag{3}$$

$$F_{\lambda\sigma} = \left\langle\lambda\left|-\tfrac{1}{2}\nabla^2 - \sum_{\alpha}[Z_\alpha/(r-R_\alpha)]\right|\sigma\right\rangle + \tfrac{1}{2}\sum_{\mu,\nu} P_{\mu\nu}\{2\langle\mu\nu|\lambda\sigma\rangle - \langle\mu\sigma|\lambda\nu\rangle\}, \tag{4}$$

$$P_{\mu\nu} = 2\sum_{\mathrm{occ}} c_{i\mu}c_{i\nu}, \tag{5}$$

$$\langle\mu\nu|\lambda\sigma\rangle = \int\chi_\mu(r_1)\chi_\nu(r_1)(r_{12})^{-1}\chi_\lambda(r_2)\chi_\sigma(r_2)\, dV_1\, dV_2. \tag{6}$$

Neglecting all multicenter and almost all bielectronic integrals, Messmer and co-workers have obtained approximate expressions for these matrix elements (Bennett *et al.*, 1971a):

$$F_{\lambda\lambda} \simeq -I_\lambda + \sum_{\substack{\mu \\ \text{atom}}} (P_{\mu\mu} - P_{\mu\mu}^{\text{atom}})\gamma_{\mu\lambda\lambda}, \tag{7}$$

$$\begin{aligned} F_{\lambda\sigma} \simeq -\tfrac{1}{2}S_{\lambda\sigma} \Bigg\{\Bigg[(I_\lambda + I_\sigma) &+ \sum_{\substack{\mu \\ \text{atom B}}} (P_{\mu\mu} - P_{\mu\mu}^{\text{atom B}})\gamma_{\mu\lambda\lambda} \\ &+ \sum_{\substack{\mu \\ \text{atom A}}} (P_{\mu\mu} - P_{\mu\mu}^{\text{atom A}})\gamma_{\mu\sigma\sigma}\Bigg]\Bigg\} \\ &- \tfrac{1}{2} \sum_{\alpha=\text{A,B}} \langle\lambda\|[Z_\alpha/(r - R_\alpha)]\|\sigma\rangle, \end{aligned} \tag{8}$$

$$\gamma_{\mu\lambda\sigma} = \langle\mu\mu|\lambda\sigma\rangle - \tfrac{1}{2}\langle\mu\sigma|\lambda\mu\rangle, \tag{9}$$

with $|\lambda\rangle$ centered on atom B, $|\sigma\rangle$ on atom A; the I's are the valence state ionization potentials (VSIPs) of the corresponding orbitals, Z_α and R_α are the core charge and position of atom α. The last term in the formula for $F_{\lambda\sigma}$ is approximated by multiplying the first by a constant (usually ≈ 1.75). In conventional EHT the differences, $P_{\mu\mu} - P_{\mu\mu}^{\text{atom}}$, between molecular and atomic occupancies are also ignored, yielding

$$H_{\lambda\sigma} \simeq -\tfrac{1}{2}KS_{\lambda\sigma}(I_\lambda + I_\sigma), \qquad H_{\lambda\lambda} \simeq -I_\lambda. \tag{10}$$

This last formulation, therefore, completely neglects electronic interactions.

One well-known shortcoming of standard EHT is the unnatural charge transfer between atoms of different electronegativities. Messmer's treatment was successful with H on graphite, less so for more electronegative atoms (Bennett *et al.*, 1971b). An obvious way of reducing this undesirable effect is manipulation of the VSIPs; this was found necessary especially for edge atoms in metal clusters (Anders *et al.*, 1973). A charge-dependent parametrization is often used to this purpose:

$$I_\lambda = aq_\alpha^2 + bq_\alpha + c. \tag{11}$$

The atomic charges q_α are calculated by using a trial parameter set; then a new set of VSIPs is obtained, and new charges are calculated, until self-consistency is reached. One obvious difficulty is finding the appropriate values of the constants a, b, c.

Anderson and Hoffmann (1974a) have proposed for diatomic molecules some modifications of EHT to improve its performance in the calculation of equilibrium bond distances and dissociation energies by the introduction of a two-body electrostatic interaction energy. By integration of the Hellmann–Feynman formula, the exact molecular energy is found to be the sum of a repulsive term due to interaction between a nucleus and a neutral atom, and an attractive term due to electronic redistributions during formation of the molecule. The first term is explicitly calculated, the second is identified with the total EH energy. This yields good agreement between experimental and calculated equilibrium internuclear distances in a number of diatomics, including H_2, Li_2, BF, CO, SiO, and NaCl. The method has been extended to polyatomic molecules by Anderson (1975a), who showed that, using a Hamiltonian that includes the two-body repulsion energies, one gets

$$H_{ij}^{\alpha\beta} \simeq \tfrac{1}{2}K(E_i^{\alpha} + E_j^{\beta})S_{ij}^{\alpha\beta} + E_{\mathrm{R}}S_{ij}^{\alpha\beta}, \qquad H_{ii}^{\alpha} \simeq E_i^{\alpha} + E_{\mathrm{R}}. \tag{12}$$

The Greek superscripts refer to atoms, the subscripts to atomic orbitals. The E's are the eigenvalues from SCF atomic calculations, E_{R} is the summation of the repulsion energies, as described above for diatomics; since this is a constant for a given molecular conformation, Eq. (12) is strictly similar to the standard EH (extended Hückel) recipe. With these adjustments, it was also found necessary, to improve the fit to experimental equilibrium distances and binding energies, to use $K = 2.25$, and to multiply the off-diagonal H matrix elements by $\exp(-0.13R)$, R being the α–β internuclear separation.

In the CNDO approximation, the set of equations (3)–(6) is reduced to:

$$F_{\mu\mu} = H_{\mu\mu} + (P_{\mathrm{AA}} - \tfrac{1}{2}P_{\mu\mu})\gamma_{\mathrm{AA}} + \sum_{\mathrm{B}\neq\mathrm{A}} (P_{\mathrm{BB}}\gamma_{\mathrm{AB}} - V_{\mathrm{AB}}), \tag{13}$$

$$F_{\mu\nu} = \beta_{\mathrm{AB}}^{0}S_{\mu\nu} - \tfrac{1}{2}P_{\mu\nu}\gamma_{\mathrm{AB}}, \tag{14}$$

$$\begin{aligned} H_{\mu\mu} &= \langle\mu| - \tfrac{1}{2}\nabla^2 - [Z_{\mathrm{A}}/(r - R_{\mathrm{A}})]|\mu\rangle \\ &\simeq -\tfrac{1}{2}(I_{\lambda} + A_{\lambda}) - Z_{\mathrm{A}} - \tfrac{1}{2}\gamma_{\mathrm{AA}}, \end{aligned} \tag{15}$$

$$P_{\mathrm{AA}} = \sum_{\mu,\,\mathrm{atom\ A}} P_{\mu\mu} \tag{16}$$

$$\gamma_{\mathrm{AB}} = \iint S_{\mathrm{A}}^{2}(1)(r_{12})^{-1}S_{\mathrm{B}}^{2}(2)\, d\tau_1\, d\tau_2, \tag{17}$$

$$\beta_{\mathrm{AB}}^{0} = \tfrac{1}{2}K(\beta_{\mathrm{A}}^{0} + \beta_{\mathrm{B}}^{0}), \tag{18}$$

$$V_{\mathrm{AB}} = \int S_{\mathrm{A}}^{2}(1)(Z_{\mathrm{B}}/r_{1\mathrm{B}})\, d\tau_1, \tag{19}$$

$K \approx 1$; Greek subscripts refer to orbitals, A and B refer to atoms. Electron interactions are included, and since the $F_{\mu\nu}$'s depend on the $P_{\mu\nu}$'s, the calculation is repeated to self-consistency. For s and p orbitals, with exponents determined by Slater's rules, a single average Coulomb integral is used (approximation 3 of Pople *et al.*, 1965). If d orbitals are to be included, a modification must be introduced by using

$$\gamma_{sd} \simeq \gamma_{ss'}, \tag{20}$$

i.e., the integrals are evaluated by using an s-type orbital (s′) with the same screening coefficient as the d orbital. This procedure is discussed by Santry and Segal (1967) and Baetzold (1971).

Among the troubles encountered in using CNDO in chemisorption model studies, convergence problems of unknown origin in the SCF procedure have been reported both by Bennett *et al.* (1971b) and by Hayns (1975). The former authors have also proposed a very interesting stratagem to include partially the periodicity of the substrate lattice. Essentially, matrix elements are modified to treat as nearest neighbors and next-to-nearest neighbors some atoms of the cluster that would be such if translational symmetry were considered. Smooth charge distributions were thus obtained for the carbon atoms of a graphite cluster.

C. Binding Energies in EHT and CNDO

In EHT the total molecular energy is calculated as the sum of the energies of the occupied orbitals. Both in EHT and in CNDO, binding energies are usually calculated as the difference between the total electronic energy of the molecule and the energy of the isolated atoms:

$$\text{binding energy (BE)} = E_{\text{tot}} - \sum_i E_{\text{atom},i}. \tag{21}$$

The reliability of the values thus obtained is very small. In an attempt to improve it, CNDO energies have been partitioned into monocentric and bicentric contributions, and the latter have been compared to known thermochemical data to obtain rescaling factors (Companion, 1973). In any case, the adsorption binding energies are then usually calculated as the difference between the total energy of the cluster–adsorbate system and the energies of the separate fragments:

$$\text{adsorption BE} = E_{\text{clust+ads}} - E_{\text{clust}} - E_{\text{ads}}. \tag{22}$$

The absolute values of the adsorption BEs are also scarcely reliable. For EHT the adsorption BE at very large separation is contaminated by unnatural electron transfers between cluster and adsorbate. Therefore, only the shapes and relative energies of BE curves can be discussed. Absolute energy values are also largely dependent on the parametrization. The performance of the various stratagems that have been described above to cope with these difficulties will be examined in the following sections.

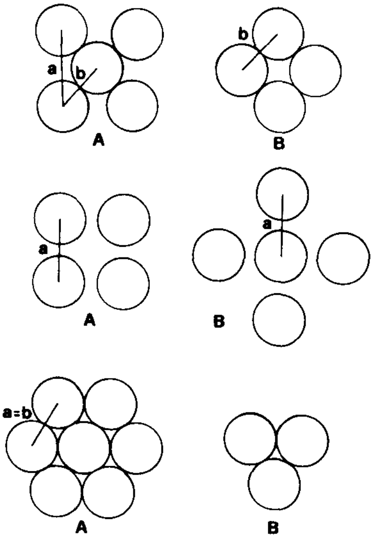


Fig. 1. Clusters formed by the superimposition of A and B layers: *a* is the cell edge, *b* the nearest-neighbor distance. From top to bottom: (100) fcc; (100) bcc; (0001) hcp.

D. How Clusters Are Built

In the cluster approach, the metal surface and first layers are to be represented by a "molecule" formed by a number of metal atoms, whose reciprocal positions can be described in terms of the edges of the crystallographic unit cell in face-centered cubic (fcc), body-centered cubic (bcc) or hexagonal close packed (hcp) lattices. All computation methods require as a starting point the Cartesian coordinates of all these atoms; finding these coordinates can be a trivial task for clusters oriented along coordinate planes (see Fig. 1) or directions for low Miller indices, but becomes more and more difficult with increasing values of the indices or complexity of the lattice. For computational purposes, it is always extremely convenient to have the face to be investigated as one of the coordinate planes in the Cartesian reference system; this implies a rotation of the cell axes, which in the case of the cubic lattices can be performed, in general, as follows. Let a be the cell edge, hkl the Miller indices of the plane under consideration, and

$$Ax + By + Cz = p \tag{23}$$

the equation of a plane (a, x, y, z, and p in Å). Using the definition of Miller indices, one readily finds that three points lying on the desired plane have the coordinates:

$$(a/h,\quad 0,\quad 0);\qquad (0,\quad a/k,\quad 0);\qquad (0,\quad 0,\quad a/l). \tag{24}$$

which immediately gives

$$A = hp/a, \qquad B = kp/a, \qquad C = lp/a. \tag{25}$$

But since A, B, and C are the direction cosines of the normal to the plane, it must be that

$$(A^2 + B^2 + C^2) = 1, \tag{26}$$

$$\begin{gathered} A = h/(h^2 + k^2 + l^2)^{1/2}, \qquad B = k/(h^2 + k^2 + l^2)^{1/2}, \\ C = l/(h^2 + k^2 + l^2)^{1/2} \end{gathered} \tag{27}$$

The equation of a line in space, passing by the origin, can be written as

$$x/\cos\alpha = y/\cos\beta = z/\cos\gamma \tag{28}$$

and therefore as

$$x/h = y/k = z/l. \tag{29}$$

Therefore, the point $x = h$, $y = k$, $z = l$ is a point on the normal to the *hkl* plane. Now, if a rotation can be found that brings, say, the x axis of the original reference system to coincide with this normal, the *yz* plane of the transformed system will be the surface plane, and the planes parallel to it will be the layers as seen by looking down the normal to the *hkl* plane. This rotation is shown in Fig. 2; a rotation of χ degrees around the original z axis, followed by a rotation of ζ degrees around the y axis will fulfill the task. Then, letting $\mathbf{x}$ be the original lattice point vectors and $\mathbf{x}'$ the transformed coordinates, one gets (refer to Fig. 2):

$$\cos\chi = h/(h^2 + k^2)^{1/2}, \qquad \sin\chi = k/(h^2 + k^2)^{1/2}, \tag{30}$$

$$\cos\zeta = (h^2 + k^2)^{1/2}/(h^2 + k^2 + l^2)^{1/2}, \qquad \sin\zeta = l/(h^2 + k^2 + l^2)^{1/2}, \tag{31}$$

$$\boldsymbol{\Omega}^{\chi} = \begin{pmatrix} \cos\chi & \sin\chi & 0 \\ -\sin\chi & \cos\chi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\Omega}^{\zeta} = \begin{pmatrix} \cos\zeta & 0 & -\sin\zeta \\ 0 & 1 & 0 \\ \sin\zeta & 0 & \cos\zeta \end{pmatrix}, \tag{32}$$

$$\boldsymbol{\Omega} = \boldsymbol{\Omega}^{\zeta}\boldsymbol{\Omega}^{\chi}, \tag{33}$$

$$\mathbf{x}' = \boldsymbol{\Omega}\mathbf{x}. \tag{34}$$

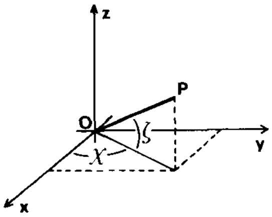


Fig. 2. Two rotations that bring the x axis to coincide with the *OP* direction (see text).

Matrix $\boldsymbol{\Omega}$ therefore represents the overall rotation that brings the x axis to coincide with the normal to the *hkl* plane. Figure 3 shows some clusters corresponding to high-Miller indices surfaces.

E. The SCF–Xα-Scattered Wave Method

This method is peculiar in that the experience gained in the study of energy bands in metals and crystals was later translated to the treatment of molecules, ions, and clusters. The suggestion for the extension to calculations of the electronic structures of systems including a relatively small number of atoms (as compared to this number in a real crystal) is due to

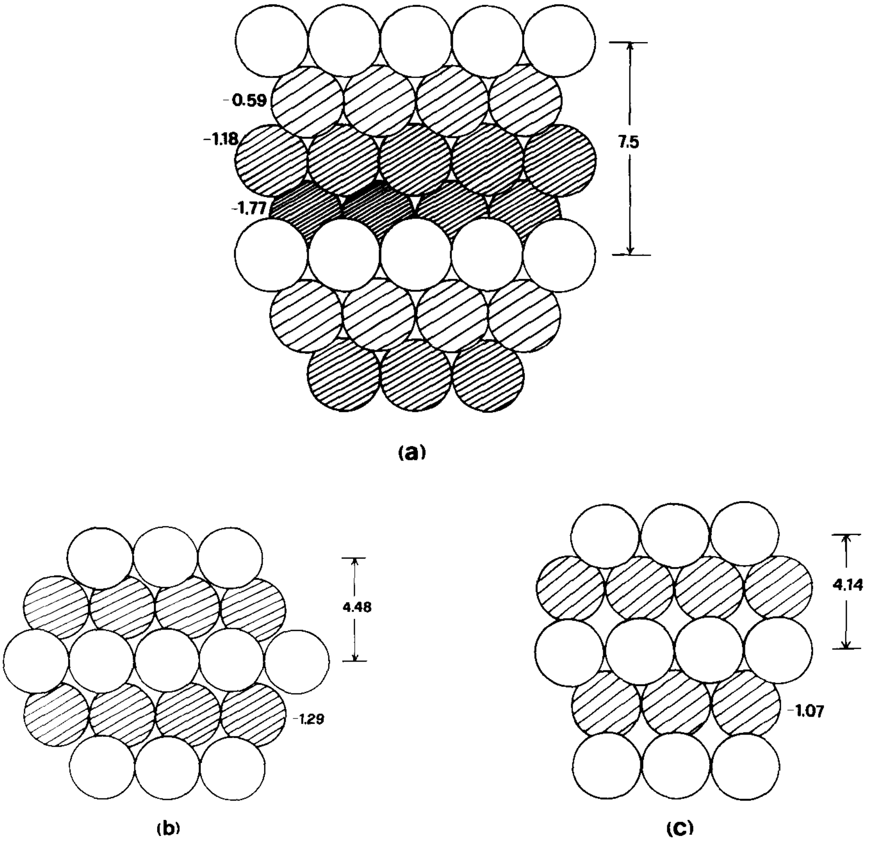


Fig. 3. Clusters corresponding to high-Miller indices planes: (a) (221) fcc nickel; (b) (211) bcc tungsten; (c) (311) fcc nickel. Circles are drawn to scale; open circles at zero level, shaded-circle depths as marked (Å).

Slater (1965). The development was started by Johnson (1966). In a sense this theory can be considered a bridge between the solid-state physics and the quantum chemistry approach. The extensive use of this type of calculation may help to increase the communicability between physicists and chemists engaged in surface science, which is badly needed at the moment. The theory has been reviewed many times in recent years, and also in a volume of this series (Johnson, 1972), but for the sake of completeness a brief presentation is included in this article.

A number of assumptions are embodied in the method. Slater (1951) pointed out that a great simplification of the Hartree–Fock (HF) method could be achieved by substituting a local exchange potential for the nonlocal exchange operator. This potential was assumed to take the value it should have in a free electron gas with local density equal to the actual density in each position, namely

$$V_{\mathrm{Xs}}(\mathbf{r}) = -6(3/8\pi)\rho(\mathbf{r})]^{1/3}, \tag{35}$$

where $\rho(\mathbf{r})$ is the electron charge density at the position $\mathbf{r}$. Since a different approach (Gaspar, 1954; Kohn and Sham, 1965) led to a different approximate potential:

$$V_{\mathrm{XGKS}}(\mathbf{r}) = \tfrac{2}{3}V_{\mathrm{Xs}}(\mathbf{r}), \tag{36}$$

the following potential has been finally proposed (Slater, 1972):

$$V_{\mathrm{X}\alpha}(\mathbf{r}) = \alpha V_{\mathrm{Xs}}(\mathbf{r}), \tag{37}$$

where α is a parameter with a value between $\frac{2}{3}$ and 1. Different techniques for its optimization are available (Slater and Wood, 1971).

In the investigation of the electronic structure of a cluster of atoms replacement of the nonlocal exchange terms in the HF equations by the local operator $V_{\mathrm{X}\alpha}(\mathbf{r})$ leads to a one-electron Schrödinger equation

$$[\nabla^2 + V(\mathbf{r})]\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{38}$$

(in atomic units), where

$$V(\mathbf{r}) = V_{\mathrm{C}}(\mathbf{r}) + V_{\mathrm{X}\alpha}(\mathbf{r}), \tag{39}$$

i.e., the potential is given by the sum of the Coulomb and the local exchange contributions.

Let us introduce next the "muffin-tin" approximation to the potential. For a cluster of N atoms the potential $V(\mathbf{r})$ is replaced by a set of spherically symmetric nonoverlapping potentials $V^i(\mathbf{r})$ centered at each atomic site $\mathbf{R}_i$. In addition, the entire cluster is surrounded with a sphere outside which $V(\mathbf{r})$ is spherically symmetric. For all other locations $V(\mathbf{r})$ is taken to have a constant value $\bar{V}$, determined as the volume average of $V(\mathbf{r})$ in this region. As seen in Fig. 4a, the space is partitioned into N spherical regions, atomic regions (I), plus an interatomic region (II) and extramolecular region (III). Thus the "muffin-tin" potential takes the form

$$V(\mathbf{r}) = \begin{cases} V^j(r_j) & \text{for atomic regions } r_j = |\mathbf{r} - \mathbf{R}_j| \leq b_j,\ j = 1, 2, \ldots, N; \\ V^0(r_0) & \text{for extramolecular region } r_0 = |\mathbf{r} - \mathbf{R}_0| > b_0; \\ \bar{V} & \text{for the interatomic region.} \end{cases} \tag{40}$$

At the start the SCF–Xα potentials centered at positions $\mathbf{R}_j$ and their associated electronic charge densities are calculated for all the atoms in the cluster (Herman and Skillman, 1963), and an intermediate potential is obtained as a superposition of such atomic potentials:

$$V(\mathbf{r}) = \sum_j V^j(r_j). \tag{41}$$

Then the potential is made spherically symmetric within each atomic sphere and in region III, while its average value $\bar{V}$ is obtained for region II:

$$\bar{V} = (\Omega_{\mathrm{II}})^{-1} \int_{\Omega_{\mathrm{II}}} V(\mathbf{r})\, d\mathbf{r}, \tag{42}$$

where Ω_{II} is the volume of the interatomic region. It should be noticed that each atomic region does not include only the contribution of the atom located at the center but also the spherically averaged contributions from all other atoms.

A detailed description of the procedure for the construction of the self-consistent potential field has been given (Weinbergen and Schwartz, 1975). The two previously described techniques, namely the use of the Xα statistical exchange term and of the muffin-tin potential, can be combined with the scattered wave method (Korringa, 1947; Kohn and Rostoker, 1954). In each spherical atomic region of radius b_j and in the outer sphere the wave function of Eq. (38) is expanded as

$$\text{(a)}\quad \psi_j(\mathbf{r}) = \sum_L C_L^j R_l^j(r_j, E) Y_L(\mathbf{r}_j), \qquad j = 1, 2, \ldots, N, \quad 0 \leq r_j < b_j;$$

$$\text{(b)}\quad \psi_0(\mathbf{r}) = \sum_L C_L^0 R_l^0(r_0, E) Y_L(\mathbf{r}_0), \qquad b_0 < r_0 < \infty, \tag{43}$$

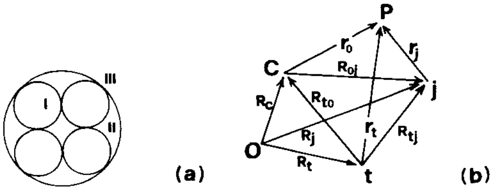


Fig. 4. (a) Regions in the Xα method. (b) Some relevant vectors (see text, Section II,E).

where $L = l, m$, C_L^j are coefficients to be determined, $Y_L(\mathbf{r})$ are real spherical harmonics, and $R_l(r, E)$ are solutions of the radial Schrödinger equation in region j or III:

$$\left[-\frac{1}{2}\frac{d}{dr}r^2\frac{d}{dr} + \frac{l(l+1)}{r^2} + V(r) - E\right]R_l(r, E) = 0, \tag{44}$$

where $r = r_j$ or r_0, $V = V^j$ or V^0, $R = R^j$ or R^0 for case a and b, respectively. The solutions of Eq. (44) can be obtained by numerical integration inward in the extramolecular region, outward in the atomic regions, imposing the condition of regularity for $R_l^j(r, E)$ at $\mathbf{R}_j$ and for R_l^0 at ∞. In region II the Schrödinger equation is

$$(-\nabla^2 + K^2)\psi_{\text{II}}(\mathbf{r}) = 0, \tag{45}$$

where $K^2 = \bar{V} - E$. In the following we shall write the equations for the more usual case, where $\bar{V} > E$, but the corresponding equations for $E > \bar{V}$ can be easily derived along the same lines.

Focusing on a particular scatterer t, the solution of Eq. (45) can be written as the sum of an incident wave and a scattered wave:

$$\psi_{\text{II}}(\mathbf{r}) = \psi_{\text{inc}}^t(\mathbf{r}) + \psi_{\text{sc}}^t(\mathbf{r}) \tag{46}$$

with

$$\psi_{\text{inc}}^t(\mathbf{r}) = \sum_L B_L^t i_l(Kr_t) Y_L(\mathbf{r}_t), \tag{47}$$

$$\psi_{\text{sc}}^t(\mathbf{r}) = \sum_L A_L^t k_l^{(1)}(Kr_t) Y_L(\mathbf{r}_t), \tag{48}$$

where $i_l(x)$ and $k_l^{(1)}(x)$ are modified spherical Bessel functions and Hankel functions of the first kind:

$$i_l(x) = i^{-l} j_l(ix), \tag{49}$$

$$k_l^{(1)}(x) = -i^{-l} h_l^{(1)}(ix), \tag{50}$$

where j_l and $h_l^{(1)}$ are ordinary spherical Bessel functions and Hankel functions of the first kind.

However, $\psi_{\text{II}}(\mathbf{r})$ can also be written in the usual multicenter representation

$$\psi_{\text{II}}(\mathbf{r}) = \sum_{j=1}^{N} \sum_L A_L^j k_l^{(1)}(Kr_j) Y_L(\mathbf{r}_j) + \sum_L A_L^0 i_l(Kr_0) Y_L(\mathbf{r}_0). \tag{51}$$

Now we want to identify $\psi_{\text{II}}(\mathbf{r})$ given in Eqs. (46) and (51). This can be accomplished by means of the following expansion theorems:

$$k_l^{(1)}(K|\mathbf{r}_2 - \mathbf{r}_1|)Y_L(\mathbf{r}_2 - \mathbf{r}_1) = 4\pi \sum_{L'} (-1)^{l+l'} \sum_{L''} I_{L''}(L, L') \times k_{l''}^{(1)}(Kr_1)Y_{L''}(\mathbf{r}_1)i_{l'}(Kr_2)Y_{L'}(\mathbf{r}_2), \quad (r_1 > r_2), \tag{52}$$

$$k_l^{(1)}(K|\mathbf{r}_2 - \mathbf{r}_1|)Y_L(\mathbf{r}_2 - \mathbf{r}_1) = 4\pi \sum_{L'} (-1)^{l+l'} \sum_{L''} I_{L''}(L, L') \times k_{l'}^{(1)}(Kr_2)Y_{L'}(\mathbf{r}_2)i_{l''}(Kr_1)Y_{L''}(\mathbf{r}_1), \tag{53}$$

$$i_l(K|\mathbf{r}_2 - \mathbf{r}_1|)Y_L(\mathbf{r}_2 - \mathbf{r}_1) = 4\pi \sum_{L'} (-1)^{l+l'} \sum_{L''} I_{L''}(L, L') \times i_{l''}(Kr_1)Y_{L''}(\mathbf{r}_1)i_{l'}(Kr_2)Y_{L'}(r_2). \tag{54}$$

The integrals

$$I_{L''}(L, L') = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta \, d\theta Y_{L''}(\theta, \varphi)Y_L(\theta, \varphi)Y_{L'}(\theta, \varphi) \tag{55}$$

are called "Gaunt numbers" and are zero unless

$$|l - l'| \le l'' \le |l + l'|, \tag{56}$$

$$l + l' + l'' = 2g, \qquad g \text{ an integer.} \tag{57}$$

They are related to the coefficients appearing in the theory of multiplets, for which tables can be found in the literature (Condon and Shortley, 1951). A detailed derivation of the expansion theorem can be found in the Appendix of an article by Johnson (1972).

From Fig. 4b, it is clear that

$$\mathbf{r}_j = \mathbf{r}_t - \mathbf{R}_{tj}, \tag{58}$$

$$\mathbf{r}_0 = \mathbf{r}_t - \mathbf{R}_{t0}. \tag{59}$$

Substituting $\mathbf{r}_j$ and $\mathbf{r}_0$ from Eqs. (58) and (59) into Eq. (51), making use of the expansion theorem, and comparing with Eqs. (46)–(48), we obtain the following set of equations (one for each value of t and L):

$$B_L^t = \sum_j \sum_{L'} G_{LL'}^{tj}(E)A_{L'}^j + \sum_{L'} S_{LL'}^{t0}(E)A_{L'}^0, \qquad t = 1, 2 \ldots, N, \tag{60}$$

where

$$G_{LL'}^{tj}(E) = (1 - \delta_{tj})4\pi(-1)^{l+l'} \sum_{L''} I_{L''}(L, L')k_{l''}^{(1)}(R_{tj})Y_{L''}(\mathbf{R}_{tj}) \tag{61}$$

and

$$S_{LL'}^{t0}(E) = 4\pi(-1)^{l+l'} \sum_{L''} I_{L''}(L, L')i_{l''}(KR_{t0})Y_{L''}(\mathbf{R}_{t0}). \tag{62}$$

The same procedure can be applied to the extramolecular region, i.e., we can write

$$\begin{aligned}\psi(\mathbf{r}) &= \psi^0_{\mathrm{inc}}(\mathbf{r}) + \psi^0_{\mathrm{sc}}(\mathbf{r}) \\ &= \sum_L B^0_L k_l(Kr_0) Y_L(\mathbf{r}_0) \sum_L A^0_L i_l(Kr_0) Y_L(\mathbf{r}_0).\end{aligned} \tag{63}$$

Again by means of the expansion theorem we obtain the equations (one for each value of L)

$$B^0_L = \sum_j \sum_{L'} S^{0j}_{LL'}(E) A^j_{L'}. \tag{64}$$

Now we wish to match continuously the wavefunction from Eq. (43a) and its first derivative at the boundary of sphere j ($r_j = b_j$) with the field in the interatomic region [Eq. (46)]. From equality of logarithmic derivatives we obtain

$$\frac{1}{R^j_l(b_j, E)} \left[\frac{dR^j_l(r_j, E)}{dr_j} \right]_{r_j = b_j} = \frac{B^j_L i_l'(Kb_j) + A^j_L k_l'(Kb_j)}{B^j_L i_l(Kb_j) + A^j_L k_l(Kb_j)}. \tag{65}$$

From this equation we can derive the following expression

$$A^j_l = t^j_l(E) B^j_L, \tag{66}$$

where

$$t^j_l(E) = \frac{[i_l(Kb_j), R^j_l(b_j, E)]}{[k^{(1)}_l(Kb_j), R^j_l(b_j, E)]}. \tag{67}$$

The square brackets in Eq. (67) symbolize Wronskian forms:

$$[f(x), g(x)] = f(x) \frac{dg(x)}{dx} - g(x) \frac{df(x)}{dx}. \tag{68}$$

From Eqs. (63) and (43b) we obtain in a similar way

$$A^0_L = t^0_l(E) B^0_L \tag{69}$$

with

$$t^0_l(E) = \frac{[k^{(1)}_l(Kb_0), R^0_l(b_0, E)]}{[i_l(Kb_0), R^0_l(b_0, E)]}. \tag{70}$$

So we have connected the amplitudes of the scattered wave and the amplitudes of the incident wave.

Matching the amplitudes of the waves at different boundaries [Eqs. (43), (46), (63)] we arrive at the following expressions:

$$A^j_L = Kb^2_j(-1)^{l+1}[i_l(Kb_j), R^j_l(b_j, E)] C^j_L, \tag{71}$$

$$A^0_L = Kb^2_0(-1)^{l+1}[R^0_l(b_0, E), k^{(1)}_l(Kb_0)] C^0_L, \tag{72}$$

where we made use of the identity (Schiff, 1949):

$$[i_l(Kb), k_l(Kb)] = (-1)^{l+1}/(Kb^2). \tag{73}$$

Now we must have the incident wave corresponding to any site $j > 0$ equal to the superposition of the scattered waves from all other scatterers and from the outer sphere, while the incident wave in the extramolecular region must equal the sum of the scattered waves from all scattering regions, $j > 0$. So combining Eqs. (66) and (69) with Eqs. (60) and (64), respectively, we obtain the following set of linear homogeneous equations:

$$\sum_j \sum_{L'} [T^{-1}(E)]^{tj}_{LL'} A^j_{L'} - \sum_{L'} S^{t0}_{LL'}(E) A^0_{L'} = 0, \tag{74}$$

$$\sum_j \sum_{L'} S^{0j}_{LL'}(E) A^j_{L'} - \sum_{L'} \delta_{LL'} [t^0_l(E)]^{-1} A^0_{L'} = 0, \tag{75}$$

where

$$[T^{-1}(E)]^{tj}_{LL'} = \delta_{tj}\, \delta_{LL'} [t^t_l(E)]^{-1} - G^{tj}_{LL'}(E). \tag{76}$$

The set of equations (74) and (75) allows nontrivial solutions when the corresponding secular determinant is zero:

$$\begin{Vmatrix} [T^{-1}(E)]^{tj}_{LL'} & -S^{t0}_{LL'}(E) \\ S^{0j}_{LL'}(E) & -\delta_{LL'}[t^0_l(E)]^{-1} \end{Vmatrix} = 0. \tag{77}$$

Since all matrix elements in the determinant are energy dependent, the determinant is evaluated as a function of E and the zeros are to be found by interpolation. For each of these molecular orbital energy values the set of equations (74) and (75) can be solved and the amplitudes A^j_L and A^0_L of the scattered waves are found. From these the amplitudes of the incident waves B^j_L and B^0_L and the coefficients C^j_L and C^0_L in Eq. (43) are calculated. Of course, the value of the determinant in Eq. (77) for a given E depends on the number of L terms that are included in the calculations. The set $L = 0, 1, 2$ is sufficient for most applications, including molecules containing transition metal atoms.

III. Results

A. Results of Semiempirical Methods

The information about chemisorption that can be obtained from theoretical calculations is, in principle, very large. First, the bonding energy can be calculated, and therefore the relative stabilities of various positions of the adsorbate over the surface can be assessed. Besides, the geometry of the adsorption can be further detailed to include distance from the surface and even molecular relaxation of the adsorbate after adsorption; finally, adsorbate–adsorbate interactions can be discussed by modeling the surface layer. Energy barriers to diffusion or reorientation of the molecule on the surface are also within the reach of the methods outlined in the preceding sections.

Any MO calculation lends itself to some kind of breakdown of the energy components and of the electronic populations. Therefore, bond-strengthening and bond-weakening phenomena occurring upon adsorption can be investigated. The composition of the molecular orbitals gives information about which atomic orbitals are used in forming adsorbate-surface bonds, and hence about orientation in the adsorption process. Moreover, the levels obtained by MO calculations can be compared with photoelectron spectroscopic assignments, and shifts induced by adsorption can be rationalized. Calculated atomic charges may be considered in connection with work function changes following adsorption.

Finally, even very crude models can reproduce to some extent the results of more accurate band calculations. Fermi levels, bandwidths, and density of states have analogs in molecular orbital results; a discussion of the effects of parametrization and cluster size on the reproduction of the electronic properties of metals can be found, e.g., in Anderson (1978a), Baetzold (1978), and references therein.

B. EHT Results

1. *Adsorption of Single Atoms*

A comprehensive set of calculations on chemisorption using a cluster model with EHT was done by Fassaert *et al.* (1972). To study H adsorption on Ni, standard EHT was used, with double-zeta d functions. Mulliken population analysis (Mulliken, 1955) was used. The following analogies between observed and calculated quantities were established:

d *bandwidth*—energy difference between highest and lowest MO with strong d character;

Fermi level—energy of the highest occupied MO;

cohesion energy—binding energy for the cluster, divided by the number of atoms in the cluster, and renormalized by multiplying by the ratio of the coordination number in the metal to the average coordination number in the cluster;

holes in d *band*—defining

$$N^{\mathrm{d}} = \sum_{\substack{K \\ \text{d orbitals}}} N_K^{\mathrm{d}} \tag{78}$$

as the amount of d character of the electrons, the summation being taken over the charges in the d orbitals of all atoms, and defining

$$n^{\mathrm{d}} = N^{\mathrm{d}}/\text{No. of atoms in cluster} \tag{79}$$

as the average d character per atom, then

$$n_{\mathrm{h}} = 10 - n^{\mathrm{d}} \tag{80}$$

is the number of holes in the d band.

Table I gives some results, as compared with experiment. Concerning the position of the adsorbate, the H atom was found to bind preferentially on top of a Ni atom, rather than in a fourfold hole. This result was found to be consistent with a general tendency of the adsorbed H atoms to have as few neighbor metal atoms as possible, also on stepped surfaces (Fassaert and van der Avoird, 1976b). The preferred adsorption site and the magnitude of the binding energy were found to be scarcely sensitive to cluster size, and even similar to those obtained from periodic crystal calculations (Fassaert and van der Avoird, 1976a). The same preferred adsorption site was obtained in calculations for H atom chemisorption on W (Anders *et al.*, 1973). However, the presence of p orbitals on N was found to change this picture (Anders *et al.*, 1975), the N atom preferring a four-coordinate adsorption site. Table II has the details of the relative adsorption energies and of the atomic charges. In these calculations, extensive adjustment of the VSIPs of the edge atoms of the clusters was found necessary to reduce charge transfers.

Bennett *et al.* (1971a) have used standard EHT to study the adsorption of atomic H on graphite. They used three different models of the graphite layer (Fig. 5), whose cohesive energies are shown in Table III. The highest cohesive energy was for C—C = 1.40 Å. Other calculated quantities are also reported in Table III. Figure 5b shows the geometry of adsorption; the most stable site (A in this figure) is on top of a carbon atom. Barriers to diffusion are: C to A, 30.9 kcal/mole; B to A, 2.1 kcal/mole.

TABLE I

CALCULATED AND EXPERIMENTAL ELECTRONIC PROPERTIES OF NICKEL[a,b]

Property	Calc. by EH	From band calc. or exptl.
d bandwidth	1.59 to 1.81	2.7 to 5
Fermi level	−7.64 to −7.72	−4.75 to −5.22
Holes in d band	0.54 to 0.68	0.6
Cohesion energy	3.8 to 4.7	4.4

[a] Fassaert *et al.* (1972).

[b] All energy values are in eV.

TABLE II

CALCULATED PROPERTIES FOR ADSORPTION OF H AND N ON W[a]

Atom	Site	Binding energy	Distance from surface	Charge on adatom
H	1CN	−3.00	1.65	−0.36
	2CN	−2.74	1.96	−0.26
	5CN	−1.72	1.61	+0.04
N	1CN	−2.85	1.88	−1.13
	2CN	−3.15	2.13	−1.36
	5CN	−3.25	2.05	−0.04
	exptl.	−6.6	—	—

[a] From Anders *et al.* (1973, 1975). Energies are in eV, distances in Å, charges in electrons.

2. *Adsorption of Diatomic Molecules*

A reference work in the study of the interaction of first period diatomic molecules with metal clusters has been published by Anderson and Hoffmann (1974b). The charge transfer between cluster and adsorbate at infinite separation is typified by the two extreme cases shown in Fig. 6; the Li—Li bond is broken after adsorption because the bonding σ orbital is emptied, while the F—F bond is weakened by population of a σ^* orbital. At bonding cluster–adsorbate distances, orbital overlap and mixing may mitigate such extreme effects for the intermediate cases of B—B and C—C, but the trends in tendency to dissociative chemisorption are found to agree with predictions based on simple electronegativity arguments. A similar effect is the basis of Baetzold's (1975) rationalization of the calculated (by EHT) barrier to dissociation of H_2 on Pd or Ag (Fig. 7). De-

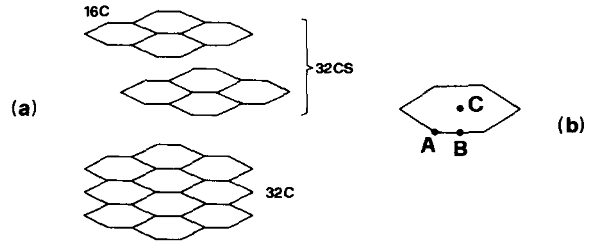


Fig. 5. (a) Graphite clusters; 16 C atoms, 32 C atoms, and 32 C atoms stacked. (b) Labeling the adsorption sites over a hexagon of C atoms.

TABLE III

CALCULATED AND EXPERIMENTAL PROPERTIES OF GRAPHITE[a]

Cluster[b]	Binding energy	Total bandwidth	Total filled bandwidth
16C	4.17	11.61	17.8
32C	4.32	12.3	17.0
32C stacked	4.17	—	—
Experimental or band calc.	5	13.6	14.3

[a] All entries except last row calculated by EH by Bennett *et al.* (1971a). Values are in eV.

[b] See Fig. 5.

stabilization of the σ level consequent to stretching of the H—H bond gives rise to a barrier, but as the σ^* level drops below the HO MO of the cluster, charge transfer takes place and the binding energy increases. The problem of unrealistic binding energy increase is a well-known difficulty at large cluster–adsorbate distances (Fassaert *et al.*, 1972; see also Gavezzotti and Simonetta, 1977). For this reason, Anderson and Hoffmann have chosen to discuss the electronic properties of the molecule + cluster system at a fixed separation of 2 Å.

An interesting remark is that the surface is "tightly packed" with available metal d orbitals, so that the formation of the chemisorption bond is rather insensitive to orientation. Relative strengths of these bonds may, however, be sensitive to orientation. Adsorption on edge atoms was found to be favorable; Fig. 8 shows some possible modes of adsorption for N_2 on a stepped W cluster. Finally, the level shifts calculated for ethylene after adsorption on Ni compared favorably with photoelectron spectral data.

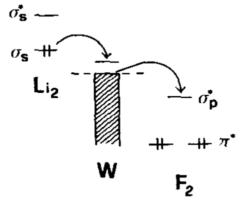


FIG. 6

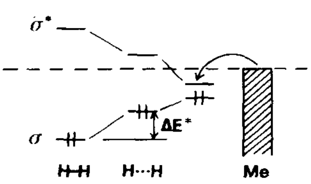


FIG. 7

Fig. 6. Schematic diagram of the relative positions of some relevant Li_2 and F_2 molecular orbitals and of the W band. Electron transfers at infinite molecule-surface separation are shown by curved arrows (Anderson and Hoffmann, 1974b).

Fig. 7. Schematic diagram of level shifts for H_2 dissociation on a metal (after Baetzold, 1975). $\Delta E^{\dagger}$ is the activation energy due to H–H stretching. The curved arrow symbolizes electron transfer.

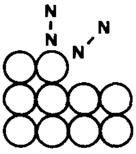


FIG. 8

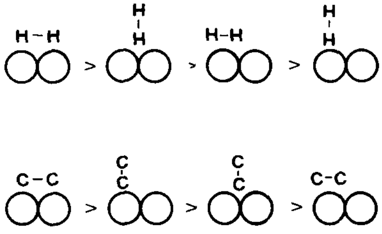


FIG. 9

Fig. 8. The most favorable adsorption sites for N_2 on stepped W clusters (after Anderson and Hoffmann, 1974b).

Fig. 9. Relative stabilities in the adsorption of C_2 and H_2 on two Fe atoms (Anderson, 1977).

The chemisorption of CO on W(100) has been studied by EHMO (Lee and Rabalais, 1978). Adsorption binding energies, relative stabilities of various adsorption sites, as well as the possibility of dissociative chemisorption with related activation energies have been obtained. On the basis of these results, a detailed interpretation of the experimental data for the various types of adsorbed CO is offered. Finally, an example of investigation of chemisorption on semiconductor surfaces may be mentioned (Nishida, 1978); the system that was considered is hydrogen on (111) clusters of Si and Ge.

3. *The Performance of Anderson's Modified EHT*

The inclusion of an electrostatic repulsion term (see Section II,B) allows, in this elaboration of the original EH formalism (Anderson, 1975a), a more confident treatment of binding energies and geometries. The method was first applied to small clusters of C, Si, Ge and Sn atoms (Anderson, 1975b) and then to clusters of Ti, Fe, Cr, Ni (Anderson, 1976a). The adsorption of ethylene and acetylene on Ni(111) was then considered (Anderson, 1976b), and the conclusion was reached that two Fe atoms gave essentially the same results as a 13-atom Fe cluster. This points strongly to a high degree of localization of the chemisorption interaction. The geometry of the adsorbate was optimized, as was the distance from the metal surface. The results are summarized in Table IV; Table V has the results for the adsorption and decomposition of small organic fragments, H_2, ethylene, and acetylene on Fe (Anderson, 1977). Again, two Fe atoms suffice to represent the essential properties of the surface interactions. Figure 9 shows the relative stabilities of various adsorbate orientations. Along the same lines, the interaction of acetylene with a Ni(111) surface was studied in detail (Anderson, 1978b) and the

TABLE IV

EH CALCULATIONS FOR THE ADSORPTION OF ETHYLENE AND ACETYLENE ON Ni(111)[a]

Property	Ethylene	Acetylene
Binding energy		
2-atom cluster	136	108
13-atom cluster	216	156
ΔE^{act} dehydrogenation	41	45
Experimental T of dehydrogenation	230 K	470 K
ΔE^{act} dissociation	10	10
Distance from surface	2.0	1.9
Back-bend angle[b]	45°	55°
Variation in C—C distance[c]	+0.22	+0.20

[a] Anderson (1976b). Energies are in kcal/mole, distances in Å.

[b] Deviation from 180° of the HCC angle in acetylene, or of the C–C–HH midpoint angle in ethylene.

[c] With respect to the free molecule.

results discussed in connection with structures and reactions of acetylene complexes.

4. *Periodic Surface Calculations*

A periodic surface calculation of H_2 on a Ni surface has been performed (Kobayashi *et al.*, 1978), in which the wave function for the Ni

TABLE V

ADSORPTION OF FRAGMENTS ON A TWO-ATOM Fe CLUSTER[a]

Fragment	Binding energy	ΔE^{act} dissociation	Energy gain after dissociation	Distance from surface[b]
H_2	15	23	12	2.7
C_2	115	7	28	1.65
CH	135	32	20	1.9
CH_2	105	24	22	1.85
CH_3	84	14	25	2.1
CH_4	43	21	32	3.0
C_2H_2	110	0	25	1.7
C_2H_4	109	1	7	2.0

[a] EH calculations by Anderson (1977). Energies are in kcal/mole, distances in Å.

[b] Defined as distance between bond midpoints or between C and Fe atoms.

surface was represented by Bloch functions constructed from the AOs of the surface atoms using the tight-binding approximation. The calculations on the Ni surface and H_2 molecule were carried out within the extended Hückel scheme and the interaction between adsorbate and surface was estimated by a perturbation method. The results are compared with those obtained with the cluster approach.

C. CNDO Results

1. *Comparison with EHT for H on Ni*

A direct comparison of the results of EHT and CNDO in dealing with the same problem is possible for H adsorption on Ni; the essential results of EHT were a preference for the "overhead" adsorption above a metal atom, with large negative charges on H, and a H—Ni distance of about 1.3 Å; d orbitals were found to play a minor role in bonding (Fassaert *et al.*, 1972). Blyholder (1975a) has used CNDO for the same investigation. His charges are much smaller, as a result of CNDO self-consistency and inclusion of electron interactions; the H—Ni distance is somewhat larger (1.6 Å). The main difference is, however, that the order of relative stability is reversed, i.e., it does not appear to be favorable to have the H atom interacting with an isolated Ni atom, and the most stable position for adsorption is in a surface hole. CNDO is in accordance with EHT concerning the small importance of d orbitals in bonding. Absolute values of CNDO binding energies being unreliable, as in EHT, except perhaps in some cases, it is disturbing that the two most popular semiempirical methods give conflicting predictions also in relative stabilities. It is to be noted that the CNDO method was, in this circumstance, especially parametrized to reproduce correctly some experimental geometrical and electronic quantities of Ni metal (see Table VI; Blyholder, 1974).

TABLE VI

CALCULATED AND EXPERIMENTAL PROPERTIES OF NICKEL[a]

Property	Ni_6 cluster CNDO[b]	Exptl. bulk Ni or band calc.
Equilibrium Ni—Ni distance	2.5	2.5
Normalized binding energy	3.7	4.4
d electrons per atom	9.46	9.4
Fermi level (HO MO)	−7.7	−5.22; −4.75
d bandwidth	3.3	2.7; 5; 4

[a] Energies are in eV, distances in Å.

[b] As calculated by Blyholder (1974).

2. *CO on Ni*

Blyholder (1975b) has used CNDO also in the case of CO over the (100) and (111) faces of Ni. Various cluster sizes have been used to ensure a sufficient number of underlying Ni atoms to model reasonably the actual environment of the CO molecule. A multicenter bond, as formed essentially by s and p orbitals, over a fourfold hole on the (100) face was judged to be the most favorable adsorption mode, on the basis of binding energy, bond order, and adsorbate–substrate distance (for details, see Table VII). Correlation with IR spectral data for the chemisorbed molecule was successful, in that the band at lower energy that develops on the surface for the C—O stretching correlates with the smaller C—O bond order for the favorable adsorption mode on the fourfold hole (see Table VII).

3. *Studies on Light Metals*

Figure 10 shows some clusters that have been used to model the various faces of Li (Companion, 1976a) and Be (Companion, 1976b). The

TABLE VII

CALCULATED AND EXPERIMENTAL PROPERTIES OF CO–Ni SYSTEMS[a]

Property	Overhead position	Bridging position	Fourfold hole
Calculated C-surface distance (Å)	1.8	1.3–1.8	0.6–1.4
Calculated binding energy (a.u.)	0.13–0.31	0.22–0.39	0.41–0.70

System	Binding energy per CO molecule	
	Calculated	Experimental
$Ni(CO)_4$ (a.u.)	0.24	0.05
Chemisorbed CO (eV)	3.5–8.4	1.4–2

System	Bond orders		Exptl. C—O stretching frequency (cm^{-1})[b]
	Ni—C	C—O	
$Ni(CO)_4$	0.76	1.14	2058
CO overhead	0.78	1.12	2075
CO multicenter	2.39	0.94	1935

[a] All calculated values are from the CNDO study of Blyholder (1975b).

[b] Assignments for the two CO surface positions were made on the basis of the different bond orders.

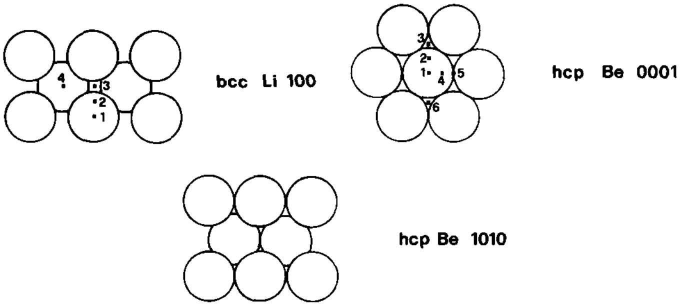


Fig. 10. Clusters of light metal atoms, as used by Companion (1976a,b) in chemisorption studies. In hcp Be, point 3 has an underlying metal atom, point 6 does not (structural "hole").

method of calculation was the diatomic energy-rescaled CNDO described in Section II,C. The scaling factors that were used are shown in Table VIII. The adsorbates were H or H_2.

The results are interesting in many respects. For H on Li, over point 1 (the numbering refers to Fig. 10) a potential well was found 1.65 Å above the surface; at point 3, on the contrary, the H atom is free to penetrate into the surface, where it experiences a periodic wavelike potential, with barriers about 2 kcal/mole high. Over point 4, another potential well is found 0.10 Å above the surface, but a barrier of only 3 kcal/mole separates this well from point 3. This treatment, though second best to having the full potential energy surface for the approach of H, is sufficient to predict that most of the incoming H atoms will eventually diffuse into the Li bulk. Also the H_2 molecule, when reaching point 3, dissociates with no activation energy to give H adatoms in 4, which are then likely to have the same fate.

TABLE VIII

CNDO DIATOMIC ENERGY RESCALING FACTORS[a]

Diatomic	Binding energy (eV)		Rescaling factor
	CNDO	Exptl.	
H—H	5.37	4.75	1.13
Li—Li	14.71	1.05	14.01
Li—H	5.90	2.52	2.34
Be—H	7.07	2.42	3.26

[a] Companion (1976a, b).

The discussion of adsorption on Be follows the same lines. The hcp structure of Be allows, however, for the presence of structural "holes" (see Fig. 10) into which it is predicted the H atoms will channel with a small energy barrier. All sites on the surface are calculated to be separated by small barriers from or completely unstable with respect to points 3 and 6. In 3 regular adsorption takes place, over the holes in 6 the previously mentioned channeling occurs. For Be, the comparison with *ab initio* calculations (Bauschlicher *et al.*, 1975) was favorable, especially if one takes into account the drastic difference in computation times, and the not so drastic difference in quantity and quality of results.

4. *Adsorption on Graphite*

Studies of chemisorption on graphite need not deal with the problem of special parametrizations for the metal atoms, which were not considered in the original formulation of such semiempirical methods as EHT or CNDO. Bennett *et al.* (1971b) have used CNDO to study the adsorption of H and first-row atoms on graphite, including an interesting boundary connection in which the C atoms layer 16C, similar to that shown in Fig. 5, is treated as a supercell, repeated by translation, and the weak interactions within this supercell are replaced by the stronger interactions between atoms within the cell and atoms repeated by translation. In this "periodically connected" graphite layer, the maximum charge on a C atom was 0.01, while without connection it was 0.11, and in EHT it was 0.35. The difference between a method that explicitly includes electronic interactions and one that does not was already evident upon comparison of the last two numbers. The inclusion of periodic connectivity improves the situation to an almost perfectly smooth and neutral layer of C atoms, without increasing the dimensions of the matrices to be diagonalized. The cohesive energy is always a maximum for a C—C spacing of 1.4 Å, although absolute values are still too large by a factor of about 5.

Table IX shows some details. Sites are labeled as in Fig. 5b. It was predicted that H would chemisorb over A by EHT and CNDO periodic connections; also for oxygen the order of stability was $B > A > C$ with connections, and $C > B > A$ without. This result may have a correlation with different adsorption modes and different catalytic activity as a function of substrate particle size. On the other hand, the consistency of the periodic connection conditions is further demonstrated by the fact that the curve of binding energy versus distance from the surface, for oxygen chemisorption on the center of a ring, was insensitive to the particular ring chosen in the model cluster.

The net charges after adsorption offer a clue to the prediction of work function changes, and also show (see the positive charge on N) that elec-

TABLE IX

CALCULATED PROPERTIES OF ADSORPTION OF VARIOUS ATOMIC SPECIES ON GRAPHITE[a]

Atom	Preferred adsorption site[b]	Net charge on adatom	Bond overlap population[c]	Number of bonds[d]
H	B	0.142	1.44	1.0
C	C	−0.02	4.84	3.3
N	C	+0.22	3.84	2.4
O	B	−0.25	1.77	1.5
F	B	−0.09	1.19	0.85

[a] CNDO study by Bennett *et al.* (1971b).

[b] See Fig. 5b.

[c] Total overlap populations summed over all the interactions of the adatom with surface atoms.

[d] Adatom-surface bonds, as obtained from overlap populations, by calibration of a scaling factor using H_2, CO, and N_2 bond orders.

tronegativity differences between adsorbate and substrate may not be always a good guide. Finally, the number of bonds formed by the adatom with the surface was found to be smaller than the conventional valency of each adatom; this confirms the high reactivity of all chemisorbed species.

Another expedient that has been used to smooth the charge distribution on graphite model clusters is to saturate the dangling valencies of the edge atoms with H atoms (Hayns, 1975). The maximum charge on a C atom was 0.04, so that this stratagem, although questionable in many respects, reaches the same goal as the periodic connection conditions. In the same work, the oxygen chemisorption was studied, with results similar to those of Bennett *et al.* (1971b); further, it was demonstrated that water is highly competitive with oxygen in chemisorption and diffusion over a graphite substrate.

A fruitful comparison can be established between the results of Messmer and co-workers (Bennett *et al.*, 1971b) for H on graphite and the results obtained using the LCAO SCF crystalline orbital method in the CNDO approximation (Dovesi *et al.*, 1976a). The method consists of defining crystalline orbitals for a periodic infinite system as linear combinations of Bloch functions, using symmetry to confine the Roothaan equations within the irreducible part of the first Brillouin zone, and then applying the CNDO approximations. This allows calculations on infinite systems consisting of a graphite crystal and a monodimensional array of adsorbate atoms. Ratios of C to H of 1 : 1 (Dovesi *et al.*, 1976c) and 1 : 2 (Dovesi *et* *al.*, 1976b) have been considered. It is interesting to note that the relative stabilities of sites A, B, and C (Fig. 5b) are the same as obtained from periodically connected cluster treatment, and hence differ from those obtained by EHT or unconnected cluster CNDO. No improvement was found in the reliability of absolute energy values, this being intimately connected with the CNDO approximations. In addition, band structure and density of states spectra for graphite and graphite + H were obtained.

5. *Dehydrogenation of Cyclohexane on Pt*

Ruiz-Vizcaya *et al.* (1978) present experimental evidence that the adsorption leading to activation for this reaction occurs over Pt sites only. A mechanism involving π–σ shifts is proposed for the dehydrogenation; CNDO calculations using a cyclohexane molecule and a six-atom Pt cluster yield an activation energy for the shift of 14 kcal/mole, in good agreement with experiment. The calculated buildup of negative charge on the σ-bound H atoms confirms the hypothesis that H_2 elimination proceeds via hydride ions.

D. Results from More Rigorous Quantum Mechanical Methods

1. *Me–H Systems*

As the number of approximations decreases, and the computational requirements increase, calculations on model systems tend to reduce the number of atoms to a bare minimum. Metal hydride calculations performed to date include MnH (Bagus and Schaefer, 1973), TiH (Scott and Richards, 1974a), ScH (Scott and Richards, 1974b), FeH (Scott and Richards, 1975), and VH (Henderson *et al.*, 1975). The idea that simple systems, including only one or two metal atoms, could be useful in understanding adsorption mechanisms was then considered. Deuss and van der Avoird (1973) used a simplified perturbation approach to study the system $H_2 + Ni_2 \rightarrow 2NiH$, keeping the H—H and Ni—Ni bonds parallel. The Ni atoms were described alternatively by one 4s or one 3d orbital each. The energy was calculated as a function of the H—H bond distance and the distance h between the two bond midpoints. By analysis of the resulting two-dimensional surface, the conclusion was reached that the d orbitals are responsible for a bonding interaction, while the s orbitals provide a repulsion giving rise to an activation energy barrier. Melius *et al.* (1976) carried out *ab initio* calculations for the same system, solving simultaneously for the interaction of all d and s metal electrons with the 1s electrons of the H atoms. The argon core of Ni was replaced by an effective potential. Whereas the distance h was kept constant at the

Ni—H bond distance and the Ni—Ni distance at the bulk nearest-neighbor separation, the H—H distance was varied. Three values were considered: 0.74 Å (the experimental free H_2 distance), 1.06 Å (a transition state), and 2.49 Å (the Ni—Ni separation). A high-energy barrier was calculated to reach the transition state, at which point a triplet state has the lowest energy. However, the main bonding effects were to be ascribed to s electrons, rather than d electrons, much the same as observed in the diatomics Ni—H and Ni_2. The 3d electrons were found to be essentially localized on the metal atoms, mixing in only very slightly.

Hartree–Fock calculations on NiH, CuH, MnH, and ScH were done by Blint *et al.* (1975, 1976). By these model calculations, a favorable bonding situation was found for the metal atom forming two sp hybrids, one of which can be used to form a bond to a supporting substrate, while the other should be available to bind the adsorbate. The Boltzmann occupation factor for the $4d^{n-1}5s^2$ state in second transition series atoms (rescaled by an empirical factor) was found to correlate with experimental trends in activity for hydrogenolysis of ethane.

Model studies using hydrides or diatomics, while very important in elucidating the electronic properties of these molecules, are probably to be used with great care in discussing surface bonding effects, although the chemisorption phenomenon may be a very localized one.

2. *CO over Ni*

A study of this very important system was undertaken by Hermann and Bagus (1977), in view of the uncertainties in assigning properly the observed peaks in UPS spectra for adsorbed CO. The orbitals that might be involved in bonding are metal d or sp manifolds, and the CO molecular 5σ or 1π levels. The total observed valence orbital shifts are thought to include two main contributions, namely, a bonding shift, due to interactions among orbitals involved in bond formation, and a relaxation contribution. These were obtained with very large basis set HF–SCF calculations of frozen orbitals or of the fully relaxed CO states. The model system was Ni—C—O, or Ni—Ni—C—O, in linear arrangements.

Results concerning binding energies and geometries were thought to be unsuitable for comparison with experimental adsorption characteristics, due to limitations of the model. In what concerns the bonding mode, CO orbitals are seen to remain unaltered except for the 5σ orbital, which forms a weak bond by mixing some $3d\sigma$ character of the nearest Ni atom (the presence of a second Ni atom was found to be immaterial to this picture). No Ni 4s levels are involved. The experimental level shifts were about 4.5 eV larger than the calculated ones, the relaxation shift being probably underestimated because of the small number of Ni atoms. However, the only large calculated bonding level shift was for the 5σ level, as expected.

3. *Cluster Calculations*

For clusters, the calculations that have been attempted so far are always restricted to light metals of the first row of the periodic table. Kunz *et al.* (1973) have used the spin-unrestricted HF method to study a cluster of Li atoms including all lattice points within 5 bohrs of an incoming H atom. A preliminary test study was carried out on Li_2 and LiH, with satisfactory results. The order of stability for the positions of the adsorbed atom was $6 > 5 > 4 > 3 > 2 > 1$ (see Fig. 11a); in 6, the minimum is at about 1 bohr below the surface. However, a perhaps more extensive study on the same system by Stoll and Preuss (1977) by the *ab initio* HF method showed the H atom to be adsorbed in position 6 at about 0.70 bohrs above the surface, and stressed at the same time the sensitivity of bond energy and geometry calculations to the choice of basis set and even to the geometry of the cluster (these authors used a variety of clusters from Li_4 to Li_{19}).

The most comprehensive effort towards calculation of chemisorption properties by *ab initio* methods is the work that Schaefer and co-workers have carried out using beryllium as a substrate. The reasons for this choice are (Bauschlicher *et al.*, 1975): Be has a closed shell ground electronic state, while Li, e.g., has an unpaired spin that could lead to difficul-

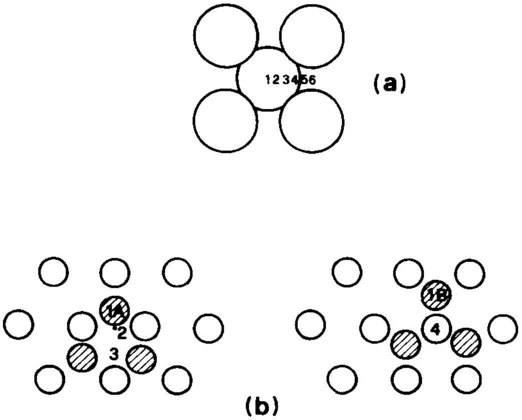


Fig. 11. (a) Position numbering for the study of H adsorption over a Li cluster (Kunz *et al.*, 1973). (b) Large Be clusters and sites for H adsorption (Bauschlicher *et al.*, 1976). 1A and 1B, "eclipsed" site; 2, bond midpoint; 3, "open" site; 4, "overhead" site. Circles representing Be atoms (the shaded ones are on the lower layer) are not drawn to scale.

TABLE X

Ab Initio RESULTS FOR SMALL (UP TO 10 ATOMS) Be CLUSTERS + H[a]

Site[b]	No. of Be atoms in cluster	Equilibrium H-surface distance[c] (Å)	Binding energy[c] (kcal/mole)
Open	3–9	1.02–1.09 (1.11–1.25)	4.2–38.4 (19.1–55.3)
Eclipsed	4–7	1.11–1.24	28.7–40.8
Bond midpoint	4–10	1.14–1.25 (1.17–1.26)	31.5–65.1 (45.3–70.1)
Overhead	7 or 10	1.42 or 1.43	16.9 or 71.3
Be–H molecule	—	1.42	46.4

[a] Bauschlicher *et al.* (1975, 1976).

[b] For site definitions, see Fig. 11.

[c] Minimum basis set results; in parentheses, results from spatially unrestricted calculations.

ties with the *ab initio* cluster approach; the HF method is exceptionally accurate for BeH dissociation energy; minimum basis set calculations yield acceptable results, as compared with larger basis sets; and finally, although experimental work on that metal has not yet been carried out, such a study is said to be feasible, to provide a check of the calculations. Figure 11b shows some clusters that have been used (the larger ones, as described in further work, see Bauschlicher *et al.*, 1976).

The results for small Be clusters are reported in Table X. Although no clear preference for one adsorption site is yet evident from these results, a number of important points can be discussed; e.g., the dramatic difference in binding energy for overhead adsorption on going from a 7- to a 10-atom cluster is partly explained if one considers the difference in cohesive energy between the two clusters (3.6 and 9 kcal/mole · atom, respectively). Thus, the Be_7 cluster sees the approaching H atom as a way of alleviating its relative internal "discomfort."

Larger clusters provided more conclusive evidence. Table XI shows the relative stabilities for different adsorption modes and confirms the intuition that shorter bonds are stronger. These results, and those of Table X, should be evaluated keeping in mind that the criteria for finding a good model cluster are that the cohesive energy be high and the ionization potential low; as seen in Table XII, Be_{13} clusters are convincing in this respect, especially because it has been verified that their properties reach convergence with respect to increasing number of metal atoms.

TABLE XI

Ab Initio RESULTS FOR LARGE Be CLUSTERS + H[a]

Site[b]	H-surface distance (Å)	Binding energy (kcal/mole)
Eclipsed A	0.90	57.2
Eclipsed B	0.94	58.3
Bond midpoint	1.05	44.8
Open	0.99	39.0
Overhead	1.41	31.1

[a] Bauschlicher *et al.* (1976).

[b] For site definitions, see Fig. 11.

One very important point concerning *ab initio* calculations on these systems is that, although the H 1s orbital is so much lower than Be 2s or 2p, a strictly covalent bond is predicted, the H electronic population always being around 1. This fact issues a further warning in the consideration of the large charge transfer usually found in semiempirical calculations.

The hybridization of Be atoms (as obtained from ratios of s to p populations) ranges from $sp^{2.6}$ to sp; an increase in 2p populations was found on Be atoms as H approaches, partly compensated by a decrease in 2s populations. The largest population on a Be atom was found for the directly approached Be atom in overhead adsorption (4.34 electrons).

Other *ab initio* studies of surface–adsorbate systems include calculations on the physical adsorption of He on model LiH surfaces, performed by the floating spherical Gaussian orbital (FSGO) method (Wood, 1978),

TABLE XII

Ab Initio CALCULATED PROPERTIES OF SOME Be CLUSTERS[a]

Cluster	Cohesive energy[b] (kcal/mole)		Ionization potential[b] (eV)	
Be_4	−9.78	+4.07	5.13	7.18
Be_7	1.49	3.64	5.21	5.58
Be_{10}	8.94	8.98	5.17	6.05
Be_{13}	14.03	14.15	4.32	4.70

[a] Bauschlicher *et al.* (1975, 1976).

[b] Minimum and maximum values for different cluster shapes.

and, as a first step toward a full *ab initio* study of clusters of transition metal atoms, the investigation of the bonding of S atoms to Ni clusters consisting of 1–4 atoms (Walch and Goddard, 1978).

4. *Valence Bond Treatment of Some Metal–Organic Fragments*

Goddard and co-workers (1977) have proposed an ingenious approach to the estimation of relative reactivities on surfaces. Using VB, these authors have evaluated the bond energies of a large number of fragments formed by the approach of an organic group to one or two metal atoms; these energies were subsequently converted into standard heats of formation for the metal–organic aggregate. In this way, activation energies for each of the steps in two different mechanisms proposed for CO methanation over Ni were evaluated, and relative velocities and rate-determining steps were calculated and identified. As an example, Table XIII reports some of the thermochemical data thus obtained. This approach is in principle very promising, though suffering from the restriction that bond energies calculated using a very small number of metal atoms in the surroundings of the adsorbate may not be realistic.

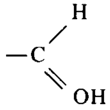
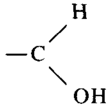
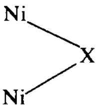
E. Results of the Xα Method

The Xα–SW method turns out to be particularly useful in the interpretation of photoemission spectra of simple molecules adsorbed on metal surfaces. A variety of models and of modifications of the standard technique have been used to this purpose. Rösch and Rhodin (1974) studied bonding of ethylene to diatomic nickel. At variance with the standard technique described before, overlapping spheres were used with the parametrization found to give excellent results for ionization potentials in ethylene and benzene (Rösch *et al.*, 1973). The C=C bond was assumed to be parallel (di-σ bonding) or perpendicular (π bonding) to the Ni—Ni bond. On the basis of the comparison of the shift of ionization potentials and change in work function obtained from experiments, the calculations favor the latter arrangement, while in the former the C=C double bond is weaker, supporting the intervention of such a complex in heterogeneous reactions.

A system that has been extensively studied is oxygen on Ni. For oxygen chemisorbed on nickel for a range of coverages UPS spectra have been reported (Eastman and Cashion, 1971). It is believed that at low oxygen coverages oxygen forms an adsorbed overlayer (Demuth *et al.*, 1973), whereas at high oxygen coverages incipient formation of surface nickel oxide is observed. The SCF–Xα–SW cluster method with spin polarization has been applied to both situations. In the case of high coverage the cluster contains an Ni^{2+} ion octahedrally coordinated by six O^{2-}

TABLE XIII

Valence Bond Calculations on Metal–Organic Aggregates[a]

Aggregate	Dissociation energies (kcal/mole)	Standard heats of formation (kcal/mole)	Ni—X distance (Å)
Ni—X			
X = —H	64	−12	1.45
—CO	27	−53	1.90
$=CH_2$	65	27	1.78
$—CH_3$	60	−25	1.87
	57	−47	—
	56	−4	—
=O	91	−31	1.60
—OH	50	−41	1.72
=S	76	−9	1.91
			
X = CO	34	−60	1.94
CH_2	122	−30	1.91
C:	91	80	—
O	101	−41	1.79
S	124	−57	2.04

[a] Goddard and co-workers (1977).

ions (Messmer *et al.*, 1974). The stabilizing effect of the surrounding Ni^{2+} ions has been modeled by surrounding the NiO_6^{10-} cluster with a spherical shell of charge +10. When spin–orbital energies are corrected for relaxation effects and the work function value of 5.5 eV is added to the experimental data, theoretical results are in nice agreement with UPS and XPS (Wertheim and Hufner, 1972) spectra. The cluster model for calculating energy levels of oxygen chemisorbed on Ni is an ONi_5 cluster, shown in Fig. 12 (Batra and Robaux, 1975). The oxygen is in a fourfold coordination site and the Ni–Ni distance is assumed the same as in the bulk. Overlapping spheres were used to optimize the virial coefficient R_{Ni} = 2.85 a.u.,

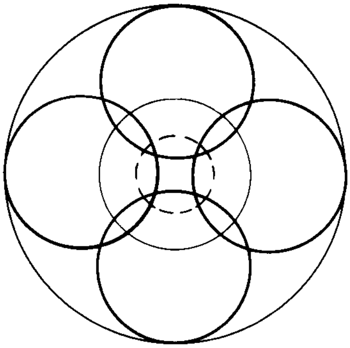


FIG. 12

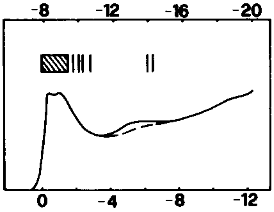


FIG. 13

Fig. 12. Model cluster for the Xα–SW study of O chemisorption on Ni (Batra and Robaux, 1975). First-layer Ni atoms (heavy circles); lower-level Ni atom (light circle); O atom, 1.67 a.u. above the first Ni layer (dashed circle).

Fig. 13. Experimental photoemission spectrum (full line, lower energy scale) and calculated electronic structure for the ONi_5 cluster (upper energy scale). The peak over the dashed line arises from chemisorbed O. All energies are in eV. Adapted from Batra and Robaux (1975).

$R_O = 1.77$ a.u. Ionization potentials were calculated by the transition state method, as described by Weinbergen and Schwartz (1975). Since the one-electron energies in the Xα method are partial derivatives of the total statistical energy with respect to orbital occupation numbers:

$$\varepsilon_i = \partial E_t / \partial n_i, \tag{81}$$

Koopman's theorem cannot be used to obtain ionization energies. These are calculated as the energy ε_i in the transition state, a state halfway between the final and initial states of the ionization process, i.e., a state in which the orbital i occupation number is taken to be $n_i = \frac{1}{2}$. The occupied levels are shown in Fig. 13, together with the experimental UPS spectrum. The calculated spectrum is shifted to produce the best overall agreement.

In the surface molecule model only local bonding is taken care of, and inclusion of delocalization effects provided by the metal would require procedures such as the embedded cluster treatment (Grimley, 1975). These procedures are very hard to apply in actual calculations and the errors embodied in the localization approximation can be less important if the trends in a series of systems are investigated, rather than absolute values in one system. To this aim SCF–Xα–SW calculations using overlapping spheres have been performed for oxygen chemisorption on Cu, Ni, and Ag clusters (Rösch and Menzel, 1976). The corresponding metal clusters were studied first, namely, M_6 and M_4 clusters shown in Figs. 14 and 15. Bond distances were chosen from bulk metal structures; 2.49,

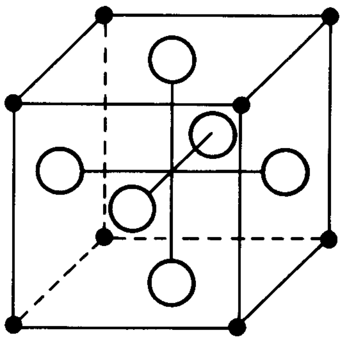


FIG. 14

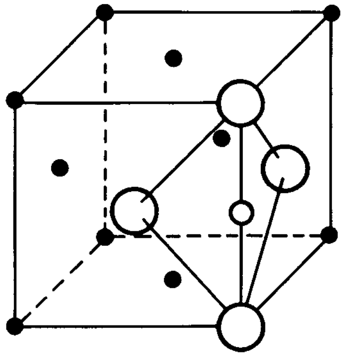


FIG. 15

Fig. 14. Octahedral model cluster in an fcc lattice.

Fig. 15. Four-atom cluster modeling a (110) substrate surface. The small circle is the position of the O atom as considered in the work by Rösch and Menzel (1976).

2.54, and 2.89 Å for Ni, Cu, and Ag, respectively, and touching spheres were used. An orbital population density analysis was performed for the octahedral cluster for both sp and d orbitals. The results are shown in Table XIV together with the results of bulk energy band calculations and experimental data. It is seen that clusters of six atoms reflect the essential

TABLE XIV

ELECTRONIC STRUCTURES OF OCTAHEDRAL CLUSTERS AND BULK METAL ENERGY BANDS[a]

	d bandwidth	sp bandwidth	Bottom sp band to top of d band	Top of d band to Fermi level
Ni				
Cluster	2.0	6.9	3.7	−0.1
Bulk	4.7	11.4	6.9	−0.02
Experiment	3.0 or 3.3	—	—	—
Cu				
Cluster	1.3	6.9	2.6	1.1
Bulk	3.3	10.8	5.7	1.8
Experiment	3.0–3.2	—	—	1.9 or 2.0
Ag				
Cluster	1.6	5.4	−0.9	3.6
Bulk	4.1	7.9	1.9	4.1
Experiment	3.2 or 3.5	—	—	3.9 or 4.0

[a] From Rösch and Menzel (1976). All energies are in eV.

electronic properties of the corresponding metals and can be used in chemisorption studies.

Chemisorption of oxygen on Ag(110) was studied by means of an Ag_4O cluster (see Fig. 15). The following model clusters were used:

(1) Ag_4 cluster with an empty sphere in plane of the O atoms;
(2) Ag_4O cluster with touching spheres;
(3) Ag_4O cluster with overlapping spheres: $R_O = 1.37$ a.u.;
(4) Ag_4O cluster as in (3) but with the O atom lifted by 0.5 a.u. along the twofold axis.

The energy levels resulting from the calculations are shown in Fig. 16. Ionization potentials were obtained assuming uniform relaxation. The spectrum obtained with model 4) shows the best agreement with experimental observation from PES spectra. To compare chemisorption on the different metals M_5O clusters were used. Comparison of the resulting electronic structures shows a decreasing strength of the oxygen–metal bond along the series. Differences in UPS spectra for the system O—M are well reproduced in calculated spectra.

A test of the reliability of the results of the SCF–Xα–SW method applied to metal clusters can be obtained by comparing density of states as obtained in bulk band structure calculations with those for clusters using the SCF–Xα–SW method. Such a comparison for a 13-atom Ni cluster is shown in Fig. 17 (Messmer *et al.*, 1975). It is clear that theoretical results have many features in common with the bulk density of states. The same kind of calculations have been carried out for Ni_8, Cu_8, Cu_{13}, Pd_{13}, and Pt_{13} clusters (with cubic geometry for M_8 clusters and the geometry shown in Fig. 18 for M_{13} clusters; Messmer *et al.*, 1976). In all cases the results show marked similarity of the electronic structures with those of the corresponding bulk metals. On the other hand, a large surface effect is in operation in the case of clusters, with the result that ionization potentials for clusters are higher than the work functions of the corresponding crystalline metals. The experience accumulated with metal clus-

Fig. 16. SCF–Xα–SW orbital energies (rydbergs) for Ag_4 and various Ag_4O clusters (1, 2, 3, 4—see text). Shaded areas represent bands of levels very close together. Adapted from Rösch and Menzel (1976).

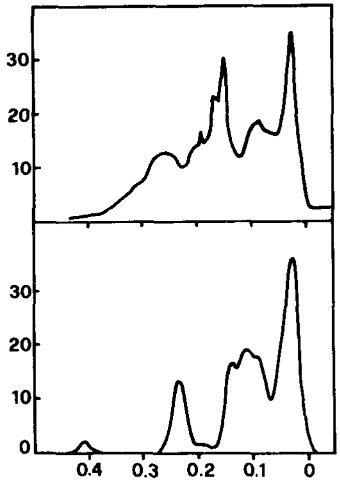


Fig. 17. Comparison between a band structure calculation of the density of states of bulk Ni (upper curve) and the results of an Xα–SW calculation using a 13-atom Ni cluster (lower curve). Energy abscissa in rydbergs, ordinate in arbitrary units. Adapted from Messmer *et al.* (1975).

ter calculations allowed interactions with chemisorbed atoms to be studied in detail. As a first example we mention oxygen chemisorption on Al(100) (Messmer and Salahub, 1977). Clusters with 5, 9, and 25 Al atoms were considered, with the oxygen atom 0.0, 2.0, and 4.0 a.u. above the metal surface, respectively. The density of states calculated for Al_{25} is compared in Fig. 19 with the photoemission spectra of clean Al. In Fig. 20 the same comparison is offered for the Al_{25}–O calculation with O at $z = 0$ a.u. and the photoemission spectrum for Al exposed to 1L (langmuir) of

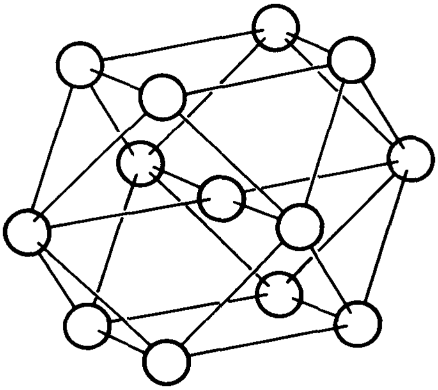


Fig. 18. Cubo-octahedral cluster used by Messmer and co-workers (1976) in an Xα–SW study of transition and noble metals.

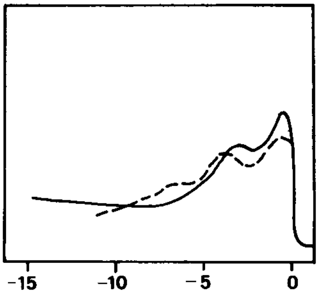


Fig. 19. Density of states for an Al_{25} cluster (dashed line, from Xα–SW calculations) and the experimental photoemission spectrum of Al (solid line). Energies are in eV. Adapted from Messmer and Salahub (1977).

oxygen. The comparisons are indeed very encouraging and support oxygen incorporation into the aluminum substrate.

The interaction of atomic hydrogen with Ni, Pd, and Pt clusters has also been investigated by the SCF–Xα–SW method (Messmer *et al.*, 1977), where relativistic effects were included, using the formalism described by Yang and Rabii (1975). Tetrahedral metal clusters were considered with metal–metal distances equal to those of the bulk metals, with or without addition of one H atom in an interstitial site. The results are in good agreement with photoemission studies of the bulk interstitial compound PdH and with photoemission spectra for hydrogen chemisorption on (111) surfaces of Ni, Pd, and Pt. For the metal clusters it is found that the Fermi level coincides with the top of the d band, and the d bandwidth, although less than in the corresponding bulk metal, increases from Ni_4 to Pd_4 to Pt_4, while the energy levels shift downward, as in crystalline

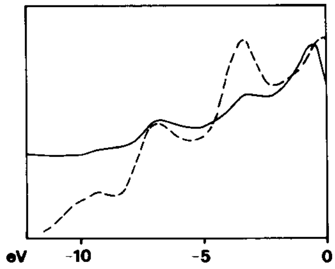


Fig. 20

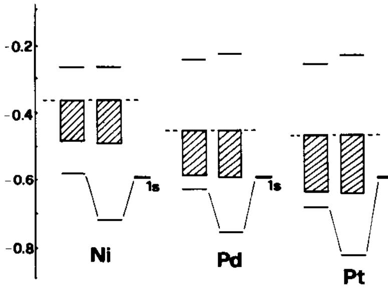


Fig. 21

Fig. 20. Experimental photoemission intensity for Al + 1 langmuir oxygen (solid line). Density of states for $Al_{25}O$ from Xα–SW calculations (dashed line). Adapted from Messmer and Salahub (1977).

Fig. 21. Nonrelativistic Xα orbital energies (rydbergs) of tetrahedral metal clusters with and without H. For each metal, at left is the structure of Me_4, at the right the structure of Me_4H. The heavy lines represent the H 1s level. Adapted from Messmer *et al.* (1977).

metals. The electronic structures of Pt_4 and Pd_4 are similar, whereas that of Ni_4 shows marked differences. This parallels differences in photoemission spectra for hydrogen chemisorbed on these metals, and in hydrogen solubility and catalytic reactivity. After addition of the H atom the levels shift as shown in Fig. 21. The 1s H SCF–Xα energy is also shown and it is clear that the position of this level with respect to the Ni orbital manifold justifies the differences in behavior previously mentioned. The influence of relativistic effects is negligible or minor in Ni and Pd, whereas in the case of Pt it produces an increase of electronegativity of the Pt aggregate, as expected from the decrease in the work function of Pt upon hydrogen adsorption. The theoretical results are reinforced by the results of photoemission spectra of hydrogen chemisorbed on the (111) faces of the three metals.

Spin-unrestricted calculations for a $CO(Ni)_5$ cluster with C_{4v} symmetry were carried on using overlapping atomic spheres (Batra and Bagus, 1975). Comparison of theoretical results with the photoemission spectra of CO chemisorbed on Ni(100) is shown in Fig. 22. A tentative interpretation of the observed spectrum was proposed.

Similar calculations have been performed for the $CO(Cu)_5$ and the COCu (linear) clusters (Yu, 1977). The Cu–C distance was varied and the

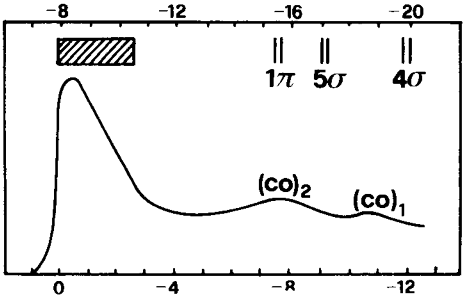


FIG. 22

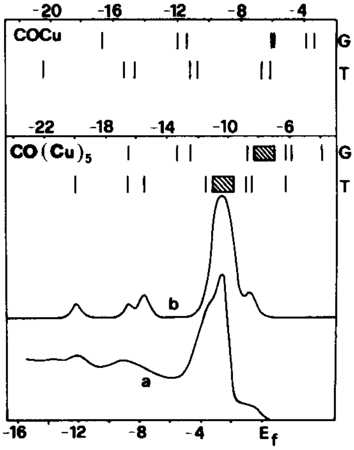


FIG. 23

Fig. 22. Experimental photoemission spectrum (lower scale) and Xα–SW energy levels (upper scale) for $CO(Ni)_5$. Energies are in eV. Adapted from Batra and Bagus (1975).

Fig. 23. Curve (a) experimental UPS spectrum for CO adsorbed on Cu(100). Curve (b) calculated valence density of states of the $CO(Cu)_5$ cluster. Ground-state (G), transition-state (T) orbital energies are shown for both COCu and $CO(Cu)_5$. Calculated quantities refer to upper scales, the experimental curve to the lower scale (eV in both cases). Adapted from Yu (1977).

total energy minimized with respect to this parameter. The results are reported for both clusters in Fig. 23. Ground-state energies, ionization energies as obtained according to the transition state method, and the experimental photoemission spectrum for CO adsorbed on Cu(100) are shown. Theoretical density of states was obtained from ionization energies by means of the following broadening function:

$$D(\varepsilon) = [(2\pi)^{1/2}\sigma]^{-1} \sum_i n_i \exp[-(\varepsilon - \varepsilon_i)^2/2\sigma^2], \tag{82}$$

where n_i is the occupation number of the ith molecular state with energy ε_i and σ is a constant chosen to be 3 eV. This curve is also shown in Fig. 23. The agreement between calculated ionization energies and spectral data is excellent, suggesting the following interpretation. The majority of CO molecules are adsorbed at fourfold symmetry sites and a smaller fraction are adsorbed on top of Cu atoms. The molecules in the C_{4v} positions produce two peaks in the spectrum below the copper d band, whereas molecules atop of Cu atoms form a single visible peak; the presence of three peaks found in experimental work is justified. The CO molecules in the two chemisorption sites are placed at different distances from the surface.

The Xα technique has been applied also to the study of various cluster–adsorbate complexes (Ni_5CO, Pt_5O, Pd_5CO, Pd_5O) by Chen *et al.* (1977).

The SCF–Xα–SW method, though satisfactory in many respects, has also some undesirable features. For example, as a consequence of the muffin-tin approximation with the assumption of constant potential in the intersphere region the potential in the bonding region between atoms is overestimated (Danese and Connolly, 1974; Danese, 1974). A partial remedy can be found in using overlapping spheres, with results as previously shown. More direct attacks on this problem can be found in the literature. The so-called Hartree–Fock–Slater (HFS) method has been proposed, in which the Xα approximation is included, but the muffin-tin approximation is avoided (Baerends *et al.,* 1973). The problem of finding approximate solutions to the one-electron HFS equation

$$\begin{aligned} h\psi_i &= \left\{ -\tfrac{1}{2}\,\nabla^2(1) + \sum_n (-z_n/r_{1n}) + V_{\mathrm{C}}(1) + V_{\mathrm{X}\alpha}(1) \right\} \psi_i(1) \\ &= \varepsilon_i \psi_i(1) \end{aligned} \tag{83}$$

is attacked by expanding the eigenfunctions into a finite set of basis functions and determining the expansion coefficients. The second part is implemented by means of a discrete variational method (Ellis and Painter, 1970) in which a linear variational function

$$\psi_i = \sum_{j=1}^{n} \chi_j c_{ji} \tag{84}$$

is used, and mean error functions are minimized:

$$\Delta_{ij} = \sum_k w(r_k)\psi_i^*(r_k)(h - \varepsilon)\psi_j(r_k), \tag{85}$$

where $w(r_k)$ is a weight factor and the summation is over a discrete set of sample points. We are led to the familiar secular equations, but matrix elements for energy and identity operators are summations, not integrals. There are no restrictions on the form of the basis functions χ_j, and STOs are as convenient as GTOs. For calculating $V_C\psi$ in Eq. (83) with

$$V_C(r_k) = \int [l(r_2)/|r_2 - r_k|]\, dr_2, \tag{86}$$

it was found useful to expand $\rho(r)$ in one-center functions:

$$\rho(r) = \sum_i a_i f_i, \tag{87}$$

where fit functions f_i are centered upon the nuclei of the system. A procedure leading to the best choice of fit functions and the determination of a_i coefficients has been worked out. The method has been successfully applied to a number of small molecules (Baerends and Ros, 1973; Heijser *et al.*, 1976). An alternative method for implementing the HFS scheme is to use the so-called self-consistent charge approximation (Rosén *et al.*, 1976). To a great extent this method follows the formalism outlined previously, including the discrete variational method to evaluate energy and overlap matrix elements. To obtain the molecular charge density, a Mulliken population analysis of the wavefunctions is performed, ending with a charge density which is the sum of spherically symmetric and overlapping atomic densities. The method has been applied to study chemisorption of first-row atoms on Ni(100) surfaces (Ellis *et al.*, 1976). The usual XNi_5 clusters were used, with X = H, C, N, O. The variation of energy levels and ground-state charge distribution with the height of X above the metal surface has been explored and the broadening and shifts of adsorbate levels due to interaction with the nickel conduction band has been discussed. The effect of embedding the surface molecule in the substrate has been explored through the use of an external potential. The results are good as far as the interpretation of photoemission spectra and electronic structure are concerned. The method still needs some improvements to make it adequate for the discussion of dissociation energies and equilibrium geometries for systems of such complexity. A further application has been explored for chemisorption of CO on the same Ni(100) surface (Ellis *et al.*, 1977). The results are consistent with CO approaching the Ni(100) plane at a C_{4v} position, at a height of about 3 a.u. The effect of relaxation and binding shifts on UPS and XPS spectra are carefully analyzed.

Non-muffin-tin corrections were also introduced in a study of oxygen chemisorption on Ni(001) (Li and Connolly, 1977). The method of including these corrections is by means of first-order perturbation theory (Danese and Connolly, 1974; Danese, 1974). The usual ONi_5 cluster was used and curves for the binding energy of oxygen on the Ni_5 cluster as a function of the distance from the surface were obtained both with and without muffin-tin corrections. They are shown in Fig. 24. The minimum is found for a height of 1.42 a.u., which corresponds to a Ni—O bond length of 3.62 a.u. to be compared with a LEED result of 3.73 a.u. and with a distance of 3.47 a.u. found in Ni–chelate complexes.

A transition-state method has been recently proposed (Ziegler and Rank, 1977) for calculating bonding energies and bond distances within the Hartree–Fock–Slater method. Applications to simple diatomic molecules and transition metal complexes led to satisfying results and the method looks very promising for the study of bonding interactions between metal fragments and organic molecules.

IV. Adsorbate Crystallography

It has been shown, mainly by LEED, that atoms or small molecules can adsorb on low Miller index crystal surfaces to form well-ordered overlayer structures (Buchholz and Somorjai, 1976). From the observation of LEED diffraction patterns the size and shape of the two-dimensional unit cell is easily established; the determination of the positions of the adsorbed atoms requires a detailed analysis of the intensities

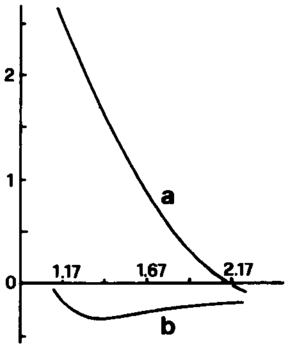


Fig. 24. Binding energy (rydbergs), calculated by the Xα method, for oxygen approaching a Ni_5 cluster. The abscissa is the oxygen cluster (distance in a.u.). Curve (a) muffin-tin approximation. Curve (b) non-muffin-tin corrections. Adapted from Li and Connolly (1977).

of the diffracted beams. This can be performed in principle by means of dynamical LEED calculations usually based on multiple scattering theory (Pendry, 1974). Several techniques, including approximations of various kinds, are available, but the actual computations are complex, so that only clean metal surfaces or simple atomic overlayers have been studied in the past; a presentation of the status of the art has recently been published (Jona, 1977).

The only example to date of a detailed analysis for molecular chemisorption is acetylene adsorbed on the (111) crystal face of platinum (Stair and Somorjai, 1976, 1977; Kesmodel *et al.*, 1976, 1977). At low exposure at 300 K acetylene forms an ordered structure with a (2 × 2) diffraction pattern. This structure is metastable, and after heating for an hour at 400 K and then cooling at 300 K it transforms into a stable structure having the same (2 × 2) pattern. Intensity versus incident electron energy curves were calculated for a number of beams under different diffraction conditions. The six models investigated (A1–C2) are shown in Fig. 25. Calculated curves were almost insensitive to the value of the C–C bond length. The best agreement between calculated and experimental *I–E* curves for the stable structure was found for the hydrocarbon molecule adsorbed at a triangular site (model C2), parallel to the metal surface, at a distance of 1.9 Å above the plane of Pt atoms. The CCH angle was estimated to be greater than 150°. The *I–E* profiles for geometries A1 and C1 are actually identical to those for A2 and C2, respectively. A quantum mechanical study has been performed by means of EH MO calculations for clusters of 9–11 Pt atoms plus one acetylene molecule (Gavezzotti and Simonetta, 1977). The same adsorption sites A1–C2 were investigated. An assumed distorted geometry for acetylene (with C—C distance of 1.3 Å and CCH angle of 140°) was found to give improved energy values with respect to the gas-phase geometry. A bonding situation was found for the

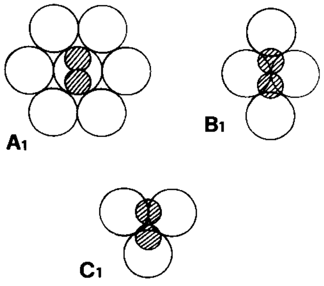


Fig. 25. A1, B1, C1 arrangements (see text) of an acetylene molecule (represented by the shaded circles) over a Pt(111) surface. A2, B2, C2 are obtained by rotating the acetylene molecule by 90° in the plane parallel to the surface plane.

B2 model with the C—C bond 2.0 Å above the Pt surface. For models A and C the energy was found to be independent of the orientation of the hydrocarbon around an axis perpendicular to the metal surface. Free rotation of adsorbed molecules was confirmed by means of molecular mechanics computations for an array of nine acetylene molecules on Pt(111). This result is in agreement with the findings of multiple scattering calculations (Kesmodel *et al.*, 1976). It may well be that the rotational freedom available to molecules adsorbed at triangular sites makes the free energy of this situation fall below that for the blocked B2 situation at room temperature and above, making EH and LEED results compatible. It is worth mentioning that UPS studies for acetylene adsorbed on Pt(111) indicate the existence of sp^2 hybridization at the carbon atoms, forming di-σ bonding (Demuth, 1977). The same system has been investigated by means of high-resolution electron energy loss spectroscopy (Ibach *et al.*, 1977a,b). It was suggested that the conversion of metastable adsorbed acetylene to the stable form is due to the formation of ethylidene (CH_3—CH), with the C—C axis perpendicular to the Pt surface. These results prompted a reinvestigation of the system by new LEED experiments and dynamical calculations (Kesmodel *et al.*, 1978). On the basis of these new results and a reinterpretation of the spectra published by Ibach *et al.* (1977b), Somorjai proposed that the stable structure is made of ethylidyne (CH_3C) groups placed on triangular adsorption sites with the C—C bond perpendicular to the metal surface. A mechanism for the formation of CH_3C through the rearrangement of acetylene to vinylidene (CH_2C) and subsequent hydrogenation was also proposed; this mechanism has been confirmed by the results of EH MO calculations (Gavezzotti and Simonetta, 1978). From these results, the catalytic effect of a Pt surface on the acetylene–vinylidene rearrangement becomes evident (see Fig. 26).

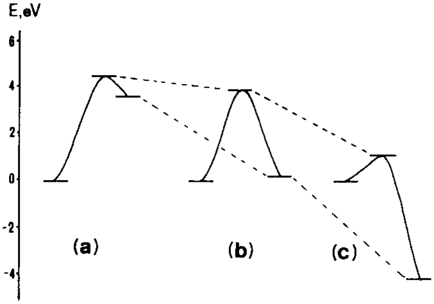


Fig. 26. Extended Hückel results (Gavezzotti and Simonetta, 1978) for the catalytic effect (lowering of the energy of the transition state) of Pt on the acetylene–vinylidene rearrangement. (a) Energy profile for the gas phase reaction; (b) the same for reaction over one Pt atom; (c) the same over a 10-atom Pt cluster.

V. Dynamical Studies

Dynamical studies for chemisorption can be carried out if a realistic potential function that is rapidly calculable is available. As usual, one is faced with the dilemma that approximate methods are of dubious reliability, whereas accurate potentials are out of the range of present computing facilities. A way out can be found if one is satisfied by model calculations, by means of which the qualitative features of the dynamical phenomenon are investigated, revealing the influence of the shape of the potential on the adsorption probability or the most effective form of energy (translational, vibrational, or rotational).

An example of an effort in this direction can be found in a series of articles by Wolken and co-workers. In one article (McCreery and Wolken, 1975a) the model potential and the method of calculation are described. The potential function is for the interaction of a diatomic molecule with a rigid solid surface; i.e., the solid surface provides the static background potential in which the gas molecule moves. To obtain this potential the London–Eyring–Polanyi–Sato (LEPS) procedure (Sato, 1955; Levine and Bernstein, 1974) is applied. This is an approximate valence bond treatment initially derived for the one-electron atoms system. The formula that gives the energy can easily be extended to treat the four-orbital four-electron systems in which six two-body interactions are present (Slater, 1931). Moreover in the case of a diatomic molecule interacting with a rigid surface only three two-body interactions are of interest, as shown in Fig. 27. Then the following energy expression is obtained:

$$E = Q_{ab} + Q_{ad} + Q_{bc} - \{\tfrac{1}{2}[(\alpha - \beta)^2 + \alpha^2 + \beta^2]\}^{1/2} \tag{88}$$

where Q_{ij} is the coulombic energy for the i–j two-electron system, $\alpha = (ab)$, $\beta = (ad) + (bc)$; (ab), (ad), and (bc) are two-center exchange integrals. Coulomb and exchange integrals are evaluated assuming that two-body interactions can be described by Morse potentials. Also a Sato parameter is included in the potential, which assumes the form:

$$V = (1 + \Delta)^{-1}\{U_1 + U_2 + U_3 - [A_1^2 + A_2 + A_3)^2 - A_1(A_2 + A_3)]^{1/2}\}, \tag{89}$$

where

$$\begin{aligned} U_i = \tfrac{1}{4}D_i[(3 + \Delta)\exp(-2\alpha_i(r_i - r_{i0})) \\ - (2 + 6\,\Delta)\exp(-\alpha_i(r_i - r_{i0}))], \end{aligned} \tag{90}$$

$$\begin{aligned} A_i = \tfrac{1}{4}D_i[(1 + 3\,\Delta)\exp(-2\alpha_i(r_i - r_{i0})) \\ - (6 + 2\,\Delta)\exp(-\alpha_i(r_i - r_{i0}))], \end{aligned} \tag{91}$$

D_i, α_i, and r_{i0} are the dissociation energy, Morse parameter, and equilibrium distance for the ith two-body interaction. The potential has been

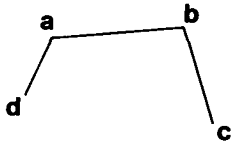


Fig. 27. Two-body interactions for a diatomic molecule plus a surface.

evaluated for the H_2 + W(100) system. The H—H Morse potential was found in the literature (Porter and Karplus, 1964), with parameters $D_{HH} = 4.7466$ eV, $r_{HH0} = 1.40083$ a.u., $\alpha_{HH} = 1.04435$ a.u. The H—W interaction was assumed to be dependent on the position of the H atom above the W(100) surface. If the surface is assumed to be the x–y plane, with the z axis pointing outward, the H—W potential is given by

$$V_{H-W} = D[\exp(-2\alpha_{HW}(z - z_0))] - 2\exp[-\alpha_{HW}(z - z_0)], \qquad (92)$$

where the parameters D, α, and z_0 are x,y-dependent. To find an explicit form of the dependence, the results of EH calculations for the same system were considered, namely the energies and equilibrium distances for the three possible adsorption sites (Anders *et al.*, 1973). The dissociation energy and equilibrium distance dependences were assumed of the form:

$$D(x, y) = D_0[1 + \delta Q(x, y)], \qquad (93)$$

$$z_0(x, y) = z_m[1 + tP(x, y)], \qquad (94)$$

where $Q(x, y)$ is chosen to have the correct symmetry for the surface, namely:

$$Q(x, y) = \cos[(2\pi x)/a] + \cos[(2\pi y)/a] - A\{\cos[(2\pi x)/a] - 1\}\{\cos[(2\pi y)/a] - 1\}, \qquad (95)$$

where $a = 5.97$ a.u. Since $Q(a/2, 0) = 0$, D_0 is the dissociation energy at the 2CN site and δ can be determined by means of the binding energy at the 1CN site ($x = 0, y = 0$). The binding energy at the 5CN site allows A to be determined. $P(x,y)$ is taken to have a similar form and the constants are determined by fitting the equilibrium heights above the surface at the three adsorption sites. The α_{HW} parameter was chosen by fitting the potential curve near the minimum to a parabola again calculated from the EH results. We obtain

$$\alpha_{HW} = [0.02844/D(x,y)]^{1/2}. \qquad (96)$$

The path of approach was chosen with the H—H bond parallel to the xy plane and to the y axis, and with one H atom moving along the z axis (this is a line through the 1CN site). The Sato parameter Δ was allowed to vary from -0.5 to 0.5, in order to obtain different surfaces with peculiar features. For the described path positive values of Δ give no barrier in the entrance channel. For values $\Delta \geq 0.2$ there is a well corresponding to one H atom in the 1CN position and the other at a distance not too far from the equilibrium distance in H_2; i.e., we see molecular adsorption. The binding energy with respect to the free molecule is 2.0 eV with $\Delta = 0.5$ or 1.0 eV with $\Delta = 0.2$. In addition, there is a barrier on the exit channel leading to a second well corresponding to the two H atoms bonded to two adjacent 1CN sites. This second well is present also for negative Δ values. A zero or negative value of Δ produces a saddle point in the entrance channel, i.e., a barrier to desorption of molecular H_2 from the surface. The height of the barrier can be used to calculate Δ by comparison with experimental work. A number of different approach paths have also been studied. As an example the potential surface for the approach with the H—H bond parallel to the x axis and the center of mass above the 1CN site is given in Fig. 28. Classical trajectories were computed for the H+H–W system with the H atom initially fixed at the 1CN position. The aiming point on the surface, the height of the starting point, and the x coordinate of the impinging atom and the collision energy were varied, while the aiming angle was kept at the constant value $\theta = 30°$ from the normal to the surface. Different outcomes from the calculated trajectories were observed, e.g., the formation and escape of a H_2 molecule, the escape of the initially fixed atom with the other atom being trapped, the impinging atom trapped in a 1CN site while the other atom is trapped and roaming the surface. Also a number of trajectories were calculated with the H atoms initially placed at two adjacent 1CN positions with total initial kinetic energy in excess of the energy needed for molecular desorption (McCreery and Wolken,

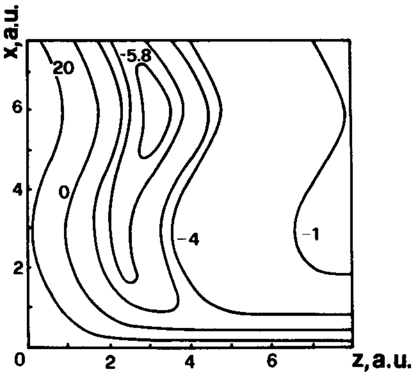


Fig. 28. LEPS potential surface (eV) for the approach of H_2 to a 1CN site over a W(100) surface. Adapted from McCreery and Wolken (1975a).

1975b, 1976a). The internal state distribution of the desorbing H_2 molecule was investigated by studying the number of molecules formed as a function of the vibration and rotation quantum numbers. The populations decrease monotonically with increasing vibrational quantum numbers, while extensive rotational excitation was observed. It was found that the percentage of molecules that desorb increases monotonically with the ratio of the initial energies of the two atoms, whereas the percentage of trajectories that result in the desorption of a H atom decreases sharply with the increase of the same ratio. No trajectories with sufficient initial kinetic energy to allow for desorption of two H atoms were computed. The angular distributions of the desorbed atoms and molecules show sharp peaks around $\theta = 0°$ and differ significantly from the cos θ form. If the closest distance of approach of the two atoms when they follow straight lines defined by the projections of their velocities in the xy plane is taken as an impact parameter, b, for the molecular desorption reactions the opacity function, i.e., the percentage of reaction as a function of b, decreases monotonically with increasing b. It was also found that the distribution of the z component of angular momentum for the desorbing molecule shows a sharp maximum for $j_z = 0$. Most of the rotation occurs in a plane perpendicular to the surface, with the two atoms tumbling over each other in the z direction. Trajectories were also calculated for the adsorption of H_2 and HD on the tungsten surface (McCreery and Wolken, 1976b,c), mainly with the purpose of establishing which form of energy is most effective in promoting adsorption. Three different values of the Sato parameter were used, to generate a potential surface with no barrier to adsorption, or a small barrier or a large one. Trajectories were calculated at constant total energy (but with different energy distributions) or with constant kinetic energy and different amounts of internal energy. It was found that translational energy is much more effective in promoting activated adsorption than internal energy. It was also found that the probability of dissociative adsorption on this surface is independent of the orientation of the projection of the beam on the surface, in agreement with experimental results. Only at high rotational levels are isotope effects observable, with HD having the larger sticking probability. The use of a rigid surface in all these calculations might be questionable, and is in part justified by the difference of mass for H and W. The H atoms move at much higher speed than the metal atoms, so that in the time needed for a trajectory, surface atoms can only adjust a negligible amount. Of course, this assumption implies that no energy between the gas and the solid is exchanged.

In more recent work (McCreery and Wolken, 1977) a LEPS surface was calculated for the interaction of three H atoms with the surface. This potential allows the study of adsorption of H_2 in the presence of an adsorbed H atom. Equipotential curves were calculated for the approach of a H_2 molecule with the H—H bond parallel to the plane of the surface. The center of mass of H_2 is approaching the adsorbed atom perpendicularly, situated on top of a 1CN site. The H—H bond can stretch in the 2CN—2CN direction. The height of the adsorbed H atom is optimized for each H_2 configuration. During the approach of the molecule the adsorbed atom is first pushed toward the surface, then it goes back to the equlibrium position. In the meantime the H—H bond stretches and the dissociation of the molecule occurs as a result of the three atoms adsorbed on the 2CN sites. Trajectory calculations were performed for different aiming areas, both for the bare and covered surfaces. A completely covered, a half-covered, and a 25% covered surface were simulated. In a completely covered surface the probability of dissociative adsorption is 15%, to be compared with a 72% probability for the bare surface. The sticking probability increases as coverage decreases. Also the dependence of the percentage of trajectories leading to dissociative adsorption on the impact parameter is very different for bare and covered surfaces: for the latter, the adsorption probability increases with increasing b; for the former, a sticking probability greater than 60% is found, regardless of b.

Exploratory classical trajectory calculations have been reported for activated dissociation of H_2 and D_2 on Cu(100) and Cu(110) surfaces (Gelb and Cardillo, 1976, 1977). Again a LEPS potential surface was used, and the interactions of two H atoms with two metal atoms were included. Morse and anti-Morse functions for H–H and H—Cu surface singlet and triplet energies were used to derive Coulomb and exchange integrals. The parameters in the hydrogen–metal functions were dependent on x, y coordinates (the xy plane is the plane of the surface) with the correct periodicity and spacing for the (100) and (110) crystal faces. Parameter calibration was adjusted in such a way to produce an energy barrier along the minimum path to dissociation of 9.7 kcal/mole for the (100) surface and 7 kcal/mole for the (110) surface. Along the minimum energy path the H_2 molecule approaches the saddle point between two Cu atoms perpendicularly, with the internuclear axis parallel to the y axis (see Fig. 29). Classical trajectories were calculated by the usual techniques (Bunker, 1971). The dissociation probabilities for the ground and first excited states of H_2 and D_2 were computed as a function of the angle of incidence, at constant kinetic energy. When a beam with vibrational and rotational distributions pertaining to reported experimental conditions is considered, the agreement between calculated and experimental data is reasonable. However, the isotope effect for the (110) surface is in the wrong direction. A lower minimum energy barrier and higher dissociation probabilities were found on the (110) surface, in accord with experiment.

The method has been modified in order to produce LEPS potentials for

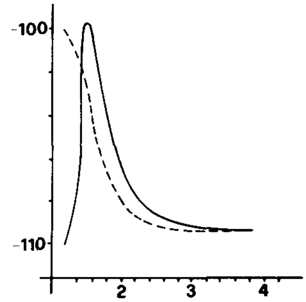


Fig. 29. Potential energy (kcal/mole) versus distance (Å) for the approach of H_2 to a Cu(100) surface over a bond midpoint (solid line); the same for impact at a fourfold site (dashed line). Adapted from Gelb and Cardillo (1977).

stepped surfaces. It has been applied (Gregory and Silbey, 1977; Gregory *et al.*, 1978) to the study of dissociative adsorption of hydrogen on copper; the (100), (110), (111) and the stepped (311) faces were considered. The latter consists of (100) terraces, two rows wide, separated by (111) steps, one atomic layer high. The calculated maximum bond energies are in the order (311) > (110) > (100) > (111). The maximum bond energy sites on these faces are those with the maximum number of nearest neighbors. The activation energy for adsorption was also estimated and the correct order for the activation energies on the faces with low indices was found.

A phenomenological model has been proposed for the collision of particles with a surface represented by a two-dimensional periodic potential (Fraser *et al.*, 1977). Model calculations have shown that the proposed method gives physically sensible results both for atomic desorption from the surface and for beam scattering. Since the form of the gas-surface potential has not been thoroughly investigated, no comparison with specific experimental data has been presented.

VI. Conclusions

The status of the art being as yet only broadly defined and rapidly changing, this review aims to be thought-provoking rather than conclusive. In this section, we try to summarize a number of relevant points emerging from the preceding analyses.

The cluster method seems to be a good intermediate stage on the way to the study of heterogeneous catalysis; many of the results reported in the previous sections were seen to illustrate the localization of the chemisorption phenomenon to the interaction of an adsorbate with a few metal atoms. If clusters of sufficient size are used, the results were seen to converge to the reproduction of metallic bulk properties.

The present limitations of *ab initio* methods (to simple atoms or to a small number of transition metal atoms) face the difficulties of semiempirical methods with parametrization. One way of using both approaches to a common goal would be to use *ab initio* to calibrate semiempirical methods. This is in principle even simpler than using empirical data to the same purpose, since, with the possible exception of UPS spectroscopy, experimental evidence in this field seldom lends itself to straightforward comparison with calculation. For instance, quantum mechanical results refer to absolute zero temperature, while chemisorption is studied in a wide range of temperatures, and catalytic processes are carried out at high temperatures. The need is felt for a way to introduce thermal effects in the calculations, as well as the effects of gas–solid energy exchange processes. For instance, model trajectories studies that have been carried out are essentially intended to generate ideas on what might occur, rather than to try to reproduce what actually occurs. The potential surface was assumed to be static; dynamic studies on vibrating surfaces might be the next step.

Some particular remarks concern the Xα–SW method. Although especially well suited to provide data for comparisons and assignments in the UPS spectroscopy field, this approach is not very satisfactory for studies of adsorption energetics. This is so mainly because of the muffin-tin approximation; refinement of the potential would, however, imply at present insurmountable difficulties in the calculations. The transition-state method is nonetheless very promising in the prediction of adsorption geometries.

Finally, although theoretical methods of the various kinds outlined in this review may prove to be promising, actual catalytic effects still seem to be out of their reach. Though yielding in principle the most detailed information on adsorbate structures, LEED itself requires ordered adsorbate domains of appreciable size, while this is certainly not the case in practical catalysis. The identification of adsorbed species is probably the best help that diffraction methods, together with appropriate spectroscopic techniques, can lend to catalysts.

ACKNOWLEDGMENT

Part of this work has been carried out while one of us (M.S.) was visiting the Department of Chemistry at the University of California in Berkeley. This author wishes to thank Professor Gabor Somorjai for his kind hospitality, for the many stimulating discussions, and for providing him with a wealth of unpublished data and information.

REFERENCES

Anders, L. W., Hansen, R. S., and Bartell, L. S. (1973). *J. Chem. Phys.* **59,** 5277.
Anders, L. W., Hansen, R. S., and Bartell, L. S. (1975). *J. Chem. Phys.* **62,** 1641.

Anderson, A. B., and Hoffmann, R. (1974a). *J. Chem. Phys.* **60,** 4271.
Anderson, A. B., and Hoffmann, R. (1974b). *J. Chem. Phys.* **61,** 4545.
Anderson, A. B. (1975a). *J. Chem. Phys.* **62,** 1187.
Anderson, A. B. (1975b). *J. Chem. Phys.* **63,** 4430.
Anderson, A. B. (1976a). *J. Chem. Phys.* **64,** 4046.
Anderson, A. B. (1976b). *J. Chem. Phys.* **65,** 1729.
Anderson, A. B. (1977). *J. Am. Chem. Soc.* **99,** 696.
Anderson, A. B. (1978a). *J. Chem. Phys.* **68,** 1744.
Anderson, A. B. (1978b). *J. Am. Chem. Soc.* **100,** 1153.
Baerends, E. J., and Ros, P. (1973). *Chem. Phys.* **2,** 52.
Baerends, E. J., Ellis, D. E., and Ros, P. (1973). *Chem. Phys.* **2,** 41.
Baetzold, R. C. (1971). *J. Chem. Phys.* **55,** 4355.
Baetzold, R. C. (1975). *Surf. Sci.* **51,** 1.
Baetzold, R. C. (1978). *J. Chem. Phys.* **68,** 555.
Bagus, P. S., and Schaefer, H. F. (1973). *J. Chem. Phys.* **58,** 1844.
Batra, I. P., and Bagus, P. S. (1975). *Solid State Commun.* **16,** 1097.
Batra, I. P., and Robaux, O. (1975). *Surf. Sci.* **49,** 653.
Bauschlicher, C. W., Liskow, D. H., Bender, C. F., and Schaefer, H. F. (1975). *J. Chem. Phys.* **62,** 4815.
Bauschlicher, C. W., Bender, C. F., Schaefer, H. F., and Bagus, P. S. (1976). *Chem. Phys.* **15,** 227.
Bennett, A. J., McCarroll, B., and Messmer, R. P. (1971a). *Surf. Sci.* **24,** 191.
Bennett, A. J., McCarroll, B., and Messmer, R. P. (1971b). *Phys. Rev. B* **3,** 1397.
Blint, R. J., Kunz, A. B., and Guse, M. P. (1975). *Chem. Phys. Lett.* **36,** 191.
Blint, R. J., Kunz, A. B., and Guse, M. P. (1976). *Chem. Phys. Lett.* **37,** 512.
Blyholder, G. (1974). *Surf. Sci.* **42,** 249.
Blyholder, G. (1975a). *J. Chem. Phys.* **62,** 3193.
Blyholder, G. (1975b). *J. Phys. Chem.* **79,** 756.
Bonzel, H. P. (1977). *Surf. Sci.* **68,** 236.
Boudart, M. (1977). *Proc. Int. Congr. Catal., 6th, 1976,* p. 1.
Brundle, C. R. (1975). *Surf. Sci.* **48,** 99.
Buchholz, J. C., and Somorjai, G. A. (1976). *Acc. Chem. Res.* **9,** 333.
Bunker, D. L. (1971). *Methods Comput. Phys.* **10,** 287.
Chen, B. H., Foyt, D. C., and White, J. M. (1977). *Surf. Sci.* **67,** 218.
Chesters, M. A., and Somorjai, G. A. (1975). *Annu. Rev. Mater. Sci.* **5,** 99.
Companion, A. L. (1973). *J. Phys. Chem.* **77,** 3085.
Companion, A. L. (1976a). *Chem. Phys.* **14,** 1.
Companion, A. L. (1976b). *Chem. Phys.* **14,** 7.
Condon, E. U., and Shortley, G. H. (1951). "The Theory of Atomic Spectra," Chapter 6. Cambridge Univ. Press, London and New York.
Danese, J. B. (1974). *J. Chem. Phys.* **61,** 3071.
Danese, J. B., and Connolly, J. W. D. (1974). *J. Chem. Phys.* **61,** 3063.
Demuth, J. E. (1977). *Chem. Phys. Lett.* **45,** 12.
Demuth, J. E., Jepsen, D. W., and Marcus, P. M. (1973). *Phys. Rev. Lett.* **31,** 540.
Deuss, H., and van der Avoird, A. (1973). *Phys. Rev. B* **8,** 2441.
Dovesi, R., Pisani, C., Ricca, F., and Roetti, C. (1976a). *Chem. Phys. Lett.* **39,** 103.
Dovesi, R., Pisani, C., Ricca, F., and Roetti, C. (1976b). *J. Chem. Phys.* **65,** 3075.
Dovesi, R., Pisani, C., Ricca, F., and Roetti, C. (1976c). *J. Chem. Phys.* **65,** 4116.
Eastman, D. E., and Cashion, J. K. (1971). *Phys. Rev. Lett.* **27,** 1520.
Elian, M., and Hoffmann, R. (1975). *Inorg. Chem.* **14,** 1058.

Ellis, D. E., and Painter, G. S. (1970). *Phys. Rev. B* **2**, 2887.

Ellis, D. E., Adachi, H., and Averill, F. W. (1976). *Surf. Sci.* **58**, 497.

Ellis, D. E., Baerends, E. J., Adachi, H., and Averill, F. W. (1977). *Surf. Sci.* **64**, 649.

Fassaert, D. J. M., and van der Avoird, A. (1976a). *Surf. Sci.* **55**, 291.

Fassaert, D. J. M., and van der Avoird, A. (1976b). *Surf. Sci.* **55**, 313.

Fassaert, D. J. M., Verbeek, H., and van der Avoird, A. (1972). *Surf. Sci.* **29**, 501.

Fraser, S., Murrell, J. N., and Mason, R. (1977). *Chem. Phys. Lett.* **46**, 399.

Gallezot, P., Bienenstock, A. I., and Boudart, M. (1978). *Nouv. J. Chim.* **2**, 263.

Gaspar, R. (1954). *Acta Phys. Acad. Sci. Hung.* **3**, 263.

Gavezzotti, A., and Simonetta, M. (1977). *Chem. Phys. Lett.* **48**, 434.

Gavezzotti, A., and Simonetta, M. (1979). *In* "Catalysis in Chemistry and Biochemistry. Theory and Experiment" (B. Pullman, ed.), p. 241. Reidel, Dordrecht.

Gelb, A., and Cardillo, M. J. (1976). *Surf. Sci.* **59**, 128.

Gelb, A., and Cardillo, M. J. (1977). *Surf. Sci.* **64**, 197.

Goddard, W. A., Walch, S. P., Rappé, A. K., Upton, T. H., and Melius, C. F. (1977). *J. Vac. Sci. Technol.* **14**, 416.

Gregory, A. R., and Silbey, R. (1977). *Chem. Phys. Lett.* **50**, 500.

Gregory, A. R., Gelb, A., and Silbey, R. (1978). *Surf. Sci.* **74**, 497.

Grimley, T. B. (1975). *In* "Electronic Structure and Reactivity of Metal Surfaces" (E. G. Derouane and A. A. Lucas, eds.), p. 113. Plenum, New York.

Hayns, M. R. (1975). *Theor. Chim. Acta* **39**, 61.

Heijser, W., Van Kessel, A. T., and Baerends, E. J. (1976). *Chem. Phys.* **16**, 371.

Henderson, G. A., Das, G., and Wahl, A. C. (1975). *J. Chem. Phys.* **63**, 2805.

Herman, F., and Skillman, S. (1963). "Atomic Structure Calculations." Prentice-Hall, Englewood Cliffs, New Jersey.

Hermann, K., and Bagus, P. S. (1977). *Phys. Rev. B* **16**, 4195.

Hoffmann, R. (1963). *J. Chem. Phys.* **39**, 1397.

Ibach, H., Hopster, H., and Sexton, B. (1977a). *Appl. Phys.* **14**, 21.

Ibach, H., Hopster, H., and Sexton, B. (1977b). *Appl. Surf. Sci.* **1**, 14.

Johnson, K. H. (1966). *J. Chem. Phys.* **45**, 3085.

Johnson, K. H. (1972). *Adv. Quantum Chem.* **7**, 143.

Johnston, H. S., and Parr, C. (1963). *J. Am. Chem. Soc.* **85**, 2544.

Jona, F. (1977). *Surf. Sci.* **68**, 204.

Jordan, R. M., and Kaufman, F. (1975). *J. Chem. Phys.* **63**, 1691.

Kesmodel, L. L., Stair, P. C., Baetzold, R. C., and Somorjai, G. A. (1976). *Phys. Rev. Lett.* **36**, 1316.

Kesmodel, L. L., Baetzold, R. C., and Somorjai, G. A. (1977). *Surf. Sci.* **66**, 299.

Kesmodel, L. L., Dubois, L. H., and Somorjai, G. A. (1979). *J. Chem. Phys.* **70**, 2180.

Kobayashi, H., Yoshida, S., Fukui, K., Tarama, K., and Kato, H. (1978). *Chem. Phys. Lett.* **53**, 457.

Kohn, W., and Rostoker, N. (1954). *Phys. Rev.* **94**, 1111.

Kohn, W., and Sham, L. J. (1965). *Phys. Rev. A* **140**, 1133.

Korringa, J. (1947). *Physica (Utrecht)* **13**, 392.

Kunz, A. B., Mickish, D. J., and Deutsch, P. W. (1973). *Solid State Commun.* **13**, 35.

Lee, T. H., and Rabalais, J. W. (1978). *Surf. Sci.* **75**, 29.

Levine, R. D., and Bernstein, R. B. (1974). "Molecular Reaction Dynamics," p. 84. Oxford Univ. Press, London and New York.

Li, C. H., and Connolly, J. W. D. (1977). *Surf. Sci.* **65**, 700.

McCreery, J. H., and Wolken, G. (1975a). *J. Chem. Phys.* **63**, 2340.

McCreery, J. H., and Wolken, G. (1975b). *J. Chem. Phys.* **63**, 4072.

McCreery, J. H., and Wolken, G. (1976a). *J. Chem. Phys.* **64,** 2845.

McCreery, J. H., and Wolken, G. (1976b). *Chem. Phys. Lett.* **39,** 478.

McCreery, J. H., and Wolken, G. (1976c). *J. Chem. Phys.* **65,** 1310.

McCreery, J. H., and Wolken, G. (1977). *J. Chem. Phys.* **66,** 2316.

Melius, C. F., Moskowitz, J. W., Mortola, A. P., Baillie, M. B., and Ratner, M. A. (1976). *Surf. Sci.* **59,** 279.

Messmer, R. P. (1977). *In* "Semiempirical Methods of Electronic Structure Calculation. Part B: Applications" (G. A. Segal, ed.), pp. 215–246. Plenum, New York.

Messmer, R. P., and Salahub, D. R. (1977). *In* "Computers in Chemical Education and Research" (E. V. Luclena, N. H. Sabelli, and A. C. Wahl, eds.), p. 171. Plenum, New York.

Messmer, R. P., Tucker, C. W., and Johnson, K. H. (1974). *Surf. Sci.* **42,** 341.

Messmer, R. P., Tucker, C. W., and Johnson, K. H. (1975). *Chem. Phys. Lett.* **36,** 423.

Messmer, R. P., Knudson, S. K., Johnson, K. H., Diamond, J. B., and Yang, C. Y. (1976). *Phys. Rev. B* **13,** 1396.

Messmer, R. P., Salahub, D. R., Johnson, K. H., and Yang, C. Y. (1977). *Chem. Phys. Lett.* **51,** 84.

Mulliken, R. S. (1955). *J. Chem. Phys.* **23,** 1833.

Nishida, M. (1978). *Surf. Sci.* **72,** 589.

Pendry, J. B. (1974). "Low Energy Electron Diffraction Theory." Academic Press, New York.

Pople, J. A., and Beveridge, D. L. (1970). "Approximate Molecular Orbital Theory." McGraw-Hill, New York.

Pople, J. A., Santry, D. P., and Segal, G. A. (1965). *J. Chem. Phys.* **43,** S129.

Porter, R. N., and Karplus, M. (1964). *J. Chem. Phys.* **40,** 1105.

Rösch, N., and Menzel, D. (1976). *Chem. Phys.* **13,** 243.

Rösch, N., and Rhodin, T. N. (1974). *Phys. Rev. Lett.* **32,** 1189.

Rösch, N., Klemperer, W. G., and Johnson, K. H. (1973). *Chem. Phys. Lett.* **23,** 149.

Rosén, A., Ellis, D. E., Adachi, H., and Averill, F. W. (1976). *J. Chem. Phys.* **65,** 3629.

Ruiz-Vizcaya, M. E., Novaro, O., Ferreira, J. M., and Gomez, R. (1978). *J. Catal.* **51,** 108.

Salem, L., and Leforestier, C. (1979). *Surf. Sci.* **82,** 390.

Santry, D. P., and Segal, G. A. (1967). *J. Chem. Phys.* **47,** 158.

Sato, S. (1955). *J. Chem. Phys.* **23,** 592.

Schaefer, H. F. (1977). *Acc. Chem. Res.* **10,** 287.

Schiff, L. I. (1949). "Quantum Mechanics," p. 78. McGraw-Hill, New York.

Scott, P. R., and Richards, W. G. (1974a). *J. Phys. B* **7,** 500.

Scott, P. R., and Richards, W. G. (1974b). *J. Phys. B* **7,** 1679.

Scott, P. R., and Richards, W. G. (1975). *J. Chem. Phys.* **63,** 1690.

Slater, J. C. (1931). *Phys. Rev.* **38,** 1809.

Slater, J. C. (1951). *Phys. Rev.* **81,** 385.

Slater, J. C. (1965). *J. Chem. Phys.* **43,** S228.

Slater, J. C. (1972). *Adv. Quantum Chem.* **6,** 1.

Slater, J. C., and Wood, J. H. (1971). *Int. J. Quantum Chem.* 4S, 3.

Somorjai, G. A. (1977). *Angew. Chem., Int. Ed. Engl.* **16,** 92.

Stair, P. C., and Somorjai, G. A. (1976). *Chem. Phys. Lett.* **41,** 391.

Stair, P. C., and Somorjai, G. A. (1977). *J. Chem. Phys.* **66,** 2036.

Stoll, H., and Preuss, H. (1977). *Surf. Sci.* **65,** 229.

Walch, S. P., and Goddard, W. A., III (1978). *Surf. Sci.* **72,** 645.

Weinberg, W. H., and Merrill, R. P. (1972). *Surf. Sci.* **33,** 493.

Weinberg, W. H., and Merrill, R. P. (1973). *Surf. Sci.* **39,** 206.

Weinberg, W. H., Deans, H. A., and Merrill, R. P. (1974). *Surf. Sci.* **41**, 312.
Weinbergen, P., and Schwartz, K. (1975). *Int. Rev. Sci., Ser. Two,* **1**, 255.
Wertheim, G. K., and Hufner, S. (1972). *Phys. Rev. Lett.* **28**, 1028.
Wood, J. C. (1978). *Surf. Sci.* **71**, 548.
Yang, C. Y., and Rabii, S. (1975). *Phys. Rev. A* **12**, 362.
Yates, J. T. (1974). *Chem. & Eng. News* **52**(34), 19.
Yu, H. L. (1977). *Phys. Rev. B* **15**, 3609.
Ziegler, T., and Rank, A. (1977). *Theor. Chim. Acta* **46**, 1.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 12

Open Shell SCF Theory: An *ab Initio* Study of Some Interstellar Molecules

RAMON CARBÓ
Departamento de Química Orgànica
Secció de Química Quàntica
Instituto Química de Sarrià
Barcelona, Spain

and

ODD GROPEN
Institute of Mathematical and
Physical Sciences
Department of Chemistry
University of Tromsø
Tromsø, Norway

I. Introduction

The recent advances in the photochemistry and spectroscopy of molecular systems (Sandorfy *et al.*, 1974; Okabe, 1978) as well as present-day computer development make the theoretical study of excited, radicalary, and ionic species more interesting. Also, from the point of view of experimental findings, the birth and growth of radioastronomy in the late 1960s has verified the existence of a remarkable set of molecular structures in some regions of interstellar space (Gordon and Snyder, 1973;

ISBN 0-12-034812-8

Cameron, 1973; Pinkau, 1974; Van Woerden, 1977). Interstellar medium conditions are in some instances very similar to these applicable to the implicit molecular environment in usual quantum chemical calculations, and in this context it is probable that quantum chemistry may be of considerable help in the study of the molecular parameters associated with interstellar molecules.

Although ground-state species are interesting and informative, the main problem from the point of view of interaction of theory and experiment concerns the *open shell structures:* excited states, radical fragments, and ions, which are difficult to reproduce and study in terrestrial laboratory conditions for a sufficiently long time. In this manner, the theoretical development of open shell techniques will provide, without doubt, a better understanding of the phenomena relevant to creation, transformation, and decomposition of excited states.

This chapter, claiming neither priority nor definitive form, is an attempt to describe in detail a *sufficiently general open shell SCF procedure* that will be: (1) easily programmable, (2) without insuperable difficulties during application, and (3) with such characteristics that the program structure may be implemented in any breed of computer; this programming trend is becoming widespread and may soon be imperative. Some *interstellar molecules* provide us with an amount of computational work sufficient for study of the appropriate features of the methodology. The first section will present a description of a restricted SCF open shell framework, where all the necessary hints will be given in order to facilitate implementation. A short survey of the interstellar molecule problem will follow this theoretical section, with the aim of justifying the chosen structures as an example of a computational application. The last two sections of this chapter will give, first, a survey of the numerical characteristics chosen in the application of the theory and, second, the results obtained, compared, when possible, with experimental and other calculations.

II. Open Shell SCF Theory

A. Historical Review

The *SCF open shell procedure* described here may be viewed as one of the last fruits of a quite old methodological tree. In order to outline the history we will give in brief some of the relevant references.

The initial work of Roothaan (1960) must be taken as the starting point which described the direction of future research. Related to this work are the early discussions of Huzinaga (1960, 1961) and the first attempt at a general view of the problem by Birss and Fraga (1963, 1964), which probably precluded the development of Dyadyuscha and Kuprievich (1965), who made an often overlooked, yet clever analysis of the open shell mathematical tools. Then a four-year lull is found in the literature. No new attempts to attack or analyze the subject are made until a study of Huzinaga (1969), followed by the article of Goddard *et al.* (1969), who clearly stated the full variational structure of the open shell problem. This opened the way for a series of articles by Peters (1972), Hirao and Nakatsuji (1973), Caballol *et al.* (1974), Hirao (1974), and Carbó *et al.* (1975) which established the present framework. The state of the art up to this point may be found, perhaps, in a recent monograph (Carbó and Riera, 1978), but the problem still is being studied, mainly as a result of the coupling operator concepts. Although these new developments are interesting, they will not be discussed here.

B. Expression of the Electronic Energy

In a common and general way, including *monoconfigurational* and some *multiconfigurational* cases, the electronic energy of a molecular system can be written (Carbó and Riera, 1978)

$$E = \sum_{i\in N} \omega_i h_{ii} + \sum_{i\in N}\sum_{j\in N} (\alpha_{ij} J_{ij} - \beta_{ij} K_{ij}), \tag{1}$$

where the constants $\{\omega_i\}$, $\{\alpha_{ij}\}$, and $\{\beta_{ij}\}$ are parameters of a given state and wave function (hereafter called *state parameters*), $\{h_{ii}\}$, $\{J_{ij}\}$, and $\{K_{ij}\}$ are the core, *Coulomb,* and *exchange integrals* defined over the set of active *molecular orbitals* (MO), whose indices appear in the summations, and N is the set of these occupied MO indices.

Representing the MO set by $\{\phi_i\}$ $(i \in N)$, the molecular integrals in Eq. (1) are defined as

$$h_{ii} = \langle i|h|i\rangle = \int \phi_i^* h \phi_i \, dV, \tag{2}$$

where h is the *core Hamiltonian* (in atomic units):

$$h = -\tfrac{1}{2}\nabla^2 - \sum_A (Z_A/r_A), \tag{3}$$

and

$$J_{ij} = (ii|jj) = \int\int \phi_i^*(1)\phi_j^*(2) r_{12}^{-1} \phi_i(1)\phi_j(2) \, dV_1 \, dV_2, \tag{4}$$

$$K_{ij} = (ij|ij) = \int\int \phi_i^*(1)\phi_j^*(1) r_{12}^{-1} \phi_i(2)\phi_j(2) \, dV_1 \, dV_2. \tag{5}$$

Equation (1), through a redefinition of integrals (4) and (5), can be rewritten with the aid of the so-called Coulomb (J_i) and exchange (K_i) operators. The definition of these operators, due to Roothaan (1951), is

$$J_i\phi_j = \int \phi_i^*(1)r_{12}^{-1}\phi_i(1)\, dV_1\phi_j(2), \tag{6}$$

$$K_i\phi_j = \int \phi_i^*(1)r_{12}^{-1}\phi_j(1)\, dV_1\phi_i(2). \tag{7}$$

In this manner one can write

$$J_{ij} = \langle j|J_i|j\rangle = \langle i|J_j|i\rangle, \tag{8}$$

$$K_{ij} = \langle j|K_i|j\rangle = \langle i|K_j|i\rangle. \tag{9}$$

Consequently, the energy expression (1) can be written

$$E = \sum_{i\in N} \langle i|\omega_i h + \sum_{j\in N} (\alpha_{ij}J_j - \beta_{ij}K_j)|i\rangle. \tag{10}$$

Furthermore, it should be noted that the equivalent equations (1) and (10) are subject to constraints related to the orthonormalization of the MO set $\{\phi_i\}$ $(i \in N)$, which we write as:

$$\langle i|j\rangle = \int \phi_i^*\phi_j\, dV = \delta_{ij}, \qquad \forall \quad i, j \in N. \tag{11}$$

The constraints (11) can be relaxed but the energy expression becomes more complicated than Eq. (1) (Löwdin, 1955).

At the same time, it should be noted that Eqs. (1) and (10) are related to a kind of wave function Φ which can be expressed as a linear combination of *Slater determinants:*

$$\Phi = \sum_I a_I\Phi_I, \tag{12}$$

subject to the conditions:

(a) $$\langle\Phi|\Phi\rangle = 1; \tag{13}$$

(b) if H is the *spinless electronic Hamiltonian,* then

$$\langle\Phi_I|H|\Phi_J\rangle = \sum_{i\in N} \left[\omega_i^{IJ}h_{ii} + \sum_{j\in N} (\alpha_{ij}^{IJ}J_{ij} - \beta_{ij}^{IJ}K_{ij})\right], \tag{14}$$

where $\{\omega_{ij}^{IJ}\}$, $\{\alpha_{ij}^{IJ}\}$, and $\{\beta_{ij}^{IJ}\}$ are constants depending on the nature of the corresponding Slater determinants Φ_I and Φ_J.

The state parameters in Eq. (1) will then be expressed as

$$\omega_i = \sum_I \sum_J a_I^* a_J \omega_i^{IJ}, \tag{15}$$

$$\alpha_{ij} = \sum_I \sum_J a_I^* a_J \alpha_{ij}^{IJ}, \tag{16}$$

$$\beta_{ij} = \sum_I \sum_J a_I^* a_J \beta_{ij}^{IJ}. \tag{17}$$

In most cases the integrals (14), when $I \neq J$, will be greatly simplified due to the nature of the functions Φ_I and Φ_J (Parr, 1963), with only a few terms remaining in the summations.

Another limitation of a SCF procedure attached to energy form (1) consists in the variational impossibility of using an energy functional corresponding to an excited state with the same symmetry as the ground state (Bethe and Salpeter, 1957). But, on the other hand, an energy equation such as (1) will be able to cope with the most common situations encountered in the study of open shell systems, including very complicated multiconfigurational cases or the simple monodeterminantal closed shell structures.

C. Euler Equations

Starting from Eq. (1) and constructing the augmented functional

$$L = E + \sum_i \sum_j \varepsilon_{ij}(\langle i|j\rangle - \delta_{ij}), \tag{18}$$

by means of the set $\{\varepsilon_{ij}\}$ of undetermined Lagrange multipliers, one can use a variational procedure in (18), which will give

$$\delta L = \delta E + \sum_i \sum_j \varepsilon_{ij}(\langle \delta i|j\rangle + \langle i|\delta j\rangle). \tag{19}$$

Recalling Eq. (1) one finds:

$$\begin{aligned} \delta E = \sum_i \omega_i\{\langle \delta i|h|i\rangle + \langle i|h|\delta i\rangle\} + \sum_i \sum_j \Big\{ & \alpha_{ij}\{\langle \delta i|J_j|i\rangle + \langle i|J_j|\delta i\rangle \\ & + \langle \delta j|J_i|j\rangle + \langle j|J_i|\delta j\rangle\} - \beta_{ij}\{\langle \delta i|K_j|i\rangle + \langle i|K_j|\delta_i\rangle \\ & + \langle \delta j|J_i|j\rangle + \langle j|J_i|\delta j\rangle\}\Big\}. \end{aligned} \tag{20}$$

Substituting (20) into (19) and taking into account that the stationary condition holds:

$$\delta L = 0, \tag{21}$$

one obtains two sets of equations ($\forall i \in N$) (*Euler equations*):

$$\tfrac{1}{2}\omega_i h + \sum_j (\alpha_{ij}J_j - \beta_{ij}K_j)|i\rangle = \sum_j \varepsilon_{ij}|j\rangle, \tag{22a}$$

$$\langle i|\tfrac{1}{2}\omega_i h + \sum_j (\alpha_{ij}J_j - \beta_{ij}K_j) = \sum_j \langle j|\varepsilon_{ij}. \tag{22b}$$

Equations (22) can be rewritten in an equivalent manner as

$$F_i|i\rangle = \sum_j \varepsilon_{ij}|j\rangle, \qquad \forall \quad i \in N, \tag{23a}$$

$$\varepsilon_{ij}^* = \varepsilon_{ji}, \qquad \forall \quad i, j \in N, \tag{23b}$$

where the operators $\{F_i\}$ $(i \in N)$, the *Fock operators*, are obviously defined by means of:

$$F_i = \tfrac{1}{2}\omega_i h + \sum_j (\alpha_{ij} J_j - \beta_{ij} K_j). \tag{24}$$

The conditions (23b) can be deduced from Eqs. (22), multiplying (22a) on the left by $\langle j|$ and (22b) on the right by $|i\rangle$; then one has, recalling the orthogonal nature of the MOs

$$\langle j|F_i|i\rangle = \varepsilon_{ij}, \tag{25a}$$

$$\langle i|F_i|j\rangle = \varepsilon_{ji}. \tag{25b}$$

Taking Eq. (25b) and conjugating both sides,

$$\varepsilon_{ji}^* = \langle j|F_i^+|i\rangle, \tag{26}$$

as done with Eq. (24) and definitions (6) and (7), it is easy to show that

$$F_i^+ = F_i. \tag{27}$$

If one considers the state parameter real, one can deduce that Eq. (26) is equivalent to (23b).

Equations (23) are the basis of the development of the formalism used here. Equation (23a) simply corresponds to a *null gradient condition on the energy*. Equations (23b) are conditions arising from the dual nature of the optimization problem. We will refer to them as the *Hermitian conditions on Lagrange multiplier matrix*.

D. Coupling Operator

There are many ways to solve Euler equations (23). Here we choose one possibility, which has as its ultimate goal the construction of a unique pseudosecular equation whose eigenvectors will constitute the MO set.

The null gradient part of the equations (23a) can be related to some operator R_0 (Caballol *et al.*, 1974) which transforms the equation for each Fock operator into a single expression. In order to do this one can define the *projectors over each MO* $|i\rangle$ as

$$P_i = |i\rangle\langle i|, \qquad \forall \quad i \in N, \tag{28}$$

Then recalling that, in practice, to the MO set there can be added a linearly independent set of vectors $\{|v\rangle\}$ $(v \in V)$, called *virtual orbitals,* which complete the basis set $\{|i\rangle\}$ $(i \in N)$ up to a given dimension, one can also define a projector

$$\Pi_i = P_i + \sum_{v \in V} |v\rangle\langle v|, \qquad \forall \quad i \in N, \tag{29}$$

over the subspace spanned by the ith MO and the virtual orbital set.

As the projectors (28) for any orbital $|p\rangle$ fulfill

$$P_i|p\rangle = \delta_{ip}|i\rangle, \tag{30}$$

one will obtain

$$\Pi_i|p\rangle = \delta_{ip}|i\rangle, \qquad \forall \quad p \in N \tag{31a}$$

$$\Pi_i|p\rangle = |p\rangle, \qquad \forall \quad p \in V. \tag{31b}$$

Using Eq. (23a) and the nature of the projectors $\{\Pi_i\}$, one can write

$$F_i\Pi_i|i\rangle = \sum_{j \in N} \varepsilon_{ij}|j\rangle, \qquad \forall \quad i \in N, \tag{32}$$

and multiplying both sides of (32) on the left by Π_i^+,

$$\Pi_i^+F_i\Pi_i|i\rangle = \sum_{j \in N} \varepsilon_{ij}\Pi_i^+|j\rangle, \qquad \forall \quad i \in N, \tag{33}$$

which, taking into account relations (31), can be transformed into

$$\Pi_i^+F_i\Pi_i|i\rangle = \varepsilon_{ii}|i\rangle, \qquad \forall \quad i \in N. \tag{34}$$

Defining a set of *projected Fock operators* $\{R_i\}$ as

$$R_i = \Pi_i^+F_i\Pi_i, \qquad \forall \quad i \in N, \tag{35}$$

Eq. (34) can be written as

$$R_i|i\rangle = \varepsilon_{ii}|i\rangle, \qquad \forall \quad i \in N. \tag{36}$$

Now, with operators (35) one can build up a unique operator

$$R_0 = \sum_{i \in N} \alpha_i R_i, \tag{37}$$

where $\{\alpha_i\}$ $(i \in N)$ is a set of arbitrary real constants which will fulfill the *pseudosecular equation*

$$R_0|i\rangle = \lambda_i|i\rangle, \qquad \forall \quad i \in N, \tag{38}$$

with

$$\lambda_i = \alpha_i\varepsilon_{ii}, \qquad \forall \quad i \in N. \tag{39}$$

At the self-consistent stage, one can obtain the *spectral decomposition*

$$R_0 = \sum_{i \in N} \lambda_i P_i + \sum_{v \in V} \left\langle v \middle| \sum_i \alpha_i F_i \middle| v \right\rangle P_v, \tag{40}$$

where

$$P_v = |v\rangle\langle v|.$$

Equations (38) will be sufficient to obtain the MO set in all the cases where conditions (23b) are automatically satisfied. This is the case when, by the system structure, one has

$$\langle i|F_i|j\rangle = \langle i|F_j|j\rangle = 0. \tag{41}$$

Conditions (41) hold when the symmetry of the system makes the beginning MOs $|i\rangle$ and $|j\rangle$ orthogonal or when for a given subset $M \subset N$ $(i, j \in M) \rightarrow F_i = F_j$. In other cases the Hermitian nature of the Lagrange multiplier matrix must be taken into account.

In order to accomplish this in a practical manner, conditions (23b) can be written as

$$\varepsilon_{ij} - \varepsilon_{ji}^* = \langle i|F_j - F_i|j\rangle = 0, \qquad \forall \quad i, j \in N, \tag{42}$$

From (42) an operator τ can be defined, using $\{\lambda_{ij}\}$ $(i, j \in N)$ arbitrary scalars (Hirao, 1974),

$$\tau = \sum_i \sum_j \lambda_{ij} P_i^+ (F_j - F_i) P_j, \tag{43}$$

which, when conjugated, gives

$$\tau^+ = -\sum_i \sum_j \lambda_{ji}^* P_i^+ (F_j - F_i) P_j. \tag{44}$$

From this a symmetric operator, called the *Hirao operator*, can be constructed:

$$R_{\mathrm{H}} = \tfrac{1}{2}(\tau + \tau^+) = \sum_i \sum_j \sigma_{ij} P_i^+ (F_j - F_i) P_j, \tag{45}$$

where

$$\sigma_{ij} = \tfrac{1}{2}(\lambda_{ij} - \lambda_{ji}^*). \tag{46}$$

Consequently

$$\sigma_{ij}^* = \tfrac{1}{2}(\lambda_{ij}^* - \lambda_{ji}) = -\sigma_{ji} \tag{47}$$

or one can say that the set of constants $\{\sigma_{ij}\}$ form a skew-Hermitian matrix. At self-consistency the operator R_{H} fulfills conditions (23b), because if $k, l \in N$, then

$$\langle k|R_{\mathrm{H}}|l\rangle = \sigma_{kl} \langle k|F_l - F_k|l\rangle. \tag{48}$$

Thus, if one constructs the operator

$$R_1 = R_0 + R_H, \tag{49}$$

and finds the self-consistent solutions of

$$R_1|i\rangle = \lambda_i|i\rangle, \qquad \forall \quad i \in N, \tag{50}$$

then

$$\langle k|R_1|l\rangle = \langle k|R_0|l\rangle + \langle k|R_H|l\rangle = \delta_{kl}\lambda_l. \tag{51}$$

In this manner one can assure that the MO set diagonalizes the operator R_1. Therefore, if $k \neq l$, then

$$\langle k|R_1|l\rangle = 0, \tag{52}$$

and taking Eqs. (40) and (48) into account, one has

$$\langle k|R_1|l\rangle = \langle k|R_H|l\rangle = \sigma_{kl}\langle k|F_l - F_k|l\rangle = 0. \tag{53}$$

when $\sigma_{kl} \neq 0$, conditions (42) are met for all MO pairs. In this sense, operator (49) and its pseudosecular equation (50) transform Eqs. (23) and allows the MOs to be obtained which make the functional (1) extremum. The operator R_1, defined in Eq. (49), will be called the *coupling operator.* To the coupling operator R_1 one can add another kind of operator which leaves the eigenvectors of Eq. (50), and thus the energy expression (1), invariant.

One can take an arbitrary set of real constants $\{\beta_k\}$ ($k \in N \cup V$) and construct a linear combination of projectors

$$Q = \sum_{k \in N \cup V} \beta_k P_k. \tag{54}$$

The new operator

$$R = R_1 + Q \tag{55}$$

fulfills

$$R|i\rangle = \mu_i|i\rangle, \qquad i \in N, \tag{56}$$

with

$$\mu_i = \lambda_i + \beta_i = \alpha_i \varepsilon_{ii} + \beta_i, \tag{57}$$

so the action of Q on the pseudosecular equation of R resides in an *eigenvalue shift* by means of the parameters $\{\beta_i\}$ (Carbó *et al.*, 1977b).

E. Matrix Representation: LCAO Formalism

The MO set $\{\phi_i\}$ can be represented by means of an *atomic orbital* (AO) basis set $\{\chi_\mu\}$, through the linear combinations

$$\phi_i = \sum_{\mu} C_{i\mu}\chi_{\mu}. \tag{58}$$

Then, using $\{\mathbf{C}_i\}$ as the column matrices isomorphic to $\{\phi_i\}$, one can define the matrices

$$\mathbf{P}_i = \mathbf{C}_i\mathbf{C}_i^{+}, \tag{59}$$

and in a similar fashion

$$\mathbf{\Pi}_i = \mathbf{P}_i + \sum_{v} \mathbf{P}_v. \tag{60}$$

The matrices (59) and (60), although not true projectors, are useful in constructing the *coupling operator matrix*

$$\mathbf{R} = \mathbf{R}_0 + \mathbf{R}_{\mathrm{H}} + \mathbf{Q}, \tag{61}$$

with obvious definitions for the matrices $\mathbf{R}_0$, $\mathbf{R}_{\mathrm{H}}$, and $\mathbf{Q}$, which, in order to be constructed, require only the *Fock operator matrix* representation, namely,

$$F_{i,\mu\nu} = \tfrac{1}{2}\omega_i h_{\mu\nu} + \sum_{\lambda}\sum_{\sigma}\{A_{i,\lambda\sigma}(\mu\nu|\lambda\sigma) - B_{i,\lambda\sigma}(\mu\sigma|\nu\lambda)\}, \tag{62}$$

where the matrix sets $\{\mathbf{A}_i\}$ and $\{\mathbf{B}_i\}$ are defined as

$$\mathbf{A}_i = \sum_{j} \alpha_{ij}\mathbf{P}_j, \tag{63}$$

$$\mathbf{B}_i = \sum_{j} \beta_{ij}\mathbf{P}_j. \tag{64}$$

The correct matrix representation of the final coupling operator must be written, however, as

$$\mathbf{R}_S = \mathbf{SRS}, \tag{65}$$

taking into account the *metric* of the AO basis set $\{\chi_{\mu}\}$, defined here as the integral

$$S_{\mu\nu} = \int \chi_{\mu}^{*}\chi_{\nu}\, dV. \tag{66}$$

The matrix representation of the pseudosecular equation (56), from which the vectors $\{\mathbf{C}_i\}$ can be obtained, can be written as

$$\mathbf{R}_S\mathbf{C} = \mathbf{SCM} \tag{67}$$

with

$$\mathbf{C} = (\mathbf{C}_1\mathbf{C}_2 \cdots \mathbf{C}_i \cdots) \qquad \text{and} \qquad \mathbf{M} = \mathrm{diag}(\mu_i,\ i \in N).$$

Equation (67) is identical to

$$\mathbf{RSC} = \mathbf{CM}, \tag{68}$$

so using an appropriate decomposition of the metric matrix S, with a general structure

$$\mathbf{S} = \mathbf{T}^{+}\mathbf{T}, \tag{69}$$

where T is an upper triangular matrix, Eq. (68) can be put in the form

$$\mathbf{R}_T\mathbf{D} = \mathbf{DM}, \tag{70}$$

with

$$\mathbf{R}_T = \mathbf{TRT}^{+} \tag{71}$$

and

$$\mathbf{D} = \mathbf{TC}. \tag{72}$$

The computational structure for solution of equations of this kind is well known (Wilkinson and Reinsch, 1971), and the algorithms are optimally programmed, so no further discussion on this subject is necessary.

In the Appendix can be found a table with the most common *state parameters,* as well as the practical energy and operator forms dealing with the usual open shell systems.

III. Interstellar Molecules

A. Historical Outline

As a justification of our interest in the study of *interstellar molecules,* we present here a brief survey of the historical development of these structures. Besides the computational role played by these molecules in this chapter, as a means of showing the performance of the already described SCF scheme in reaching variationally correct and accessible excited states, great interest in this subject was aroused by the development of *radioastronomical detection techniques.* Many of the molecules studied here have been also treated in many ways by other authors from a quantum mechanical point of view, as will be discussed later, when results are presented. Some interstellar structures have not been treated in the current *ab initio* literature (e.g., see Richards *et al.,* 1971, 1974, 1978); others, such as formaldehyde, have been the object of numerous studies, yet no comprehensive work, with a coherent methodology and basis set, has been performed on all of these molecular species to date. This may in itself be ample justification for our interest, but also one can find in the history of interstellar molecules—the radioastronomical history—an even better reason.

The early discovery of the first interstellar molecules was made in the period 1937–1949, using *optical techniques*, and evidence of the presence of diatomic structures in interstellar space was obtained (Adams, 1948). The interest in finding new molecules was aroused by this preliminary search and, some time later, the OH radical emission line was detected by means of *radioastronomical observation* (Weinreb *et al.*, 1963). The actual list of known structures has about forty items and contains a wide variety of molecules ranging from diatomics to medium-sized (up to ten second-row atoms) compounds. Each year this set grows in size. *General discussions* on this subject were presented within a radioastronomical point of view by Buhl and Snyder (1971) and Buhl (1973), and *general characteristics* were covered by Gordon and Snyder (1973), Margulis (1973), Shimizu (1973), Solomon (1973), Drake (1974), Mezger (1975), Nicholls (1977), and Gammon (1978). The term *interstellar chemistry* was coined by Khare and Sagan (1973), and was also studied with respect to *formation and related reactions of interstellar matter* by McNally (1972), Dalgarno and Black (1976), Herbst and Klemperer (1976), Watson (1976, 1977, 1978), Mitchell *et al.* (1978), and Sagan and Khare (1979). The related problem of the *origin of life* in connection with interstellar molecules was developed in the articles by Werner (1971), Buhl and Ponnamperuma (1971), Buhl (1974), and Ponnamperuma (1978). Although an initial attempt to promote interest in the *link between quantum chemical points of view and interstellar molecules problems* was published in 1976 (Caballol *et al.*, 1976), sufficient information on *recent developments* and trends can be found in the article by Green and Herbst (1979).

B. Interstellar Chemistry

From the theoretical point of view, interstellar chemistry may be studied in a wide variety of ways, which can be summarized as follows: (a) the computation of *properties* and molecular parameters of known interstellar structures; (b) the computation of *formation or decomposition paths* from known or plausible molecular fragments in order to establish the direction of the search for new structures; (c) the search for information on *possible structures* (isomers or others) not yet detected in order to interpret interstellar emission lines not yet identified. Here we have chosen the first point of view.

C. The Chosen Structures

From the set of interstellar molecules we have taken those which, being medium-sized, can be useful for the theoretical purposes outlined earlier. The structures studied are presented in Table I. Selection of these structures has been governed partly by their relevance to *cosmochemistry*

TABLE I

DESCRIPTION OF MOLECULES STUDIED

Formaldehyde	HCHO	Acetaldehyde	CH_3CHO
Formic acid	HCOOH	Methylamine	CH_3NH_2
Methanol	CH_3OH	Cyanamide	NH_2CN
Dimethyl ether	CH_3OCH_3	Formamide	NH_2CHO
Ethanol	CH_3CH_2OH	Methylenimine	CH_2NH

and partly by their potential for comparison with experimental and previously published data.

The ten structures have a wide variety of organic functions—aldehyde, alcohol, ether, amine, acid, cyano, imine—which adds an ulterior interest to this theoretical description. Although some molecules in Table I have been widely studied from many points of view (e.g., formaldehyde of formamide), others (e.g., methylamine), as far as we know, have received little or no attention from the theoretical side.

D. Computational Details

All molecules chosen have a *closed shell ground state*, and the *geometry* attached for this state is, if possible, in accordance with available experimental data. The nonredundant coordinates for the atoms are presented for all molecules in Table II, where, at the same time, *gross atomic populations* are given for the four states (ground state, first excited singlet and triplet, and positive ion) studied on each structure.

Except for formaldehyde, whose symmetry was C_{2v}, all the remaining structures possess C_S group characteristics; in addition to the ground state, the excited singlet and triplet states, as well as the positive ion have been studied. In all calculations the geometry used was the same as the one employed for the ground state, so our results must be interpreted as *vertical excitations or ionizations*. Thus, a source of error in the results presented may be associated with the lack of *geometric relaxation*. Open Shell states were studied from a *monoconfigurational* point of view, so differences in correlation energy should also be expected between the ground state and these states.

The *basis set* chosen is formed by Cartesian GTO functions of *double zeta quality,* taking a (9s, 5p) set contracted to a (3s, 2p) basis of atomic functions, as described by Dunning and Hay (1977). To this basis a diffuse s-type orbital was added to the second-row atoms C, N and O.

Calculations were performed by means of the MOMPOS system (Carbó *et al.,* 1977a), which uses a modified version of the program

TABLE II

NUCLEAR COORDINATES (a.u.) AND GROSS ATOMIC POPULATIONS

Molecule	x	y	z	$Q(\mathrm{GS})^a$	$Q(\mathrm{S})^b$	$Q(\mathrm{T})^c$	$Q(\mathrm{PI})^d$
Formaldehyde							
C	0	0	0	5.93	6.29	6.34	6.01
O	0	0	2.2835	8.35	8.14	8.10	7.09
H1	1.793	0	−1.110	0.85	0.79	0.78	0.65
Formic acid							
O1	1.086	−2.087	0	8.47	8.18	8.17	7.95
O2	1.145	2.199	0	8.50	8.51	8.50	8.34
C	0	0	0	5.61	5.88	5.91	5.50
H1	−2.050	0	0	0.79	0.77	0.76	0.62
H2	2.914	1.897	0	0.62	0.66	0.66	0.58
Methanol							
O	0	0	2.697	8.62	8.49	8.44	8.02
C	0	0	0	6.16	6.47	6.50	6.30
H1	0	1.947	−0.706	0.88	0.79	0.79	0.75
H2	1.686	−0.973	−0.706	0.84	0.73	0.72	0.69
H4	0	1.710	3.281	0.67	0.81	0.83	0.54
Dimethyl ether							
O	0	0	0	8.55	7.52	7.52	8.03
C1	1.507	2.213	0	6.16	7.05	7.04	6.30
C2	1.507	−2.213	0	6.17	7.04	7.04	6.33
H1	3.533	1.801	0	0.87	0.74	0.74	0.70
H2	1.169	3.411	1.651	0.84	0.73	0.73	0.72
H4	0.381	−3.947	0	0.83	0.74	0.74	0.71
H5	2.745	−2.338	1.651	0.87	0.73	0.73	0.72
Ethanol							
O	2.544	0	3.789	8.61	8.20	8.26	8.02
C1	0	0	0	6.42	6.10	6.19	6.42
C2	0	0	2.989	6.02	7.03	6.73	6.23
H	−1.947	0	−0.688	0.86	0.82	0.83	0.78
H	0.974	1.686	−0.688	0.84	0.81	0.81	0.79
H	−0.974	1.686	3.587	0.87	0.72	0.73	0.71
H	2.529	0	5.612	0.66	0.79	0.91	0.54

TABLE II (*Continued*)

Molecule	*x*	*y*	*z*	*Q*(GS)[a]	*Q*(S)[b]	*Q*(T)[c]	*Q*(PI)[d]
			Acetaldehyde				
O	0	1.989	3.984	8.39	8.17	8.13	7.92
C1	0	0	0	6.49	6.44	6.44	6.50
C2	0	0	2.836	5.81	6.12	6.17	5.77
H1	0	1.947	−0.706	0.80	0.83	0.83	0.74
H2	1.686	−0.973	−0.706	0.82	0.83	0.83	0.71
H4	0	−1.823	3.888	0.85	0.78	0.77	0.63
			Methylamine				
N	0	0	2.785	6.95	6.91	6.95	6.37
C	0	0	0	6.32	6.34	6.32	7.20
H	1.947	0	−0.689	0.87	0.78	0.78	0.73
H	−0.974	1.686	−0.688	0.84	0.80	0.81	0.74
H	0.828	1.562	3.505	0.75	1.17	1.15	0.62
			Cyanamide				
N1	0	0	2.226	7.20	7.22	7.17	6.82
N2	0	0	−2.509	7.58	7.46	7.46	7.32
C	0	0	0	5.84	5.91	5.93	5.63
H1	1.537	0	−3.397	0.68	0.70	0.70	0.61
			Formamide				
O	−1.257	1.869	0	8.44	8.14	8.12	8.15
C	0	0	0	5.69	5.99	6.02	5.77
N	2.600	0	0	7.67	7.68	7.69	7.20
H1	−0.821	−1.914	0	0.85	0.78	0.78	0.72
H2	3.475	1.705	0	0.67	0.70	0.70	0.57
H3	3.515	−1.658	0	0.69	0.69	0.69	0.59
			Methylenimine				
C	0	0	0	6.16	6.37	6.45	6.01
N	2.415	0	0	7.42	7.38	7.29	7.08
H1	3.644	1.502	0	0.73	0.68	0.68	0.58
H2	−1.387	1.594	0	0.92	0.82	0.82	0.69
H3	−1.853	−0.898	0	0.78	0.75	0.74	0.62

[a] Ground state.

[b] Excited singlet.

[c] Excited triplet.

[d] Positive ion.

MOLECULE for integral calculation (Almlöf, 1972), whereas the SCF part, fully programmed by the authors, and adapted to integral and symmetry handling of MOLECULE, is based on the theory outlined in Section II. *Double precision* computations were carried out throughout the SCF cycles. Convergence in all cases reached 10^{-8} in the Euclidean norm of the difference between density matrices in two successive iterations.

Medium-sized structures were chosen, ranging from 4 to 9 atoms, or preferably from 24 to 42 AO, in our previously described double zeta basis set. Small interstellar molecules like OH, H_2O, etc. were discarded because they have been studied extensively, and for this reason are of somewhat less interest in quantum chemistry. Bigger structures were discarded due to problems of time. At the same time, linear or highly symmetrical structures with the appropriate size were also excluded due to the complexity of the possible attached excited states.

The different molecular states were computed by a choice of trial vectors and the freezing of the appropriate state symmetry by means of a *shift operator.* The use of shifting is very time-consuming—the SCF-iteration time increases and convergence is substantially slower. Convergence is assured, unless intrinsic divergent states are encountered, in cases where oscillating behavior is present by the use of shifting. Sometimes a considerable shift must be used at the start, so the convergence rate still is decreased. For this reason we found it convenient to decrease the shift as convergence progresses, with some simple criterion. We found that if a modification such as

$$\text{new shift} = \text{old shift}/(\text{iteration number})^{\alpha}$$

is used, where α is a control parameter which can be used to decrease shift values at any convenient speed per iteration step, then one can obtain the optimal SCF performance for each state and molecule.

Table III shows the time spent for the integral part and the SCF cycle timing for ground state, and excited singlet and triplet; time is roughly doubled in each state as the number of shells increases one unit, so *relative time* between the three states will be approximately 1 : 2 : 3 for each cycle. For each molecule and state studied the number of SCF iterations is also given. Thus the reader can have a good knowledge of the cost for each calculation; a CDC 371 computer was used here.

Furthermore, as it should be expected, computation time is very sensitive to the number of basis functions used. The number of SCF cycles is increased when going from ground state to excited states, and does not seem to bear a direct relation to basis set size, but to the need to use larger shifts in order to get smooth convergence features.

TABLE III

COMPUTING TIME[a] FOR DIFFERENT STRUCTURES AND STATES

Molecule	Total integral		Time per SCF cycle/number of cycles		
	Number of AO	Calculation time	Ground state	Excited singlet	Excited triplet
Formaldehyde	24	351	36/16	108/20	65/20
Methylenimine	26	461	52/12	114/20	95/24
Methanol	28	671	85/11	246/23	181/20
Methylamine	30	840	125/12	221/15	220/15
Cyanamide	34	1140	128/9	404/28	242/25
Formic acid	34	1246	153/14	449/19	263/22
Formamide	36	1500	190/9	573/22	344/22
Acetaldehyde	38	1992	163/13	504/20	302/20
Dimethyl ether	42	3014	445/9	1323/18	816/21
Ethanol	42	3052	425/15	1279/25	803/22

[a] In seconds.

TABLE IV

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS AND PREVIOUS CALCULATIONS FOR FORMALDEHYDE

Property	Present work	Exptl.	Davidson (1976)	Goodman (1975)
Ground state ($^1A_1$)				
Total energy (a.u.)	−113.82798		−113.8951	−113.8211
Dipole moment (D)	3.07	2.34[a]		3.08
$\langle r^2 \rangle$	61.29			
First excited singlet ($^1A_2$)				
Excitation energy (eV)	2.83	3.8–4.2[b]	3.14	
Dipole moment	1.71	1.56[d]		
$\langle r^2 \rangle$	61.63			
First excited triplet ($^3A_2$)				
Excitation energy	2.46	3.30–3.60[b]	2.75	2.40
Dipole moment	1.56			1.56
$\langle r^2 \rangle$	61.63			
Positive ion ($^2A_2$)				
Ionization potential (eV)	9.69	10.87[c]		9.63
$\langle r^2 \rangle$	53.56			

[a] McClellan (1963).
[b] Chuitjian (1974).
[c] Brundle (1967).
[d] Freeman and Klemperer (1966).

E. Results and Discussion

1. *Introduction*

The *total energies, dipole moments, mean square distances of the electrons to the center of mass* $\langle r^2 \rangle$, *excitation energies,* and *ionization potentials* obtained are presented in Tables IV–XIII, with available experimental information and results of previous calculations. Also *atomic charges* can be found in Table II, as previously stated. In this manner we hope to give a sufficiently extensive summary of the calculations carried out on the molecules listed in Table I. In general, one can say that the results given in the tables may be comparable in accuracy to existing computations and in some cases provide for the first time information on excited states of various structures—theoretical information not yet found in the literature. In what follows some general remarks will be made on the results obtained for the parameters shown in the tables.

2. *Charges*

Upon excitation N and O atomic charges tend to be decreased or remain practically unchanged; on the contrary C atomic charges tend to be increased. It seems that this situation is sufficiently general, and in

TABLE V

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS AND PREVIOUS CALCULATIONS FOR FORMIC ACID

Property	Present work	Exptl.	Previous calc.
Ground state ($^1A'$)			
Total energy (a.u.)	−188.6952		−188.7119[d]
Dipole moment (D)	1.96	1.41[a]	1.419[d]
$\langle r^2 \rangle$	135.30		
First excited singlet ($^1A''$)			
Excitation energy (eV)	3.89	4.77–5.51[b]	5.46[f], 5.24[e]
Dipole moment	0.65		0.994[f]
$\langle r^2 \rangle$	134.48		
First excited triplet ($^3A''$)			
Excitation energy	3.61		5.11[f], 4.60[e]
Dipole moment	0.67		
$\langle r^2 \rangle$	134.44		
Positive ion ($^2A'$)			
Ionization potential (eV)	10.03	11.05[c]	12.53[e]
$\langle r^2 \rangle$	124.98		

[a] Kim *et al.* (1962). [b] Herzberg (1966). [c] Vedeneyev *et al.* (1966). [d] Ha and Keller (1975). [e] Demoulin (1976).

TABLE VI

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS AND PREVIOUS CALCULATIONS FOR METHANOL

Property	Present work	Experiment	Wadt and Goddard (1976)
Ground state ($^1A'$)			
Total energy (a.u.)	−115.00345		
Dipole moment (D)	2.44	1.71[a]	
$\langle r^2 \rangle$	85.08		
First excited singlet ($^1A''$)			
Excitation energy (eV)	6.38	6.68[b]	6.72
Dipole moment	3.28		
$\langle r^2 \rangle$	104.25		
First excited triplet ($^3A''$)			
Excitation energy	6.18	6.50[c]	6.23
Dipole moment	3.23		
$\langle r^2 \rangle$	103.39		
Positive ion ($^2A''$)			
Ionization potential (eV)	9.79	10.85[d]	
$\langle r^2 \rangle$	75.73		

[a] McClellan (1963).
[b] Tam and Brion (1979).
[c] Knoop *et al.* (1972).
[d] Harrison *et al.* (1959).

aldehydes, e.g., may be related to the well-known carbonyl behavior, in which a polarity change in the CO bond is usually expected.

Charges on H atoms do not suffer appreciable increments, except for in methylamine, for which there seems to be a charge transfer from N to the attached amine hydrogens. No great differences are encountered between charges on the excited singlets and triplets, where charge redistribution due to excitation is governed mainly by the open shell orbitals acting in a more important manner than spin structure.

3. *Dipole Moments*

The general trend of dipole moments upon excitation is to decrease for carbonyl molecules and to increase for noncarbonyl structures. In the first case, μ (excited triplet) $\leq$ μ (excited singlet) $<$ μ (ground state) was found; the sequence is reversed in the second group, with the exception of methylenimine, which behaves as a carbonyl-like structure from this point of view. When experimental results were available a correlation between

TABLE VII

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS AND PREVIOUS CALCULATIONS FOR DIMETHYL ETHER

Property	Present work	Exptl.	Wadt and Goddard (1976)
Ground state ($^1A'$)			
Total energy (a.u.)	−154.0049		
Dipole moment (D)	2.19	1.33[a]	
$\langle r^2 \rangle$	190.06		
First excited singlet ($^1A''$)			
Excitation energy (eV)	6.21	6.67[b]	6.76
Dipole moment	2.45		
$\langle r^2 \rangle$	231.89		
First excited triplet ($^3A''$)			
Excitation energy	6.18		6.31
Dipole moment	2.47		
$\langle r^2 \rangle$	231.58		
Positive ion ($^2A''$)			
Ionization potential (eV)	8.97	10.00[c]	12.74
$\langle r^2 \rangle$	176.94		

[a] McClellan (1963).
[b] Tam and Brion (1974).
[c] Bernecker and Long (1961).

TABLE VIII

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS FOR ETHANOL

Property	Present work	Exptl.
Ground state ($^1A'$)		
Total energy (a.u.)	−154.02290	
Dipole moment (D)	2.21	1.73[a]
$\langle r^2 \rangle$	197.00	
First excited singlet ($^1A''$)		
Excitation energy (eV)	6.32	6.83[b]
Dipole moment	3.85	
$\langle r^2 \rangle$	222.86	
First excited triplet ($^3A''$)		
Excitation energy	6.05	
Dipole moment	4.36	
$\langle r^2 \rangle$	218.78	
Positive ion ($^2A'$)		
Ionization potential (eV)	9.66	10.5[b]
$\langle r^2 \rangle$	183.78	

[a] McClellan (1963). [b] Harrison *et al.* (1959).

TABLE IX

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS FOR METHYLAMINE

Property	Present work	Exptl.
Ground state ($^1A'$)		
Total energy (a.u.)	−95.174309	
Dipole moment (D)	1.17	1.29[a]
$\langle r^2 \rangle$	95.88	
First excited singlet ($^1A''$)		
Excitation energy (eV)	8.05	
Dipole moment	3.14	
$\langle r^2 \rangle$	106.12	
First excited triplet ($^3A''$)		
Excitation energy	7.57	
Dipole moment	3.21	
$\langle r^2 \rangle$	105.69	
Positive ion ($^2A''$)		
Ionization potential (eV)	8.11	9.41[b]
$\langle r^2 \rangle$	89.90	

[a] McClellan (1963).
[b] Vedeneyev *et al.* (1966).

computed dipole moments and experimental values was obtained. The following linear relation was found:

$$\mu(\text{exptl.}) = 0.8517\mu(\text{calc.}) - 0.1661$$

for $n = 11$ and $r = 0.9652$. A very similar correlation can be reproduced from the more ample data in the article by Snyder (1974).

4. $\langle r^2 \rangle$

The $\langle r^2 \rangle$ parameter has been found to decrease or increase slightly upon ionization, or, in contrast, to dramatically increase upon excitation.

Carbonyl molecules show the first character, the remainder the second. In carbonyls a small decrease is found in some cases. Large changes may be interpreted as the result Rydberg-type excitations. Changes are, as a general rule, less pronounced in triplets than in singlets, but differences between both states are not very significant.

5. *Energies*

The excited states will, as mentioned earlier, suffer mainly from the lack of geometric optimization and the inclusion of correlation energy.

TABLE X

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS AND PREVIOUS CALCULATIONS FOR ACETALDEHYDE

Property	Present work	Exptl.	Ha and Keller (1975)
Ground state ($^1A'$)			
Total energy (a.u.)	−152.86079		−152.8916
Dipole moment (D)	3.46	2.69[a]	3.17
$\langle r^2 \rangle$	166.73		
First excited singlet ($^1A''$)			
Excitation energy (eV)	3.05	4.28[b]	4.27
Dipole moment	1.88		2.40
$\langle r^2 \rangle$	166.23		
First excited triplet ($^3A''$)			
Excitation energy	2.70		3.94
Dipole moment	1.72		
$\langle r^2 \rangle$	166.16		
Positive ion ($^2A'$)			
Ionization potential (eV)	9.02	10.18[c]	
$\langle r^2 \rangle$	154.03		

[a] Kim *et al.* (1962).
[b] Ha and Keller (1975).
[c] Vedeneyev *et al.* (1966).

TABLE XI

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS AND PREVIOUS CALCULATIONS FOR CYANAMIDE

Property	Present work	Snyder (1974)	Snyder and Basch (1972)
Ground state ($^1A'$)			
Total energy (a.u.)	−147.84687		−147.8437
Dipole moment (D)	4.86	4.27	4.91
$\langle r^2 \rangle$	143.11		142.63
First excited singlet ($^1A''$)			
Excitation energy (eV)	6.18		
Dipole moment	5.15		
$\langle r^2 \rangle$	143.58		
First excited triplet ($^3A''$)			
Excitation energy	5.52		
Dipole moment	4.75		
$\langle r^2 \rangle$	143.30		
Positive ion ($^2A'$)			
Ionization potential (eV)	10.03		
$\langle r^2 \rangle$	131.38		

TABLE XII

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS AND PREVIOUS CALCULATIONS FOR FORMAMIDE

Property	Present work	Exptl.	Davidson (1977)
Ground state ($^1A'$)			
Total energy (a.u.)	−168.87647		−168.9857
Dipole moment (D)	4.57	3.71[a]	4.10
$\langle r^2 \rangle$	147.16		
First excited singlet ($^1A''$)			
Excitation energy (eV)	4.57	5.65[b]	4.75 (5.63)[d]
Dipole moment	2.16		2.09
$\langle r^2 \rangle$	145.99		
First excited triplet ($^3A''$)			
Excitation energy	4.27	5.3[b]	4.46 (5.34)[d]
Dipole moment	2.02		1.91
$\langle r^2 \rangle$	145.93		
Positive ion ($^2A'$)			
Ionization potential (eV)	9.89	10.16[c]	11.9
$\langle r^2 \rangle$	134.21		

[a] McClellan (1963).
[b] Davidson (1977).
[c] Vedeneyev *et al.* (1966).
[d] With CI.

Correlation energy will, as a general rule, be larger for the ground state due to an additional electron pair to correlate; this error will then reduce the excitation energy by ≈ 1 eV. Geometric relaxation, on the other hand, is probably most important for the excited states, and this will also reduce the excitation energy. As these errors are opposite in sign one should expect the results to be not too far from experimental values.

In fact, excitation energies will appear to be displaced between 0.5–1 eV in comparison with experimental values. By inspecting the tables it is apparent that alcohols and ether show displacements of about 0.5 eV lower whereas aldehydes show a similar trend, only doubled. Previous calculations including correlation energy have improved this situation for both kinds of molecules, and CI calculations provide results very close to experimental values without including geometric optimization. These results however, may be, fortuitous because in many cases the geometry will be strongly perturbed following excitation. A closer study of the tables reveals that both the alcohols and ether studied have excitation

TABLE XIII

CALCULATED MOLECULAR PROPERTIES COMPARED WITH EXPERIMENTAL RESULTS AND PREVIOUS CALCULATIONS FOR METHYLENIMINE

Property	Present work	Johnson and Lovas (1972) (exptl.)	Sandorfy (1973)
Ground state ($^1A'$)			
Total energy (a.u.)	−93.913996		−93.8824
Dipole moment (D)	2.39	1.98	2.44
$\langle r^2 \rangle$	74.67		
First excited singlet ($^1A''$)			
Excitation energy (eV)	3.18		4.21
Dipole moment	1.36		
$\langle r^2 \rangle$	74.31		
First excited triplet ($^3A''$)			
Excitation energy	2.56		3.49
Dipole moment	1.28		
$\langle r^2 \rangle$	74.25		
Positive ion ($^2A'$)			
Ionization potential (eV)	8.40		9.05
$\langle r^2 \rangle$	63.56		

energies of the same magnitude. The calculated wave functions show that these transitions are due to the promotion of a π lone pair electron on oxygen to a delocalized σ orbital for all three structures. This is consistent with the charge displacement from the O atom to the neighboring C atom, as mentioned previously. In these molecules and for this kind of transition, because no bonding orbitals are fully involved, it should be expected that geometry changes in excitation will not be very strong. For the aldehydes, on the contrary, the situation is very different. Excitation on these molecules appears to involve the promotion of a lone pair on the oxygen to an antibonding π^* orbital covering the carbonyl group, which is localized in the CO bond as in formaldehyde or extended over the whole structure as in acetaldehyde or formamide. The effect is heavily pronounced when extension of the delocalized system is present. Thus there is a small gap between formaldehyde and acetaldehyde, which increases to 1.5 eV when comparing this last structure with formamide. In any case, it is rewarding to learn that, even with the previously mentioned lack of geometric optimization and correlation energy, the well-known trends are reproduced on the aldehydes. With respect to the three remaining structures little can be said since cyanamide, methylamine, and methylenimine are very different from the other molecules and also have little in common with each other. Because experimental information about these molecular species is scarce as far as excited states are concerned, our results, in comparison, are difficult to evaluate. It was difficult to find the lowest excited state for cyanamide, as this molecule has several plausible excitations. Primarily, one would expect a $n \rightarrow \pi^*$ or a $\pi \rightarrow n^*$ transition, but in our computational context we find the singlet $\pi \rightarrow n^*$ transition lower than the first one. Also, the difference between the two transitions $\pi \rightarrow a_1^*$ and $\pi \rightarrow b_1^*$ was only 0.66 eV, and the second one was the lowest.

A correlation was tried, finally, between the experimental singlet transition energies and the computed ones. The result:

$$S(\text{exptl.}) = 0.7509S(\text{calc.}) - 2.04,$$

with $n = 7$ and $r = 0.9928$, shows a fairly good trend, which confirms our valuation of the reasonably accurate predictions of monoconfigurational open shell computations.

6. *Ionization Potentials*

A final word must be said on ionization potentials, a molecular quantity closely related to interstellar chemistry. The experimental values of the ionization potentials (IP) for our molecules are in the range 9.5–11 eV, whereas the computed values are in the range 8–10 eV, so a mean displacement of ≈ 1 eV is present in our results. A correlation function

$$\text{IP}(\text{exptl.}) = 0.7367\text{IP}(\text{calc.}) - 3.4561,$$

with $n = 8$ and $r = 0.7812$, gives a mean absolute error in the prediction of ≈ 0.2 eV, which is reasonable, although less convincing than in the dipole moment or singlet excitation correlation linear functions. In any case, the present results confirm the general trend of ionization potentials found by other authors in a SCF context (Schwartz, 1977); it seems that introduction of electron cloud rearrangement upon ionization produces results ≈ 1 eV too low in comparison with experimental values. The previous linear correlation shows, however, that this situation can be easily remedied.

7. *Corollary*

One can conclude that the procedure we have presented here is the first step in the right direction, and it is satisfying to note that *ab initio* open shell SCF calculations can now be performed in a quite straightforward manner, as has long been the case for closed shell structures. It is, however, reasonable to think that in order to obtain refined results with a higher degree of accuracy, it is necessary to include geometric relaxation and, to some extent, correlation energy. Nevertheless, the SCF procedure provides sufficiently useful information about excited states in sequences of molecular structures.

Appendix. State Parameters

In Table A.I the state parameters used in our calculations are presented. The energy expression (1) may also be written as

$$E = \sum_{I\in S} \omega_I \left(\sum_{p\in I} h_p \right) + \sum_{I\in S} \sum_{J\in S} \left\{ \alpha_{IJ} \left(\sum_{p\in I} \sum_{q\in J} J_{pq} \right) - \beta_{IJ} \left(\sum_{p\in I} \sum_{q\in J} K_{pq} \right) \right\}, \qquad \text{(A.1)}$$

TABLE A.I

STATE PARAMETERS[a]

Closed shell $S = \{C\}$

	ω	α: C	β: C
C	2	2	1

Doublet $S = \{C, O\}$

		α		β	
	ω	C	O	C	O
C	2	2	1	1	0.5
O	1		0		0

Triplet $S = \{C, O\}$

		α		β	
	ω	C	O	C	O
C	2	2	1	1	0.5
O	1		0.5		0.5

Singlet $S = \{C, O_1, O_2\}$

		α			β		
	ω	C	O_1	O_2	C	O_1	O_2
C	2	2	1	1	1	0.5	0.5
O_1	1		0	0.5		0	−0.5
O_2	1			0			0

[a] S, set of shells; C, closed shell orbital indices; O, open shell orbital indices.

where S is the *set of shells* appearing in the state, $\{\omega_I\}$ are the *occupation numbers* for the MOs belonging to shell I, and $\{\alpha_{IJ}\}$ and $\{\beta_{IJ}\}$ are the *Coulomb and exchange parameters* between shell I and J. From Eq. (A.1) *Fock operators* may be written in *shell form* through

$$F_I = \tfrac{1}{2}\omega_I h + \sum_{J\in S} (\alpha_{IJ}J_J - \beta_{IJ}K_J), \tag{A.2}$$

where F_I stands for a Fock operator used in the same form for all MOs belonging to shell I. At the same time there are defined *Coulomb and exchange shell operators*

$$J_J = \sum_{p\in J} J_p, \tag{A.3}$$

$$K_J = \sum_{p\in J} K_p. \tag{A.4}$$

As the most usual case occurs when energy takes the form (A.1), preferably programs should take this simplification into account in order to obtain optimal computation times.

REFERENCES

Adams, W. S. (1948). *Astrophys. J.* **109**, 354.

Almlöf, J. (1972). USIP Report 72-09. University of Stockholm.

Bernecker, R. R., and Long, F. A. (1961). *J. Am. Chem. Soc.* **65**, 1565.

Bethe, M. A., and Salpeter, E. E. (1957). "Quantum Mechanics of One and Two Electron Atoms." Academic Press, New York.

Birss, F. W., and Fraga, S. (1964). *J. Chem. Phys.* **40**, 3203, 3207, 3212.

Brundle, P. E. (1967). *Chem. Commun.* p. 314.

Buhl, D. (1973). *Sky & Telescope* **45**, 2.

Buhl, D. (1974). *Origins Life* **5**, 29.

Buhl, D., and Ponnamperuma, C. (1971). *Space Life Sci.* **3**, 157.

Buhl, S., and Snyder, L. E. (1971). *Technol. Rev.* **73**, 3.

Caballol, R., Gallifa, R., Riera, J. M., and Carbó, R. (1974). *Int. J. Quantum Chem.* **8**, 373.

Caballol, R., Carbó, R., Gallifa, R., Hernández, J. A., Martín, J. A., and Riera, J. N. (1976). *Origins Life* **7**, 163.

Cameron, A. G. W., ed. (1973). "Cosmochemistry," Astrophys. Space Sci. Libr., Vol. 40. Reidel Publ., Dordrecht, Netherlands.

Carbó, R., Riera, J. M. (1978). *Lect. Notes Chem.* **5**.

Carbó, R., Gallifa, R., and Riera, J. M. (1975). *Chem. Phys. Lett.* **30**, 43.

Carbó, R., Gropen, O., and Huzinaga, S. (1977a). "Input Description of the MOMPOS Program System." Department of Chemistry, University of Alberta, Edmonton, Canada.

Carbó, R., Hernández, J. A., and Sanz, F. (1977b). *Chem. Phys. Lett.* **47**, 581.

Chuitjian, A. (1974). *J. Chem. Phys.* **61**, 4279.

Dalgarno, A., and Black, J. H. (1976). *Rep. Prog. Phys.* **39**, 573.

Davidson, E. (1976). *J. Chem. Phys.* **64**, 4699.

Davidson, E. (1977). *Theor. Chim. Acta* **44**, 405.

Demoulin, D. (1976). *Chem. Phys.* **17**, 471.

Drake, J. F. (1974). *Proc. R. Soc. London, Ser. A* **340**, 457.

Dunning, T. H., and Hay, P. J. (1977). *Mod. Theor. Chem.* **3**, 1.

Dyadyuscha, G. G., and Kuprievich, V. A. (1965). *Theor. Exp. Chem. USSR* **7**, 729.

Freeman, D. E., and Klemperer, W. (1966). *J. Chem. Phys.* **45**, 52.

Gammon, R. H. (1978). *Chem. Eng. News,* 21.

Goddard, W. A., Dunning, T. H., and Hunt, W. J. (1969). *Chem. Phys. Lett.* **4**, 231.

Gordon, M. A., and Snyder, L. E., eds. (1973). "Molecules in the Galactic Environment." Wiley (Interscience), New York.

Green, S., and Herbst, E. (1979). *Astrophys. J.* **229**, 121.

Ha, T., and Keller, L. (1975). *J. Mol. Struct.* **27**, 225.

Harrison, A. S., Cederholm, B. J., and Terrwillinger, M. A. (1959). *J. Chem. Phys.* **30**, 355.

Herbst, E., and Klemperer, W. (1976). *Phys. Today,* 32.

Herzberg, G. (1966). "Molecular Spectra and Molecular Structure," Vol III, p. 624. Van Nostrand-Reinhold, Princeton, New Jersey.

Hirao, K. (1974). *J. Chem. Phys.* **60**, 3215.

Hirao, K., and Nakatsuji, H. (1973). *J. Chem. Phys.* **59**, 1457.

Hoyle, F., Olavsen, A. H., and Wickramasinghe, R. (1978). *Nature* (*London*) **271**, 229.

Huzinaga, S. (1960). *Phys. Rev.* **120**, 866.

Huzinaga, S. (1961). *Phys. Rev.* **122**, 131.

Huzinaga, S. (1969). *J. Chem. Phys.* **51**, 3971.

Johnson, D. R., and Lovas, F. J. (1972). *Chem. Phys. Lett.* **15**, 65.

Johnson, D. R., and Lovas, F. J. (1972). *Chem. Phys. Lett.* **15**, 65.

Khare, B. N., and Sagan, C. (1973). *In* "Molecules in the Galactic environment" (M. A. Gordon and L. E. Snyder, eds.), p. 399. Wiley (Interscience), New York.

Kim, H., Keller, R., and Gwinn, W. D. (1962). *J. Chem. Phys.* **37**, 2748.

Knoop, F. W. E., Brongersma, H. H., and Osterhoff, L. J. (1972). *Chem. Phys. Lett.* **13**, 20.

Löwdin, P. O. (1955). *Phys. Rev.* **97**, 1474.

McClellan, A. L. (1963). "Tables of Experimental Dipole Moments." Freeman, San Francisco, California.

McNally, D. (1972). *Adv. At. Mol. Phys.* **8**, 1.

Margulis, L., ed. (1973). "Proceedings of the IVth Conference on Origins of Life," pp. 219–248. Springer-Verlag, Berlin and New York.

Mezger, P. G. (1975). *In* "Structure and Evolution of Galaxies" (G. Setti, ed.), p. 143. Reidel Publ. Co., Dordrecht, Netherlands.

Mitchell, G. F., Ginsburg, J. L., and Kuntz, P. J. (1978). *Astrophys. J., Suppl. Ser.* **38**, 1612.

Nicholls, R. W. (1977). *Am. Rev. Astron. Astrophys.* **15**, 197.

Okabe, H. (1978). "Photochemistry of Small Molecules." Wiley Interscience, New York.

Özkan, I., Chu, S. Y., and Goodman, J. (1975). *J. Chem. Phys.* **63**, 3195.

Parr, R. H. (1963). "Quantum Theory of Molecular Electronic Structure." Benjamin, New York.

Peters, D. (1972). *J. Chem. Phys.* **57**, 4351.

Pinkau, K. (1974). "The Interstellar Medium," NATO Adv. Study Inst. Ser., Vol. C6. Reidel Publ., Dordrecht, Netherlands.

Ponnamperuma, C. (1978). *In* "Proceedings of the R. A. Welch Foundation Conferences on Chemical Research. XXI. Cosmochemistry" (W. O. Milligan, ed.), p. 137.

Richards, W. G., Walker, T. E. H., and Hinkley, R. K. (1971). "A Bibliography of 'ab initio' Molecular Wave Functions." Oxford Univ. Press (Clarendon), London and New York.

Richards, W. G., Walker, T. E. H., Farnell, L., and Scott, P. R. (1974). "Bibliography of 'ab initio' Molecular Wave Functions. Supplement for 1970–1973." Oxford Univ. Press (Clarendon), London.

Richards, W. G., Scott, P. R., Colbourn, E. A., and Marchington, A. F. (1978). "Bibliography of 'ab initio' Molecular Wave Functions. Supplement for 1974–1977." Oxford Univ. Press (Clarendon), London and New York.

Roothaan, C. C. J. (1951). *Rev. Mod. Phys.* **23**, 69.

Roothaan, C. C. J. (1960). *Rev. Mod. Phys.* **32**, 179.

Sagan, C., and Khare, B. N. (1979). *Nature* (*London*) **277**, 102.

Sandorfy, C. (1973). *Theor. Chim. Acta* **29**, 1.

Sandorfy, C., Ausloos, P. J., and Robin, M. B., eds. (1974). "Chemical Spectroscopy and Photochemistry in the Vacuum UV," NATO Adv. Study Inst. Ser., Vol. 8. Reidel Publ., Dordrecht, Netherlands.

Schwartz, M. E. (1977). *Mod. Theor. Chem.* **4**, 357.

Shimizu, M. (1973). *Prog. Theor. Phys.* **49**, 153.

Snyder, L. S. (1974). *J. Chem. Phys.* **61**, 747.

Snyder, L. S., and Basch, H. (1972). "Molecular Wavefunctions and Properties." Wiley, New York.

Solomon, P. M. (1973). *Phys. Today*, 32.

Tam, W. C., and Brion, C. E. (1974). *J. Electron Spectrosc. Relat. Phenom.* **3**, 263.

Van Woerden, H., ed. (1977). "Topics in Interstellar Matter," Astrophys. Space Sci. Libr., Vol. 70. Reidel Publ. Co., Dordrecht, Netherlands.

Vedeneyev, V. I., Gurvich, L. V., Kondrat, V. N., Medvedev, V. A., and Frankevich, Ve. L. (1966). "Bond Energies, Ionization Potentials and Electron Affinities." Arnold, London.

Wadt, W. R., and Goddard, W. A. (1976). *Chem. Phys.* **18**, 1.

Watson, W. D. (1976). *Rev. Mod. Phys.* **48**, 513.

Watson, W. D. (1977). *Acc. Chem. Res.* **10**, 221.

Watson, W. D. (1978). *Annu. Rev. Astron. Astrophys.* **16**, 585.

Weinreb, S., Barrett, A. H., Meeks, M. L., and Henry, J. C. (1963). *Nature* (*London*) **200**, 829.

Werner, M. W. (1971). *Icarus* **15**, 352.

Wilkinson, J. H., and Reinsch, C. (1971). "Linear Algebra," Vol. 186, p. 303. Springer-Verlag, Berlin and New York.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 12

The Spin-Projected Extended Hartree-Fock Method[1]

ISTVÁN MAYER
Central Research Institute for Chemistry
of the Hungarian Academy of Sciences
Budapest, Hungary

---

[1] Dedicated to Professor Per-Olov Löwdin on the occasion of the twenty-fifth anniversary of the publication of his classical series of papers (Löwdin, 1955a,b,c) initiating the development of the field.

ISBN 0-12-034812-8

I. Introduction

A. The One-Electron Approximation and the DODS Idea

Almost all the methods used in quantum chemistry are based on the well-known variation theorem (Eckart, 1930). Choosing the type of the trial wave function used in the variation procedure, one must take into account a great number of different factors, e.g. the given class of the wave functions must be adequate to describe the most important physical features of the system, contain as many free variational parameters as feasible from the computational point of view, etc. It is also rather desirable to build up the wave function of simpler constituting elements which may be visualized to some extent, thus facilitating the interpretation of the results obtained and the elaboration of a realistic physical picture for the system studied. The one-particle approximations—the Hartree–Fock method and its generalizations—are of upmost importance in this respect. This is one of the reasons why just the Hartree–Fock (HF) or self-consistent field (SCF) method is the basic one in quantum chemistry: the many-electron wave function is approximated as a single Slater determinant, i.e., as an antisymmetrized product of functions each of which depends on coordinates of only one electron. These one-electron functions (orbitals) are then optimized in a variational way.

The HF method in its most usual form applies the restriction for the one-electron functions that (for closed-shell systems) two electrons with opposite spins are put on the same spatial orbital. (Besides this double-occupancy restriction, often restrictions are introduced also for the spatial symmetry properties of the orbitals.) For this reason the conventional HF method is often called "restricted" HF (RHF), especially if one wants to make a clear distinction between the different related approaches.

The use of the double-occupancy restriction has the advantage that the determinantal many-electron wave function becomes automatically an eigenfunction of the spin operators $\hat{S}_z$ and $\hat{S}^2$. The RHF method gives results which are often satisfactory in many respects, but the correlation energy (the error in the energy calculated by the RHF method) is quite significant, being at least comparable with the energy changes and differences important from the chemical point of view. Another well-known drawback of the RHF method is that it is inapplicable to the description (even qualitative) of the dissociation processes: at large interatomic distances (if one can get convergence in the SCF procedure at all) the RHF wave function does not correspond to the neutral odd-electron atoms or groups, but contains also large ionic VB terms as a necessary consequence of the double-occupancy restriction. It is well known that there is a great variety of other problems too, for which the RHF approximation is not adequate; it is sufficient to mention here the magnetic (or more generally, spin-dependent) phenomena. There is consequently a need for methods which go beyond the RHF approximation and take into account, at least partly, the correlation effects.

The method of constructing the HF wave functions is the reason why it is also often called the one-electron or independent-particle model: one may consider that each electron moves in the field of the nuclei and the average field due to the other electrons, with the intimate correlation between the motion of different electrons not taken into account. Strictly speaking, this is fully valid only for the simplest Hartree method, in which the Pauli principle is taken into account only partly; if one uses the properly antisymmetrized HF wave function, then the probability of finding two electrons with the same spin at the same point of the space is always zero ("Fermi-hole"). Accordingly, in the HF method only the motion of the electrons with opposite spins may be considered as fully uncorrelated. Nevertheless, owing to the specific importance of the RHF method, it is usual, following Löwdin's (1959) proposition, to define the correlation energy as the deviation of the RHF energy from the exact nonrelativistic value, and to call correlation effects only those which are not included by the antisymmetrization into the RHF wave function.

Some of the methods which aim at taking into account a part of the correlation energy can be considered as generalizations of the usual HF approach since they remain within the framework of the one-electron approximation. Their important advantage is in conserving the conceptually simple, visualizable picture of assigning at least in some sense one orbital to each electron. (We shall consider in this article only methods of this type.) These methods are based on the idea of using "different orbitals for different spins" (DODS), which was probably first explored in the papers of Slater (1930, 1951) concerning problems of solid-state physics and worked out for molecular systems mainly by Löwdin and co-workers.

The basis of the DODS philosophy consists in the assumption that important correlation effects can be described if one does not "force" pairs of electrons having opposite spins to occupy identical spatial orbitals. Thus these electrons may "avoid" each other, their average distance increases, and therefore it becomes possible to describe not only the Fermi-hole but, at least to some extent, also the "Coulomb" one. Thus a part of correlation between the electrons with opposite spins can be taken into account in this conceptually simple way. The flexibility of the wave function increases as the number of optimized spatial orbitals is doubled with respect to the RHF case.

B. DODS Methods

The first step in order to realize the DODS idea is to remove the double-occupancy restriction of the RHF method and yet conserve the single Slater-determinant wave function. Thus one gets the (spin-) unrestricted HF (UHF) or simple DODS method[2] (e.g., Coulson and Fischer, 1949; Berthier, 1954a,b; Pople and Nesbet, 1954; Pratt, 1956; Löwdin, 1958a; Nesbet, 1961; Amos and Hall, 1961; Amos and Snyder, 1964; Snyder and Amos, 1965; Koutecký, 1967; Claxton and Smith, 1971; Mayer, 1971a, 1973a, 1974b; Fukutome, 1972, 1973; Ladik and Otto, 1975; Tiňo and Klimo, 1967a,b; Tiňo *et al.*, 1978; Kertész *et al.*, 1979; Lunell, 1979). As the orbitals with different spins now can be varied independently of each other, the application of the UHF method means that a trial wave function of increased flexibility is used in the variation procedure. Accordingly, the UHF method often gives better (i.e., lower) energies than the RHF one, thus really permitting account to be taken of a part of the correlation energy. The "dissociation catastrophe" characteristic of the RHF method also does not appear in the UHF case. However, in many cases, especially for small and intermediate internuclear distances (i.e., just around the equilibrium geometry), one cannot find UHF solutions differing from the RHF one (e.g., Coulson and Fischer, 1949; Fukutome, 1972; Mayer, 1978). In these cases the decrease of the interelectronic repulsion which can be obtained within the framework of the UHF method cannot compensate for the slight increase in the average distance between the electrons and nuclei and/or the increase in the kinetic energy of the electrons. In these cases the use of the UHF formalism is useless, of course.

There are problems with the UHF method even in the most favorable cases when it gives lower energies than the RHF one. Though lower energy means better energy according to the variation principle, it does not follow that the UHF wave function is also automatically better than the RHF one. Often the opposite is true. (A characteristic numerical example will be shown in Section IV,A). In fact, using the single DODS determinant UHF wave function, one encounters the serious difficulty that it is not an eigenfunction of the total spin operator $\hat{S}^2$, i.e., it does not correspond to any definite spin multiplicity, but is a mixture of components having different multiplicities (e.g., singlet, triplet, etc.).

In order to solve this problem Löwdin (1955c) proposed to use the "spin projection operator" bearing his name, which permits one to obtain from a wave function of such mixed character a new wave function having the desired multiplicity. As will be seen, the spin projection operator "annihilates" (multiplies by zero) all the components with unwanted multiplicities, leaving unchanged that component which corresponds to the

---

[2] The UHF equations were first derived (Berthier, 1954a,b; Pople and Nesbet, 1954) for treating open-shell systems (radicals) and not in order to deal with the correlation problem.

required multiplicity, i.e., it performs the projection into the subspace of the appropriate spin eigenfunctions.

If one solves the UHF equations and subjects to spin projection the UHF wave function obtained, then usually one gets an improvement with respect to both the energy and the wave function. This approach may be called "UHF with subsequent spin projection" (UHF+SP) procedure. In this case, however, there appears the contradiction that one optimizes variationally the single determinant wave function which one has before the projection and not the projected one which is really applied. (In Section IV,A we shall demonstrate an example showing that this inconsistently variational approach may lead to serious difficulties.)

Owing to these problems, it is desirable to apply the spin-projected wave function directly as trial function in the variation procedure, and to determine the orbitals in such a way as to obtain a minimum energy for the wave function which one has after the projection. This is the essence of the spin-projected "extended Hartree-Fock" (EHF) method—the central subject of this article—and explains why the EHF method is often called "projected HF" (PHF) also.[3]

The "alternant molecular orbital" (AMO) method represents a simplified version of the EHF one, since the orbitals are not varied quite freely but are constructed according to some recipe. It will be discussed in some detail in Section II,C. Some other methods of DODS type will be mentioned in Section II,K.

We may note here that the variational wave functions used in some well-known classical calculations may also be considered as specific versions of the EHF wave function. This is the case, e.g., when the wave function of the He atom is built up by using two 1s orbitals with different exponents (Hylleraas, 1929; Eckart, 1930). Specific attention must be paid to the fact that Weinbaum's (1933) wave function for the H_2 molecule, containing the ionic terms with an optimized weight (the importance of which, for the development of our understanding of the chemical bond, may hardly be overemphasized), can also be presented in the form of a spin-projected determinant. This means that for the case of two electrons and two basis orbitals the EHF wave function coincides with the full CI one (cf. also Davidson and Jones, 1962). This is, of course, not valid for larger systems.

---

[3] There is a significant inconsistency in the nomenclature. Thus the EHF method is also called "spin-extended," simply "generalized," "GF" (originating from a group-theoretical notation), etc. At the same time the notation EHF is sometimes used for quite different approaches (e.g., Löwdin, 1955a; Das and Wahl, 1966; Jug, 1973). In the papers of the present author the nomenclature proposed by Löwdin (1958a) was consistently used; it will be followed, therefore, also in the present article.

C. The "Symmetry Dilemma" and the Projection Method

It is worthwhile to look at the EHF problem from a somewhat more general point of view. As Löwdin (1955c, 1962c, 1963, 1964, 1966, 1969, 1976) emphasized, the exact solutions of the Schrödinger equation are automatically eigenfunctions of the operators corresponding to the constants of motion and thus commuting with the Hamiltonian.[4] However, if one uses approximate wave functions, no such properties follow from the variation principle. On the contrary, requiring the fulfillment of some symmetry property represents an additional restriction (constraint) from the point of view of the variational problem, thus decreases the flexibility of the trial function, and, in general, leads to an increase of the energy (also see, e.g., Musher, 1970). Löwdin introduced the term "symmetry dilemma" to describe such situations. In fact, remaining within the given class of functions, one must choose between obtaining a lower energy or having a wave function with correct symmetry properties (e.g., an eigenfunction of the operator $\hat{S}^2$). In order to solve this symmetry dilemma Löwdin proposed the "component analysis," which may be performed by means of projection operators. The energy of a wave function constructed in an unrestricted manner with respect to the given constant of motion represents the weighted average of the energies of the pure components. As these energy values are usually different, one can find a component with an energy lower than that of the starting wave function. Thus the use of the projection operator formalism permits one not only to ensure the correct symmetry but also to get a further improvement in the energy.[5] There is also another important advantage of this approach: if the starting wave function is a Slater determinant, then one conserves the immediate connection with the independent-particle model. (Not only the well-visualizable orbital picture is conserved, but also properties such as Koopman's theorem may be generalized—cf. Section III,F). In this respect one has to consider the projected determinant not as a linear combination of a large number of different determinants (in the form of which it can be explicitly expressed) but as a *single conceptual entity*. The situation is quite similar to the case of a single Slater determinant which is also nothing but the antisymmetric projection of a Hartree product. Thus the projected HF method in the most general sense (the spin-projected EHF method is a specific case of it) represents such a kind of generalization of the independent-particle model in which the trial wave function used in

---

[4] In the case of degenerate energy levels one may also choose the wave functions in such a way as to fulfill this requirement.

[5] Moreover, if the wave function possesses automatically a correct symmetry without imposing any restriction on it, this does not mean, in general, that there is no "broken symmetry" wave function which will have a lower energy after the projection is performed.

the variation procedure is a Slater determinant projected according to the appropriate symmetry. Based on a rather limited number of data available at that time, Löwdin (1966) expected that in this manner one could take into account about 95% of the correlation energy. (Unfortunately, this expectation was fulfilled in the actual calculations performed at the *ab initio* level only for the case of two-electron systems—see Sections II,I and IV,A.)

II. Some Theoretical Aspects of the EHF Method: A Survey

After the pioneering work of Pratt (1953), the EHF method was proposed by Löwdin (1955c) and since that time a very large number of scientists have worked on the problems connected with different aspects of the method or its simplified and generalized versions. They used a great variety of quite different approaches and formalisms. It is completely impossible to give a detailed account of all these works; in this survey we can discuss only those main steps in the development of the field which were most important from the point of view of the approach used by the present author and summarized in Section III. Some of the results (mainly those which will be explicitly utilized in the next sections) are discussed in some detail; many others can only be briefly summarized or mentioned. These differences in presentation, however, do not in any way correspond to similar differences in the importance of the contributions to the field due to the different authors.

A. Pratt's Spin Operator

Pratt (1953) was the first to construct an operator which, when applied to a product of spin functions of $2N$ electrons, containing N α-s and N β-s (i.e., corresponding to $S_z = 0$), produces a singlet eigenfunction of the operator $\hat{S}^2$. Though the spin projection formalism introduced by Löwdin (1955c) shortly thereafter is much more elegant and general, it is instructive to discuss Pratt's work briefly. Its importance consists not only in its pioneering character, but also the physical picture [the spin-coupling scheme (Pauncz, 1967)] which can be connected with the EHF wave functions appears here in the most immediate way. The essence of Pratt's approach is as follows.[6]

The electrons are divided into two groups (subsystems) A and B, each of which contains N electrons and is in the state of maximum multiplicity

---

[6] Originally Pratt considered the wave function together with its spatial part and applied specific orthogonality conditions for the orbitals; from our point of view it is sufficient to consider the spin functions alone.

($S = S_A = S_B$); these are then coupled together according to the vector coupling scheme to form a resulting singlet:

$$\Psi(S = 0) = \sum_{M_S=-S}^{S} (-1)^{S-M_s} \varphi_A(S, M_s)\varphi_B(S, -M_s). \tag{1}$$

Here the factors $(-1)^{S-M_s}$ are the Clebsch–Gordan coefficients; $\varphi_A(S, M_s)$ and $\varphi_B(S, -M_s)$ are normalized spin functions of the subsystems A and B, respectively, corresponding to the different z-projections $S_z = \pm M_s$ which are possible for the case of the maximum multiplicity $S = N/2$. These spin functions may easily be constructed from the functions $\varphi_A(S = \frac{1}{2}N, M_s = \frac{1}{2}N) = \alpha(1_A)\alpha(2_A) \cdots \alpha(N_A)$ and $\varphi_B(S = \frac{1}{2}N, M_s = -\frac{1}{2}N) = \beta(1_B)\beta(2_B) \cdots \beta(N_B)$ by the repeated use of the "step operators"

$$\hat{S}_{A-} = \sum_{i=1}^{N} \hat{S}_{i_A-} \quad \text{and} \quad \hat{S}_{B+} = \sum_{i=1}^{N} \hat{S}_{i_B+}$$

corresponding to the two subsystems and by introducing appropriate normalization coefficients. Accordingly

$$\Psi(S = 0) = \hat{O}_1[\varphi_A(S = \tfrac{1}{2}N, M_s = \tfrac{1}{2}N)\varphi_B(S = \tfrac{1}{2}N, M_s = -\tfrac{1}{2}N)], \tag{2}$$

where

$$\hat{O}_1 = \sum_{M=0}^{N} (-1)^M[(N - M)!/(N!M!(N + 1)^{1/2})](\hat{S}_{A-}\hat{S}_{B+})^M. \tag{3}$$

As a result of applying this operator, one obtains a sum of products of spin functions (or, when the spatial part and the antisymmetry of the wave function are also taken into account, a sum of Slater determinants). According to the previous discussion, applying the operator $\hat{O}_1$ to a product of spin functions containing N α-s and N β-s, one has to consider the electrons with spins α as belonging to subsystem A, and the electrons with spins β as belonging to subsystem B. In order to get this division automatically, Pratt (1953) modifies the operator $\hat{O}_1$ and obtains the more general operator[7]

$$\hat{O} = (N+1)^{-1/2} \sum_i \left[\left(1 + \sum_{M=1}^{N} \frac{(-1)^M (N-M)!}{N!M!} \{(\hat{S}_{A_i-}\hat{S}_{B_i+})^M + (\hat{S}_{A_i+}\hat{S}_{B_i-})^M\} \right) \right.$$
$$\left. \times \prod_{j_{A_i}, k_{B_i}}^{N} (\hat{S}_{j_{A_i}+}\hat{S}_{j_{A_i}-}\hat{S}_{k_{B_i}-}\hat{S}_{k_{B_i}+} + \hat{S}_{j_{A_i}-}\hat{S}_{j_{A_i}+}\hat{S}_{k_{B_i}+}\hat{S}_{k_{B_i}-}) \right] \cdot \tag{4}$$

Here the summation according to i goes over all the possible divisions of $2N$ electrons into two groups A and B, each containing N electrons. The

---

[7] In Eq. (4) we added a pair of parentheses, which are absent in the original article (Pratt, 1953) due perhaps to a misprint.

essence of this modification is the following: the operators after the symbol Π eliminate all wave functions in which, for the given division i, the two groups consist not of only α-s or β-s. Thus only that term of the sum over i will survive which corresponds to the required division of the electrons (except, possibly, the interchange of the groups A and B, which needs not to be considered as a separate case). Pratt (1953) proved that the operator $\hat{O}$ has the fundamental properties: (a) it is Hermitian; (b) it commutes with the operator $\hat{S}^2$; (c) it commutes with the spin-free Hamiltonian; (d) its square is proportional to itself: $\hat{O}^2 = \text{const} \cdot \hat{O}$; and (e) it commutes with the antisymmetrizer.

We shall not discuss those aspects of Pratt's work (1953) which are not essential from our point of view; we only call attention to the fact that Pratt, though he did not explicitly realize that his operator $\hat{O}$ is (up to a constant) a projection operator, already obtained the most important properties of the spin projection operator.[8]

B. Löwdin's Spin Projection Operator

In his classical series of papers Löwdin (1955a,b,c), in the course of the detailed analysis of density matrices, natural spin orbitals, and CI expansions, derived the formulas for different matrix elements taken between Slater determinants built up of not mutually orthogonal orbitals. (These formulas were used in the derivations discussed in Section III). Löwdin also performed a detailed analysis of the conventional HF method and of the properties of wave functions consisting of a single Slater determinant. He pointed out that all the essential information concerning the single determinant wave function is contained in the first-order density matrix, and that the latter (similarly to the wave function) is invariant under the linear transformations of the orbitals. In other words, the single determinant wave function is uniquely determined by the linear space spanned by the occupied orbitals. The single determinant wave function in the general (DODS) case is not an eigenfunction of the operator $\hat{S}^2$ but is a mixture of terms with different multiplicities. In the third part of the series Löwdin (1955c) introduced his spin projection operator as a specific case of the projection operators which can be assigned to the operators commuting with the Hamiltonian:

$$\mathcal{O}^S = \prod_{l \neq S} \frac{\hat{S}^2 - l(l+1)}{S(S+1) - l(l+1)}. \tag{5}$$

This operator permits selection from a wave function of that component which belongs to the eigenvalue $S(S+1)$ of the operator $\hat{S}^2$ (having

---

[8] Löwdin's spin projection operator (Section II,B) is, however, strictly idempotent.

the spin multiplicity $2S + 1$). In fact, each term $\hat{S}^2 - l(l + 1)$ of this operator product annihilates that component of the wave function which belongs to the eigenvalue $l(l + 1)$ of the operator $\hat{S}^2$ whereas the other components are multiplied by a constant, e.g., the component corresponding to the eigenvalue $k(k + 1)$ is multiplied by $k(k + 1) - l(l + 1)$. For the component with the desired eigenvalue $S(S + 1)$ this is just compensated by the introduction of the appropriate denominators. In the operator (5) the values of l run from 0 (if the number of electrons is even) or $\frac{1}{2}$ (if the number of electrons is odd) to the maximum value $N/2$, which is possible for N electrons, except the desired value S. Thus, after applying this operator, only this component of the wave function survives.

By considering the case $S_z = 0$ for an even number of electrons ($N = 2n$) and presenting the operator $\hat{S}^2$ in the well-known form $\hat{S}^2 = -\frac{1}{4}N(N - 4) + \sum_{i<j}\hat{P}^\sigma_{ij}$ (here $\hat{P}^\sigma_{ij}$ is the operator interchanging the spin coordinates of the ith and jth electrons), Löwdin (1955c) demonstrated that the spin-projected single-determinant wave function may be given as a sum of Slater determinants:

$$\hat{O}^S\Phi = \hat{O}^S T_0 = \sum_{k=0}^{n} c_k T_k. \tag{6}$$

Here $T_0 \equiv \Phi$ is the original Slater determinant, while T_k is the sum of all the $\binom{n}{k}^2$ Slater determinants which can be obtained from T_0 by replacing k spins α by β and, simultaneously, k spins β by α. In order to prove the expansion (6), Löwdin utilized the fact that the spin projection operator is a polynomial of the operator $\hat{S}^2$ and that $\sum_{i<j}\hat{P}^\sigma_{ij}$ and, therefore, $\hat{S}^2$ (and $\hat{O}^S$) commute with the antisymmetrizer. By counting the number of interchanges corresponding to the electrons with identical and different spins, respectively, it can be found that $\hat{S}^2 T_k$ is a linear combination of T_{k-1}, T_k, and T_{k+1}:

$$\hat{S}^2 T_k = (n - k + 1)^2 T_{k-1} + [n(2k + 1) - 2k^2]T_k + (k + 1)^2 T_{k+1} \tag{7}$$

(by using the definition $T_{-1} = T_{n+1} = 0$ this expression remains valid also for $k = 0$ or $k = n$). Because $\hat{O}^S$ is a polynomial of the order $n - 1$ of $\hat{S}^2$, all the T_k-s ($0 \leq k \leq n$) must appear at the right-hand side of Eq. (6)

The spin-projected wave function $\hat{O}^S T_0$ is an eigenfunction of $\hat{S}^2$:

$$\hat{S}^2\hat{O}^S T_0 = S(S + 1)\hat{O}^S T_0. \tag{8}$$

Substituting Eqs. (6) and (7) into Eq. (8) Löwdin (1955c) obtained the recursion formula

$$(n - k)^2 c_{k+1} + [n(2k + 1) - 2k^2 - S(S + 1)]c_k + k^2 c_{k-1} = 0 \tag{9}$$

for the coefficients c_k (these coefficients are often called "Sanibel coefficients").

Already in this basic paper, Löwdin (1955c) gave also the explicit expressions of the Sanibel coefficients for the singlet ($S = 0$) case and for the $S_z = 0$ component of the state of the highest multiplicity $S = n$:

$$c_k^{(0)} = c_0^{(0)}(-1)^k \binom{n}{k}^{-1} \tag{10}$$

$$c_k^{(n)} = c_0^{(n)} \qquad \text{(does not depend on } k\text{)}, \tag{11}$$

respectively.

The value of the coefficient c_0 is usually unimportant, as in general the spin-projected determinant wave function is not normalized anyway. Nevertheless, Löwdin (1955c) gave its value for the singlet case [$c_0^{(0)} = (n+1)^{-1}$] for which it can be easily obtained by considering a determinant built up of doubly filled orbitals, which is automatically a pure singlet.

Löwdin (1955c) gave also the expectation value of the energy for the specific case of the spin-projected Slater determinant built up of strictly orthonormalized spatial orbitals, and discussed how the results can be used for the interpretation of Hund's rule. At the same time he pointed out that it is necessary to consider also the general nonorthogonal case, which, however, became possible only after the so-called "pairing theorem" (Section II,D) was proven.

The expansion of type (6) occurs also in the more general cases, and the corresponding Sanibel coefficients[9] were the subject of extensive investigations (e.g., Percus and Rotenberg, 1962; Sasaki and Ohno, 1963; Smith, 1964; Manne, 1966; Pauncz, 1967; Kouba and Öhrn, 1969; Smith and Harris, 1969; van Leuven, 1969; Mano, 1970; Sarma, 1975). As a result, one can determine the values of Sanibel coefficients in all cases by using closed formulas, finite sums or recursion relations. We shall not discuss here these derivations,[10] but only mention that in the most important "principal case" $S = M$ (i.e., when the eigenvalue M of the operator $\hat{S}_z$ is the largest value compatible with the given eigenvalue $S(S+1)$ of the operator $\hat{S}^2$) the Sanibel coefficients may be most easily obtained (e.g., Pauncz, 1967) by using the following interesting presentation (Löwdin, 1958b) of the spin projection operator

$$\hat{O}_S^S = (2S+1)! \sum_{p=0}^{N/2-S} (-1)^p \frac{\hat{S}_-^p \hat{S}_+^p}{p!(2S+p+1)!} \tag{12}$$

---

[9] In order to distinguish between the different cases, more detailed notations are also often used for the Sanibel coefficients: e.g., $C_k(S, S_z, \mu)$, where μ stands for the number of electrons with spin α and the meaning of the other symbols is obvious.

[10] A detailed discussion of this problem and of many other aspects of the spin projection operator formalism may be found in Pauncz's book (1967).

(the subscript S indicates that we are dealing with the principal case). Because (to the extent of this author's knowledge) the derivation of Eq. (12) cannot be found in the literature, it seems worthwhile to sketch it very briefly (Mayer, 1977).

The spin projection operator commutes with the antisymmetrizer and does not act on the spatial functions. We may therefore restrict ourselves to the consideration of spin functions. Let us start from a product of spin functions containing μ α-s and ν β-s ($\mu + \nu = N, \mu - \nu = 2M = 2S$):

$$\vartheta_1(\mu,\nu) = \alpha(1)\,\alpha(2)\cdots\alpha(\mu)\,\beta(\mu+1)\,\beta(\mu+2)\cdots\beta(N).$$

$\vartheta_1(\mu, \nu)$ is an eigenfunction of operator $\hat{S}_z$ with an eigenvalue M, it contains, therefore, only components with $S \geq M$. Accordingly, in order to obtain the projection $S = M$, the spin projection operator must annihilate only the components which are of higher multiplicity:

$$\hat{O}_S^S = \prod_{l=S+1}^{N/2} \frac{\hat{S}^2 - l(l+1)}{S(S+1) - l(l+1)} = \prod_{l=S+1}^{N/2} \frac{\hat{S}_-\hat{S}_+ + \hat{S}_z^2 + \hat{S}_z - l(l+1)}{S(S+1) - l(l+1)}. \tag{13}$$

The second equality in (13) follows from the identity $\hat{S}^2 = \hat{S}_-\hat{S}_+ + \hat{S}_z^2 + \hat{S}_z$. As the operators $\hat{S}_z$ and $\hat{S}^2$ commute, the application of the annihilators in Eq. (13) produces again spin functions ϑ which are eigenfunctions of $\hat{S}_z$ with the same eigenvalue $M = S$. Thus on the right-hand side of Eq. (13) we may replace $\hat{S}_z^2 + \hat{S}_z - l(l+1)$ by $S(S+1) - l(l+1)$, and obtain after simple algebraic manipulations

$$\hat{O}_S^S = \prod_{q=1}^{N/2-S} \{1 - \hat{S}_-\hat{S}_+/[2qS + q(q+1)]\}. \tag{14}$$

Based on the identity $\hat{S}_+\hat{S}_- = \hat{S}_-\hat{S}_+ + 2\hat{S}_z$, it is easy to prove by induction that the following relationship holds for the case $S = M$, when $\hat{S}_z\hat{S}_+^k\vartheta = (S+k)\hat{S}_+^k\vartheta$:

$$\hat{S}_-^q\hat{S}_+^q\hat{S}_-\hat{S}_+ = \hat{S}_-^{q+1}\hat{S}_+^{q+1} + q(2S+q+1)\hat{S}_-^q\hat{S}_+^q. \tag{15}$$

Now, it is easy to see immediately that Eq. (12) is equivalent to Eq. (14) if one considers only one annihilator ($q = 1$) or the product of the first two annihilators ($q = 1, 2$). For the general case ($q = 1, 2, \ldots, N/2 - S$) this equivalency may be easily proven by induction on the basis of Eq. (15).

It is worthwhile to mention that the Sanibel coefficients in the "principal case" can be given (Pauncz, 1967) explicitly as

$$c_q = (-1)^q \frac{2S+1}{\mu+1}\binom{\mu}{q}^{-1} = (-1)^q\binom{\mu}{q}^{-1} c_0, \tag{16}$$

i.e., they are interrelated exactly in the same manner as in the singlet case.

The relationships between the above two formalisms, i.e., between Pratt's spin operator and Löwdin's spin projection operator were investigated by Berencz (1960). He showed that the operator $\hat{O}_1$ in Eq. (3) representing the basis of Pratt's approach and Löwdin's spin projection operator $\hat{O}^S$ in Eq. (5) give results coinciding up to a constant when applied to a determinant with the spin distribution $T_0 = \{\alpha\alpha . . . | \beta\beta . . . \}$. To do this, he proved the relationship $T_k = (k!)^{-2}(\hat{S}_{A-}\hat{S}_{B+})^k T_0$. Berencz (1960) gave also an independent derivation of the recursion relation (9) based on the identity $\hat{S}^2 = \hat{S}_+\hat{S}_- + \hat{S}_z^2 - \hat{S}_z$. It may be also noted in this connection that Berencz and Pauncz (1958; Berencz, 1958, 1963) have further developed Pratt's formalism, which is based on the consideration of two subsystems, and applied it in constructing the spin functions belonging to different eigenvalues of the operator $\hat{S}^2$ and in treating CI problems.

C. The AMO Method

By generalizing some ideas of Slater (1930) concerning antiferromagnetism, Löwdin (1955c) introduced the alternant molecular orbital (AMO) method as a specific version of applying the DODS idea and the spin projection method for treating the correlation problem. The AMO method was quite successful in describing the π-electron structure of conjugated molecules (Yoshizumi and Itoh, 1955; Itoh and Yoshizumi, 1955; Lefebvre *et al.*, 1960; Dearman and Lefebvre, 1961; Pauncz *et al.*, 1962a,b; de Heer, 1962a,b, 1963; Fischer, 1962; Pauncz, 1962, 1963, 1964, 1967, 1969; Coulson, 1963; Moskowitz, 1963; de Heer and Pauncz, 1963; Swalem and de Heer, 1964; Silberman *et al.*, 1968; Lindner and Lunell, 1967; Lunell and Lindner, 1968; Tamir and Pauncz, 1968; Laskowski and van Leuven, 1971; Tyutyulkov, 1975, 1978) and, at least to some extent, in treating the electronic structure of crystals (Löwdin, 1956, 1962a; Dermit, 1962; Calais, 1965, 1967, 1979; Goscinski and Calais, 1965; Calais and Sperber, 1973; Sperber and Calais, 1973; Sperber, 1973; Larson and Thornson, 1966; Dugay and Thomas, 1975; for some different applications, see, e.g., Brandas, 1967, 1968). The AMO method and the results which can be obtained by its use are described in detail in Pauncz's book (1967). Accordingly, we shall show only that the AMO method may be considered as a simplified version of the EHF method, differing from the latter in the way of performing the variational procedure for the orbitals. In the AMO method one does not consider the orbital coefficients as free variational parameters, but the DODS orbitals are obtained by combining with each other the occupied and empty (virtual) RHF (or, sometimes, Hückel or "topological") orbitals according to a specific recipe. This approach reduces the number of variational parameters which are to be optimized. By using the appropriate coupling scheme and optimizing the value(s) of the parameter(s), one may ensure that the electrons having different spins will "avoid each other," as a consequence of the fact that the orbitals filled before the projection by spins α and β, respectively, are more or less localized on different subsystems of the atoms.[11] This is especially evident for the π-electron systems of alternant hydrocarbons, for which the occupied and virtual Hückel or RHF orbitals appear by pairs (Coulson and Rushbroke, 1940; Coulson and Lounget-Higgins, 1947; Pople, 1953; Lounget-Higgins, 1966; Hall and Amos, 1969). Each pair consists of an occupied (ψ_{ib}) and a virtual (ψ_{iv}) orbital, and their LCAO coefficients corresponding to the atomic orbitals (AO-s) χ_μ belonging to the two subsystems (I and II) of atoms are either identical or differ only in the sign:

$$\psi_{ib} = \sum_{\mu \in \mathrm{I}} c_{\mu i}\chi_\mu + \sum_{\mu \in \mathrm{II}} c_{\mu i}\chi_\mu, \qquad \psi_{iv} = \sum_{\mu \in \mathrm{I}} c_{\mu i}\chi_\mu - \sum_{\mu \in \mathrm{II}} c_{\mu i}\chi_\mu,$$

By forming the linear combination of these orbitals

$$a_i = \psi_{ib} \cos \vartheta + \psi_{iv} \sin \vartheta, \qquad b_i = \psi_{ib} \cos \vartheta - \psi_{iv} \sin \vartheta, \tag{17}$$

one may produce new pairs of orbitals which are localized on the two subsystems to an optimum extent. It follows from the orthogonality of the functions ψ that

$$\begin{gathered} \langle a_i|a_j\rangle = \langle b_i|b_j\rangle = \delta_{ij}, \qquad \langle a_i|b_i\rangle = \lambda_i = \cos 2\vartheta, \\ \langle a_i|b_j\rangle = 0 \qquad (i \neq j). \end{gathered} \tag{18}$$

The trial wave function used by the AMO method in the variation procedure is the spin-projected DODS determinant built up of the spin orbitals $a_i\alpha$ and $b_i\beta$ constructed according to Eqs. (17). Depending on whether one uses identical "mixing parameters" ϑ (i.e., λ) for all the orbital pairs, or one seeks the energy minimum by optimizing the mixing parameter for each pair of orbitals separately, we get the "one-parameter" and "many-parameter" versions of the AMO method. The former is evidently much simpler, while for the latter the flexibility of the trial function is much larger, though its optimization is also much more difficult.

D. The Pairing Theorem

As we have seen, the spatial orbitals a_i and b_i used in the AMO method, although not all of them are mutually orthogonal, have neverthe-

---

[11] According to Adams (1963), the wave functions of the AMO (and EHF) type can describe, at least partly, not only the short- and long-term correlation of the electrons with opposite spins, but also the short- (Fermi hole) and long-term correlation of the electrons with identical spin.

less some far-reaching orthogonality properties: according to Eqs. (18) the orbitals which are filled before projection with the same spin form two orthonormalized subsets, and the orbitals a_i and b_j filled with different spins have also nonzero overlaps only by pairs. These orthogonality properties did much facilitate the treatment of the AMO formalism. Following Löwdin's proposition, Amos and Hall (1961) proved the so-called "pairing theorem" (see also Löwdin, 1962a; Pauncz, 1967; Ukrainsky *et al.*, 1972) which permits the generalization of this type of limited orthogonality for all DODS determinant wave functions built up of arbitrary spin orbitals $a_i\alpha$ and $b_i\beta$. This means that one can always find such transformations of the orbitals which ensure—without changing the resulting many-electron wave function—that the orbitals become orthonormalized and "paired":

$$\langle a_i|a_j\rangle = \langle b_i|b_j\rangle = \delta_{ij}, \tag{19a}$$

$$\langle a_i|b_j\rangle = \lambda_i\delta_{ij}. \tag{19b}$$

The orbitals satisfying Eqs. (19) are often called "corresponding orbitals."

The pairing theorem has a fundamental importance for the EHF method, as it is the key to obtaining explicit expressions for the different matrix elements appearing in the theory.

It follows from the known properties of the Slater determinants that the orbitals filled with the same spin may be assumed orthonormalized [Eq. (19a)] with no loss of generality. (The linear transformations providing the orthonormalization may influence at most the normalization of the many-electron wave function.) Thus Amos and Hall (1961) had to prove only that one can also always obtain the pairing (19b) by performing appropriate unitary transformations within the two sets of orbitals a_i and b_i separately.

In order to do this, Amos and Hall (1961) proved that a rectangular matrix S of order $a \times b$ ($a \leq b$) having a rank a can be diagonalized by means of two unitary matrices[12]:

$$\mathbf{U}^+\mathbf{S}\mathbf{V} = \mathbf{T}. \tag{20}$$

Here **U** and **V** are unitary matrices and **T** is a diagonal matrix. The matrix formed from the overlaps $\langle b_i|a_j\rangle$ changes just according to this formula, if one subjects the orbitals b_i and a_j to the unitary transformations **U** and **V**, respectively.

The Hermitian matrix $\mathbf{S}^+\mathbf{S}$ can be diagonalized by the unitary matrix **V** of order $b \times b$:

$$(\mathbf{S}^+\mathbf{S})\mathbf{V} = \mathbf{V}\mathbf{D}. \tag{21}$$

---

[12] The case $a \neq b$ occurs when the number of orbitals filled with spins α and β, respectively, is different.

Here **D** is the diagonal matrix containing the eigenvalues. One obtains after multiplying Eq. (21) from the left by **S**:

$$(\mathbf{SS}^+)(\mathbf{SV}) = (\mathbf{SV})\mathbf{D}. \tag{22}$$

This is again an eigenvalue equation with the same eigenvalues as Eq. (21). The rank of **S** and, thus, of **SV** is a, therefore the number of linearly independent columns of **SV** is also a. It follows from Eq. (21) that $(\mathbf{SV})^+(\mathbf{SV}) = \mathbf{D}$. Based on this result Amos and Hall (1961) proved that the remaining $b - a$ columns of **SV** contain only zeros and correspond to zero eigenvalues in **D**. (The other eigenvalues are equal to the $|\lambda_i|^2$ values.) Furthermore, they showed that the a linearly independent columns are also orthogonal to each other. Therefore, the linearly independent columns of **SV**, after being normalized to unity, form the unitary matrix **U** of order $a \times a$. This may be written as

$$\mathbf{SV} = \mathbf{UT}, \tag{23}$$

where **T** is a diagonal matrix. After multiplying Eq. (23) from the left by $\mathbf{U}^{-1} = \mathbf{U}^+$ we arrive at Eq. (20).

One has to mention (though it does not influence the correctness of the pairing theorem) that the above algorithm becomes partly inapplicable to finding the paired orbitals if the rank of **S** is less than a, i.e., if there are one or more zero λ_i-s in (19b). Owing to the round-off errors, one may expect a numerical instability already in the case of small, though not strictly zero λ_i values.

Pauncz (1967) demonstrated that a necessary condition of the pairing is that the unitary matrices **U** and **V** must diagonalize the Hermitian matrices $\mathbf{S}^+\mathbf{S}$ and $\mathbf{SS}^+$, respectively. This is, of course, in agreement with the above derivations. Such a pairing procedure was used, e.g., by Gorlov and Ukrainsky (1973, 1974), since it is applicable also in the case of zero λ values. It is, however, easy to see (Mayer *et al.*, 1973) that the two diagonalizations performed independently of each other do not guarantee the sufficient conditions for the pairing if there are degenerate eigenvalues, i.e., if two $|\lambda|$ values are equal. This may happen due to symmetry reasons (as it is the case, e.g., for the π-electrons of benzene). In fact, in such a case the unitary matrices obtained in the diagonalization procedure are not uniquely determined. No similar problem can occur for the above algorithm of Amos and Hall (1961), as the matrix **U** is built up by using the matrix **V**.

These difficulties with the pairing algorithms may be solved (Mayer, 1977) by utilizing the fact that the problem of pairing four orbitals (two of spin α and two of spin β) can be solved analytically. On this basis one could construct also a complete pairing procedure, by using the method of two by two rotations; it is, however, sufficient to apply this approach for the subspaces of the orbitals corresponding to the $|\lambda|$ values lower than some threshold (if the Amos–Hall algorithm is used) or to the degenerate eigenvalues (for Paunecz's method).

It is to be emphasized once again that the conditions (19) do not represent any restriction (constraint) from the point of view of the EHF variational problem.[13] In fact, one may always satisfy these conditions without changing the determinantal many-electron wave function (except, possibly, the normalization); accordingly, the component of the wave function corresponding to a given multiplicity and selected out by the projection operator is also invariant under these transformations of the orbitals. With other words (e.g., Löwdin, 1976), the projected wave function may be essentially characterized by the same fundamental quantities (projectors into the subspaces spanned by the two sets of filled orbitals or the corresponding density matrices if the LCAO formalism is used) which define the single determinant. This fact has many important consequences.

E. The Spatial Symmetry Properties of the EHF Wave Function

Popov (1970) and Gerratt (1971) investigated the spatial symmetry properties of the EHF wave functions. They showed that for the molecules having spatial symmetry and an even number of electrons the (singlet) spin-projected determinant wave function may correspond to the symmetry of the Hamiltonian in two cases. The first arises when each of the two orbital sets a_i and b_i, filled before projection by spins α and β, respectively, transform according to irreducible representations of the full symmetry group G of the molecule. The second case becomes possible if the group G has at least one "halving subgroup" g (i.e., a subgroup the number of elements of which is exactly a half of the original group G). In this case the spin-projected determinant wave function will be symmetry-adapted also if the orbitals a_i and b_i separately are basis functions of irreducible representations only of the subgroup g, while the other operations of the group G interchange the two sets of orbitals. (The orbitals a_i and b_i must span equivalent irreducible representations of the subgroup g.) The first case is quite similar to the situation in the RHF method and is an evident consequence of the fact that the single determinant wave function, and thus also the projected one, is uniquely determined by these two subspaces of orbitals. It is also very easy to understand the second case, if one takes into account that due to the symmetry (Pauncz, 1967) of the Sanibel coefficients (n is the number of electron pairs):

$$c_{n-k} = (-1)^{n+S} c_k, \tag{24}$$

---

[13] This is valid for any type of wave function that can be obtained from a Slater determinant by applying a linear operator.

the interchange of the sets of orbitals filled before projection with spins α and β, respectively, can change the spin-projected wave function by not more than a phase factor. (In contrast to the first case, the unprojected determinant has no strict spatial symmetry in the second one.) It should be noted that Gorlov and Ukrainsky (1973, 1974) erroneously assumed that the EHF wave function is necessarily symmetry-adapted and corresponds to one of the aforementioned two cases. In fact, though rarely, one may encounter the spatial "symmetry dilemma" also in the EHF case (e.g., Mayer and Kertész, 1975), because the spin-projected EHF method does not involve any explicit restrictions for the spatial symmetry of the wave functions.

F. The Limit of a Very Large Number of Electrons

The relations between the EHF and UHF (or projected and unprojected AMO) methods in the limit of a very large ($N \rightarrow \infty$) number of electrons was investigated by many authors (e.g., Pauncz *et al.*, 1962a; Pauncz, 1962, 1967; Löwdin, 1962a; Ukrainsky, 1972; G. Biczó, unpublished results; Martino and Ladik, 1971; Kruglyak and Ukrainsky, 1970; Ukrainsky and Kruglyak, 1970; Misurkin and Ovchinikov, 1972; Mayer and Kertész, 1976; Mayer, 1977). The most general results were published perhaps by Ukrainsky (1972). He investigated the interrelations which hold between the quantities (denoted in Section III as A_b^a) occurring in the expression for the total energy (E_{proj}) of a spin-projected determinant for the case of a singlet or of other very low ($S \ll N$) multiplicities.[14] He showed that in the limit of $N \rightarrow \infty$ the expression for E_{proj} goes over into the expression for the energy (E_{DODS}) of the unprojected DODS determinant. Based on this consideration Ukrainsky (1972) concluded that

$$\lim_{N \rightarrow \infty} (E_{\text{proj}} - E_{\text{DODS}}) = 0. \tag{25}$$

This is, however, not true (Mayer and Kertész, 1976), because one must properly take into account that both of the energy values diverge in the limit of $N \rightarrow \infty$. It is true that the terms which appear in the projected case but are absent in the unprojected one become negligible as compared to the terms present in both cases, but these small terms, each of which tends to zero when $N \rightarrow \infty$, occur in double sums, thus their number goes to infinity. As the total energies diverge, the Eq. (25) assumed by Ukrainsky (1972) does not follow from his results. In fact (Mayer and Kertész, 1976) the difference between the total projected and unprojected energies may tend to a constant nonzero limit or diverge slowly (e.g.,

---

[14] If all λ_i fulfill the relation $0<|\lambda_i|<1$, one obtains a nonzero limit for the ratio A_b^a/A_0^0 if $b = 0$, but this ratio vanishes (Mayer, 1977) as $1/N^b$ if $b > 0$ (N is the number of electrons).

logarithmically). Nevertheless, Ukrainsky's statement is valid for that quantity which has a real physical importance, i.e., for the value of the energy per particle:

$$\lim_{N\to\infty} [(E_{\text{proj}} - E_{\text{DODS}})/N] = 0. \tag{26}$$

The situation may be well illustrated by Figs. 1 and 2 showing the results of some model calculations (Mayer and Kertész, 1976) performed for the π-electrons of some polyene chains by using the PPP integral approximation.

The relation (26) permits determination of the different energetic characteristics of very large systems without performing the spin projection explicitly. [There is also a useful relation for the spin densities (Kruglyak and Ukrainsky, 1970; Ukrainsky and Kruglyak, 1970; Ukrainsky, 1972)]. Thus one can, so to speak, "legalize" the use of the simple UHF method in the problems of solid-state physics, which is of great importance in treating antiferromagnetism and the correlation of mobile electrons in long chains or crystals (e.g., Kertész, 1977; Kertész *et al.*, 1976, 1979).

G. The Method of the "Composite Hamiltonian"

After this consideration of the most important properties of the wave functions used in the EHF method, we shall give a brief survey of the

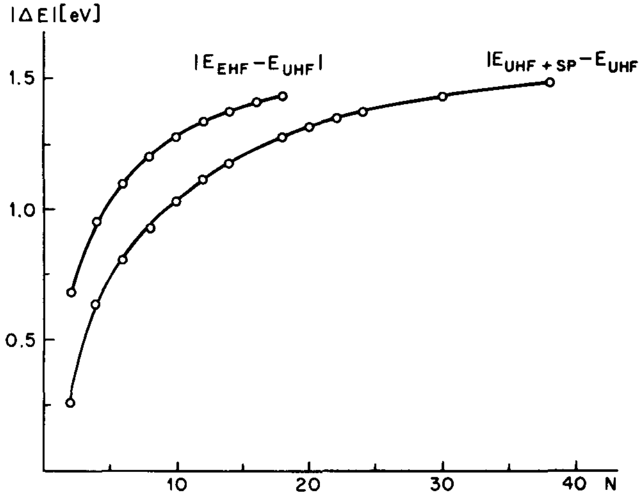


Fig. 1. The difference between projected (UHF+SP) and unprojected UHF energies and between the EHF and UHF energies for polyenes as a function of the chain length (number of π-electrons). From Mayer and Kertész (1976).

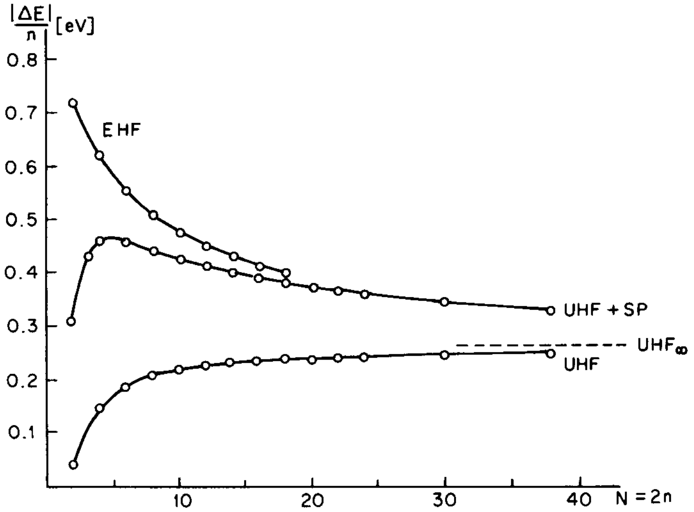


Fig. 2. Energy improvement per electron pair with respect to the RHF solution given by some DODS methods as a function of the polyene chain length (number of π-electrons). The UHF value for an infinite chain is also indicated. From Mayer and Kertész (1976).

different approaches used to treat the variation problem. (The AMO method was considered separately owing to its historical importance and to the specific role it played in the development of the field.)

In order to treat the general variation problem going beyond the AMO formalism, Löwdin (1955c, 1966) proposed the method of the "composite Hamiltonian." For the spin-projected determinant $\hat{O}^S\Phi$ the matrix element of the Hamiltonian $\hat{H}$ may be written as

$$\langle \hat{O}^S\Phi|\hat{H}|\hat{O}^S\Phi\rangle = \langle\Phi|\hat{O}^{S+}\hat{H}\hat{O}^S|\Phi\rangle, \tag{27}$$

i.e., it has the same form as the matrix element of the "composite Hamiltonian" $\hat{O}^{S+}\hat{H}\hat{O}^S$ for the original (unprojected) determinant. Therefore, if one takes the normalization $\langle\hat{O}^S\Phi|\hat{O}^S\Phi\rangle = \langle\Phi|\hat{O}^{S+}\hat{O}^S|\Phi\rangle$ properly into account, then the problem may be reformulated as if the Hamiltonian were the "composite Hamiltonian" containing also many-electron terms, but the wave function is a single Slater determinant and not a sum of many determinants.[15] Though Löwdin (1955c) gave a formal generalization of

---

[15] By taking into account the properties of the spin projection operator, one can get the following simplifications:

$$\hat{O}^{S+}\hat{H}\hat{O}^S = \hat{H}\hat{O}^S, \qquad \hat{O}^{S+}\hat{O}^S = \hat{O}^S.$$

These may be used also to obtain (e.g., Löwdin, 1976) from this approach the generalized Brillouin theorem discussed in Sections II,J and III,B.

the HF equations for the case of Hamiltonians containing many-electron interactions, this approach remained of only abstract theoretical importance, because it was not possible to write down explicitly and treat the terms of the "composite Hamiltonian" corresponding to the n-particle interactions and the necessary density matrix elements. Exceptions were the applications to some two- and three-electron systems, but the method encountered serious convergence difficulties even in these cases (Lunell, 1970).

H. Goddard's GF Equations

As it is known, one of the fundamental approaches to the problem of spin functions is based on the investigation of the irreducible representations of the "symmetric group" S_N. It is enough to refer here to the "branching diagram" of the Kotani–Yamanuchi group and their "genealogical spin functions" forming an orthonormalized basis in the given S, S_z subspace (Kotani *et al.*, 1955). These functions were essentially rederived by Goddard, who considered the different antisymmetric spin eigenfunctions by using Young's orthogonal representations as well as the corresponding variational problems (Goddard, 1967a,b,c, 1968a–e, 1969, 1970, 1972; Ladner and Goddard, 1969; Blint and Goddard, 1970; Guberman and Goddard, 1970).

The detailed equations were obtained for the so-called GF wave function[16] (Goddard, 1968a), which is equivalent (Goddard, 1967b, 1968a; Löwdin and Goscinski, 1970; Mano, 1970; Snook, 1972) to a spin-projected Slater determinant in the "principal case" $S = S_z$. Accordingly—though the formalisms are entirely different—the GF method is fully equivalent with the EHF in the most important "principal case." (The EHF method is, however, applicable not only to the "principal case" but also to any pairs of S and S_z values.) Though Goddard (1967c, 1968b,e, 1969) successfully performed some rather interesting calculations, one encountered serious difficulties concerning his GF equations. Disregarding some minor ambiguities in their published form (cf. Mayer *et al.*, 1973), the main defect of the GF equations is that they are given only for the use in a complicated iterative scheme which is, however, not convergent in most cases (e.g., I. Ukrainsky, private communication, 1972). [The convergence difficulties are so severe, that, e.g., Gorlov and Ukrainsky (1973, 1974) gave up the attempts (Ukrainsky, 1971;

---

[16] The treatment of the variational problem for the other "GI" spin functions is not feasible, except the case of a very small number of electrons, because there are no relationships like the pairing theorem. The so-called "generalized valence bond" (GVB) method (Goddard *et al.*, 1973) represents a simplified version of the "GI" method (spin-coupling scheme in which the spins of each electron pair are coupled together into a singlet), as the orthogonality properties of type (19) are introduced as additional constraints.

Ukrainsky and Kruglyak, 1973) to use the GF equations and turned to the direct numerical minimization of the energy.] As Goddard did not derive explicitly the conditions of a stationary energy, his results (despite their great theoretical importance) were not well applicable as a basis for further analysis and for searching for algorithms with better convergence properties.

According to the previous discussion, the EHF equations described in Section III are not identical with the GF equations but are equivalent with them in some sense: if one succeeds in finding the solutions of the GF equations, then they are related to the EHF orbitals by unitary transformations. (Thus the relationship between the GF and EHF orbitals resembles that between the canonic and corresponding UHF ones.) For a detailed analysis of the relations between the GF and EHF equations see the article by Mayer *et al.* (1973).

I. Some Other Approaches

The matrix elements of different operators, appearing in the framework of the AMO method (Pauncz *et al.*, 1962a; de Heer, 1962a; Pauncz, 1962, 1967; de Heer and Pauncz, 1963) or for the general spin-projected DODS determinants (Löwdin, 1958b; Harris, 1966, 1967a,b,c), were determined by some authors based on the consideration of different permutations. (Despite the differences in the formalisms, Goddard's method is also closely related to this approach.) These authors do not consider the spin-projected determinant wave function as a linear combination of a number of Slater determinants, but rather, treat the spatial and spin parts of the wave function separately. They utilize the Hermiticity of the antisymmetrizer and its projection operator character, and present the permutations occurring in the antisymmetrizer as products of permutations each acting only on the spatial and only on the spin coordinates, respectively. Then they find all those permutations for which the result of the integration over the spatial coordinates does not vanish due to the orthogonality (19a) or pairing (19b). The summation over the spin functions gives for each permutation one of the Sanibel coefficients. Consequently, the different matrix elements contain some quantities depending in a somewhat complicated manner on the Sanibel coefficients and on the overlap integrals λ appearing in the spatial integrations. [These quantities we shall denote as A_b^a (Mayer *et al.*, 1973) in Section III.] Thus Harris (1966, 1967a,b,c) derived the expectation values of different spin-free and spin-dependent operators for the spin-projected determinant wave functions.

We may note at this point that the derivation of the equations to be discussed in Section III required a number of other matrix elements also.

In order to obtain the different matrix elements of spin-free operators, the present author (Mayer, 1971b, 1974d, 1977; Mayer *et al.*, 1973) found it convenient to work directly with expansion (6).[17]

Harriman and co-workers derived the expressions[18] for the first- and second-order density matrices corresponding to the spin-projected determinants (Harriman, 1963a,b, 1964; Hardisson and Harriman, 1967), as well as to the determinants projected according to the spatial symmetry (Simons and Harriman, 1969). Their most important results consist in the finding that the natural orbitals corresponding to the spin-projected DODS determinant (Harriman, 1963b, 1964) are exactly the same as for the unprojected single determinant (Amos and Hall, 1961). The natural orbitals in both cases are the normalized sums and differences of the paired ("corresponding") orbitals and the spin projection changes only their occupation numbers: for the unprojected determinant (Amos and Hall, 1961) the occupation numbers are $\frac{1}{2}(1 \pm \lambda_i)$, while for the spin-projected determinant one has, in the notation of Section III, the expression (I. Mayer, unpublished):

$$1 \pm \lambda_i[A_0^1(i) + A_1^1(i)]/A_0^0.$$

Harriman and co-workers (Harriman, 1967; Sando and Harriman, 1967; Harriman and Sando, 1968) also worked out a rather complicated energy minimization procedure based on the density matrices and the steepest descent technique, and applied it to investigate the π-electron structure of some conjugated molecules and radicals.

Mestechkin's method (Mestechkin, 1967, 1973; Mestechkin and Whyman, 1974) is also based on the consideration of the density matrices.

---

[17] Owing to the extremely great number of determinants appearing in this case and to the absence of the full orthogonality between the orbitals, it was assumed in the literature that this expansion cannot be managed and its use must be avoided even at the expense of introducing further complications into the formalism (e.g., Lefebvre and Prat, 1967, 1969; Burden, 1972). However, by utilizing the Hermiticity and idempotency properties of the spin projection operator, as well as that it commutes with the spin-free operators, it was possible to work out a simple (though sometimes tedious) technique based on Löwdin's (1955a) general formulas (see also, e.g., McWeeny and Sutcliffe, 1969) for the matrix elements between the determinants built up of nonorthogonal orbitals. It seems that it is much easier to realize this method than to explain (Mayer, 1977) its details by words. In fact, this technique is straightforward to such an extent that the author set up a computer program performing all the time-consuming steps of derivations directly in symbols. (This program was used to check once again the EHF equations discussed in Section III, derived originally by hand.)

[18] The minor inaccuracies found in them are noted, e.g., by Ukrainsky (1972), G. Biczó (unpublished) and Philips and Schug (1974b). As a consequence, the formula for the spin density given in Pauncz's book (1967) contains also a minor error (I. Mayer and T. Pollák, unpublished); the correct formula may be found in the work of Ukrainsky (1972) and Harris (1966); see also Section III,F.

In this approach, however, the central elements of the theory are not the Sanibel coefficients and the quantities formed from them and from the overlap integrals λ, but the relative weights of the components with different multiplicities present in the unprojected determinant. (The orbitals do not appear explicitly in the formalism at all.) Accordingly, this is a line which is completely disconnected from the approaches used by other authors. For this reason we shall not discuss this theory [described also in a recent monograph of Mestechkin (1977) on density matrices] in detail, but only point out that it has been developed into a rather powerful method. Thus Klimo and Tiňo (1976, 1977, 1978; V. Klimo and J. Tiňo, private communication, 1979; Tiňo *et al.*, 1979) used it to perform a series of rather interesting calculations for different radicals both at the semiempirical and *ab initio* levels.

Rosenberg and Martino used a direct energy minimization procedure based on the conjugated gradient and steepest descent techniques in performing some very interesting calculations at the *ab initio* level (Rosenberg, 1975; Rosenberg and Martino, 1975). They pointed out (Rosenberg and Martino, 1975) that their condition for the vanishing energy gradient is equivalent to the equations (Mayer *et al.*, 1973) described in Sections III,B and C, and, therefore, their method represents an indirect way of solving these equations.

The different approaches (e.g., Poshuta and Kramling, 1968; Sullivan, 1968, 1972; Musher, 1969; Hameed *et al.*, 1969; Matsen and Cantu, 1969; Ruedenberg and Poshuta, 1972; Lim, 1975; Gallup, 1968, 1969a,b, 1970, 1973a,b,c; Heikes and Gallup, 1970) based on the group-theoretical formalism of the so-called "spin-free quantum mechanics" (Matsen, 1964) exhibit a considerable similarity to Goddard's methods. We call the attention especially to the articles by Gallup, who also investigated the variational problem corresponding to the EHF method; his iterative scheme is, however, also not free from convergence difficulties.

One must discuss separately the specific case of two electrons, for which the solution of the variational problem does not represent any serious difficulty. In the two-electron case the performance of the EHF method is extremely good: up to 90–98% of the correlation energy corresponding to the given symmetry species (e.g., of the radial correlation for He) can be taken into account (Schull and Löwdin, 1956a,b, 1959; Löwdin, 1966; Hurst *et al.*, 1968; Chong, 1966; Goddard, 1968b). If one performs projection also according to the symmetry (cf. Section II,K), then a similar percentage of the total correlation energy can be obtained (Auffray and Percus, 1962). These very good results may be connected with the following specific property (Davidson and Jones, 1962; Bunge, 1967) of the projected wave functions, which is, unfortunately, valid only in the two-electron case: One can always present in the form of an appropriately projected single determinant the wave function resulting from a full CI between all the configurations which can be set up by taking for each symmetry species the first two natural orbitals of the exact wave function, having the largest occupation numbers. The two-electron case was investigated thoroughly by Shull and Löwdin (1956a,b, 1959) and Coulson and co-workers[19] (Coulson, 1964, 1966; Coulson and Hibbert, 1967; Hibbert, 1967, 1968; Froese, 1966). An interesting calculation by using the "method of moments" was done by Szondy and Kapuy (1976).

J. Brillouin Theorem and Generalized Brillouin Theorem

The well-known Brillouin theorem of the RHF or UHF methods ("the singly excited configurations do not interact with the HF ground state") is usually derived as a consequence of the HF equations. It is, however, known (Lefebvre, 1959; Dahl *et al.*, 1970; Mayer, 1971a) that this procedure may also be reversed: one can at first prove that the Brillouin theorem is equivalent to the variational principle and then obtain the HF equations from this theorem. Since the author's approach to the EHF variation problem represents a generalization of this idea, it seems worthwhile to sketch here very briefly a derivation of this type (Mayer, 1971a); this may be of interest also because this is perhaps the simplest possible way of obtaining the UHF equations.

At first, we shall show that the Brillouin theorem must hold for the single-determinant wave function Φ_0 having the absolute minimum energy, or else one could construct another single-determinant wave function which would have an even lower energy. In fact, the wave function Φ_0 and the singly excited wave function Φ_1 [in which one of the orbitals (say a_k) is replaced by an arbitrary orbital c orthogonal to all the filled orbitals of the same spin] are determinants differing in only one column. Therefore, according to the properties of the determinants, their linear combination $\Phi = c_0\Phi_0 + c_1\Phi_1$ can be also presented as a single determinant. We may assume with no loss of generality that both Φ_0 and Φ_1 are normalized and we may determine the coefficients c_0 and c_1 in such a way as to obtain a minimum energy for the wave function Φ. As $\langle\Phi_0|\Phi_1\rangle = 0$, the lowest root of the secular equation is

$$E_2 = H_{00} + \tfrac{1}{2}(H_{00} - H_{11})\left[\left(1 + \frac{4|H_{01}|^2}{(H_{00} - H_{11})^2}\right)^{1/2} - 1\right]. \qquad (28)$$

where $H_{ij} = \langle\Phi_i|\hat{H}|\Phi_j\rangle$. If we assume that $H_{01} \neq 0$, we have from Eq. (28) $E_2 < H_{00}$. But H_{00} is the energy corresponding to Φ_0, and it was assumed

---

[19] They used the notation UHF in the sense of our EHF.

to be the *absolute* minimum. We arrive at a contradiction; the indirect proof of the theorem is completed.

Thus the Brillouin theorem gives a necessary condition for the single-determinant wave function with the lowest energy. It is also easy to show (e.g., Mayer, 1971a) that this theorem gives also the necessary and sufficient conditions for the energy to be stationary: $\delta(\langle\Phi|H|\Phi\rangle/\langle\Phi|\Phi\rangle) = 0$. Expressing the Brillouin theorem by means of the one-electron orbitals and introducing the usual Coulomb and exchange operators we have

$$\begin{aligned}\langle\Phi_1|\hat{H}|\Phi_0\rangle &= \langle c|\hat{H}^N|a_k\rangle + \sum_{\substack{j\\(j\neq k)}} ([ca_j|a_ka_j] - [ca_j|a_ja_k]) \\ &\quad + \sum_j [cb_j|a_kb_j] \\ &= \int c^* \left\{\left[\hat{H}^N + \sum_{\substack{j\\(j\neq k)}} (\hat{J}_j^a - \hat{K}_j^a) + \sum_j \hat{J}_j^b\right] a_k\right\} dv = 0. \qquad (29)\end{aligned}$$

For the notations see, e.g., the article by Mayer (1974b) or Section III.

As c is an arbitrary orbital orthogonal to all the a_j, the function in braces may have components only in the subspace of the filled orbitals a_j, and we obtain the UHF equations:

$$\hat{F}_k^a a_k = \left[\hat{H}^N + \sum_{\substack{j\\(j\neq k)}} (\hat{J}_j^a - \hat{K}_j^a) + \sum_j \hat{J}_j^b\right] a_k = \sum_j \varepsilon_{kj}^a a_j. \qquad (30)$$

One may add the self-repulsion terms and diagonalize the matrix $\boldsymbol{\varepsilon}^a$ in order to get the UHF equations in the canonical form.

In the case of other classes of wave functions the Brillouin theorem must be replaced by the generalized Brillouin theorem[20] (e.g., Lefebvre and Smeyers, 1967; Kaldor, 1968a–d; Smeyers and Pinto Suarez, 1968; Lunell, 1973; Kaldor *et al.*, 1968; Levy and Berthier, 1968; Coulson, 1971; Grein and Chang, 1971; Mayer, 1971b, 1973b, 1977; Mayer *et al.*, 1973; Löwdin, 1976). As pointed out in Epstein's (1974) monograph, it has become a generally accepted approach to treat the variational problem connected with the different classes of trial wave functions on the basis of the corresponding generalized Brillouin theorem.

For the spin-projected determinant wave functions the generalized Brillouin theorem may be formulated as follows. If Ψ is a spin-projected determinant having a stationary energy value and Ψ_1 is any "singly excited" spin-projected determinant (i.e., which can be obtained from Ψ by replacing one of the filled orbitals by an arbitrary one) then

$$\langle\Psi_1|\hat{H} - E|\Psi\rangle = 0. \qquad (31)$$

---

[20] Sometimes it may be suitable to use not the ordinary Brillouin theorem but the generalized one also in the UHF case (Lunell, 1973; Mayer, 1974b).

Here $\hat{H}$ is the total Hamiltonian of the system and E is the (stationary) energy value belonging to Ψ.

The EHF equations discussed in Section III were derived by the author (Mayer, 1971b, 1974a,c; Mayer *et al.*, 1973; Mayer and Kondász, 1975) based on this theorem. Therefore we will not discuss here the proofs of the theorem, as a quite general one will be presented in Section III,B. We mention only that for the class of wave functions considered, the most general proof (applicable to an arbitrary number of electrons) was perhaps previously given by Kaldor (1968a). Based on the method of the "composite Hamiltonian," he gave some formal EHF equations containing 1, 2, . . . , N-electron terms and obtained the generalized Brillouin theorem as their consequence. It is surprising, therefore, that by using the generalized Brillouin theorem Kaldor (1968a) obtained a wave function for the Li atom which did not satisfy the equations from which the theorem was derived. This situation is most probably due (Mayer *et al.*, 1973) to some minor inaccuracies in the summations, occurring in Kaldor's equations. (The derivation of the generalized Brillouin theorem was, however, not influenced by these errors.)

Lefebvre and Smeyers (1967), Kaldor (1968a–d; Kaldor *et al.*, 1968), and others (e.g., Lunell, 1970, 1973; Grein and Chang, 1971; Smeyers and Pinto Suarez, 1968) have solved the variational problem by using algorithms based immediately on the generalized Brillouin theorem. In these algorithms one repeatedly finds corrections for the orbitals (either for all the orbitals simultaneously or for one orbital at a time) by requiring the fulfilment of the theorem in the first order of the corrections. To do this, one has to perform in each step a limited CI calculation between the "ground" and "singly excited" configurations and update the orbitals according to the CI coefficients obtained for the excited configurations. The procedure is then continued until the generalized Brillouin theorem is fulfilled within the given convergence criterion.

K. Some Simplified Versions and Generalizations of the EHF Method

Martino and Ladik (1970) derived EHF equations for the so-called "one λ" approximation which corresponds to the one-parameter version of the AMO method: only equal λ_i values are permitted in (19b). In order to exclude the off-diagonal Lagrangian multipliers the "coupling operator formalism" (Birs and Fraga, 1963; Huzinaga, 1969) was applied (H. Sklenar and J. Ladik, unpublished). The numerical realization, however, met unsolvable convergence difficulties. Then complying with Ladik's request, the present author performed an independent rederivation and analysis of the "one λ" EHF equations and found that due to a not quite consistent use of the Lagrangian multiplier technique Martino and Ladik

(1970) lost an important additional condition (cf. Section II,L). This condition must be added to their equations to ensure the optimum common value for the λ-s, but it was found to be too complicated to be programmed. Instead, Martino and Rosenberg (F. Martino, personal communication, 1973) used a technique of direct minimization of the energy, but the "one λ" approximation was shortly given up. [Its generalization to the crystal orbital method (Ladik and Martino, 1970) was superfluous in light of the later work of the same authors (Martino and Ladik, 1971) and the other results discussed in Section II,F.]

It represents an approximation in another respect if one uses instead of the full spin projection operator (5) only one of its terms. This is the so-called "simple (or single) annihilation" method (Amos and Hall, 1961; Amos, 1962; Amos and Snyder, 1964; Snyder and Amos, 1964, 1965; Hall and Amos, 1965). In this method one annihilates only the component of the DODS determinant wave function which has the closest multiplicity to that of the ground state and is, therefore, usually of the highest weight among the components with unwanted multiplicities. (Usually one excludes the triplet in the even-electron case and the quartet for the odd-electron one.) As the wave function becomes a sum of a number of determinants even in the case of the simple annihilation, this method—after a period of popularity—is now used only rarely and only for subsequent "correction" of the UHF wave functions (e.g., Salotto and Burnelle, 1970; Philips and Schug, 1974a; Yonezawa *et al.*, 1969; Kruglyak *et al.*, 1973; Brown and Williams, 1973; Stösser *et al.*, 1975; Cremaschi *et al.*, 1976).

More recently, however, Smeyers and co-workers worked out an extremely simple and ingenious variant of the partial annihilation, called the "half projected HF" (HPHF) method (Smeyers, 1971; Smeyers and Doreste-Suares, 1973; Smeyers and Delgado-Barrio, 1974, 1976, 1977, 1978; Smeyers and Buceña, 1978; Cox and Wood, 1976). In this method one uses a variational wave function consisting of only two determinants which can be obtained from each other by interchanging all spins α and β. They showed that due to the symmetry property (24) of the Sanibel coefficients each second multiplicity is absent from this type of wave function (e.g., for the case of an even number of electrons there remain the singlet, quintet, etc. components, but the triplet, heptet, etc. are absent). The method gives results which are close to the EHF ones and may be considered as a happy compromise between the spin "symmetry dilemma" and other troubles connected with the UHF method, on the one hand, and the complexity of the EHF equations, on the other.

It may be worthwhile to note here that all the considerations discussed in Section III for the EHF case are essentially applicable also in the HPHF one. One may obtain the equations for the HPHF orbitals by performing some trivial substitutions in the EHF equations (I. Mayer, unpublished): one has to set equal to zero all the Sanibel coefficients except $c_0 = \frac{1}{2}$ and $c_n = (-1)^{n+S}/2$. As a consequence, only those A_b^a coefficients [see Eqs. (53)] survive for which $b = 0$ and/or $a = b$. The quantities B_r^a occurring in the definition of the A_b^a-s will also reduce either to unity or to a simple products of some $|\lambda_i|^2$-s. This leads to a considerable simplification of the equations, of course.

As discussed already, the AMO method is essentially a simplified version of the EHF one. We may note additionally that one may, in turn, consider the EHF method a generalization of the AMO in which not only the mixing parameters are optimized but also one performs simultaneously a variational optimization of the "starting" molecular orbitals (they are also the natural orbitals for the DODS determinant and its spin-projected counterpart). The recent "simplified EHF" procedure of Lengsfield and Schug (1978) is essentially an AMO calculation in which the starting orbitals are obtained from the natural orbital analysis of the UHF wave function.

There are two lines of generalization of the spin-projected EHF method. One of them is an immediate generalization into the projected HF method in a more general sense, in which one applies, besides the spin projection, also projection according to the spatial symmetry (e.g., Löwdin, 1964, 1966; Auffray and Percus, 1962; Kancerevičius, 1968; Bendazzoli *et al.*, 1971; Bunge and Bunge, 1971; Jucys and Lazauskas, 1974; Laskowski and Lunell, 1975; Lunell, 1979). In the case of atoms, e.g., one can take into account a part of the angular correlation in this way.

Another line of generalization of the EHF method (Brigman and Matsen, 1957; Burke and Mulligan, 1958; Kerwin and Burke, 1962; Poshuta and Kramling, 1968; Hameed *et al.*, 1969; Kaldor and Harris, 1969; Kaldor, 1970a,b, 1972; Kunik and Kaldor, 1971, 1972; Ladner and Goddard, 1969; Blint and Goddard, 1970; Goddard, 1972; Heikes and Gallup, 1970; Lunell, 1973) resembles in some sense the multiconfigurational SCF method, though it remains within the framework of a generalized one-electron picture. In this approach one takes into account that for more than two electrons there is, in general, more than one linearly independent spin function belonging to the same pair of S and S_z values ("spin degeneracy"). Accordingly, the wave function is constructed as a linear combination of these different spin functions. Then in performing the variational procedure one has to optimize simultaneously both the orbitals and the coefficients at the different spin functions. The absence of the orthogonality between the orbitals and the rapidly increasing number of the independent spin functions seems to represent an impassable barrier to the wider use of the wave functions of such increased flexibility.[21] There is a similar "nonorthogonality problem" also in the case when one builds up the starting wave functions of spin orbitals which are not of pure spin α or β but have general spin functions ("general spin orbitals," GSO). In this case (Löwdin, 1976) the number of spatial orbitals is essentially doubled (the spatial parts corresponding to spins α and β, respectively, may be different) and one has to apply projections according to eigenvalues of both operators $\hat{S}_z$ and $\hat{S}^2$ (e.g., Lunell, 1968, 1970, 1972a; Beebe and Lunell, 1975). (For an unprojected version of this approach, see, e.g., Delgado-Barrio *et al.*, 1977.) The nonorthogonality problem was the reason for the failure of some other methods also. We may mention the "nonpaired spatial orbital" (NPSO) method (for a review, see Pauncz, 1967, 1969), which was proposed by Linnett as a generalization of the "semilocalized orbitals" used by Coulson and Fischer (1949); if the nonorthogonality problem could be solved, the NPSO method would represent a rather interesting and useful application of the DODS idea.

L. The Problem of Lagrangian Multipliers

The difficulty encountered in the "one λ" EHF theory of Martino and Ladik (1970) and other problems induced us (I. Mayer and G. Biczó, unpublished) to consider the questions connected with the use of the Lagrangian multipliers in deriving the different SCF-type equations. These problems were repeatedly debated in the literature (e.g., Goddard *et al.*, 1969; Dahl *et al.*, 1970; Morikawa, 1974) and their consistent solution becomes of utmost importance if one wants to use this technique in more complicated cases such as the EHF variational problem.

The point is that, according to the standard method of variational calculus, the variation of the auxiliary functional (obtained by adding to the original one the constraints multiplied by the appropriate Lagrangian multipliers) must be performed as for a free variational problem and the auxiliary conditions (constraints) must be substituted only afterwards. Instead of this, in deriving the HF-type equations, one usually starts with the energy expression obtained by making explicit use of the orthonormality conditions and performs its variation only. This means that the overlap integrals between the orbitals, which are implicitly present in the energy formula (as zero or unity values), are fixed and not varied together with the orbitals. It is mentioned by Goddard *et al.* (1969) that the additional terms appearing from the variation of the overlap integrals are such that

---

[21] The "mixing" of different spin functions was investigated also within the framework of the AMO formalism (Pauncz, 1967, 1969; Lindner and Lunell, 1967; Lunell and Lindner, 1968; Tamir and Pauncz, 1968) but was found to be of usually minor importance.

they can be combined with the undetermined Lagrangian multipliers to form some new ones. This is, of course, true for the usual RHF or UHF case; one might only add that, e.g., in the UHF case (I. Mayer, unpublished), the "original" off-diagonal Lagrangian multipliers (i.e., those which one has if the overlap integrals are also consequently varied) are, in fact, all zero,[22] while the diagonal ones are all equal to $-E$ (E being the total many-electron energy) and thus are not related to the orbital energies. (One may trivially reproduce these results, as all the necessary matrix elements are given in Mayer, 1973a.)

In the EHF case, however, the overlap integrals λ_i between the corresponding orbitals also appear explicitly in the energy formula. As there is no restriction in (19b) for the value of λ_i in the "diagonal" case $j = i$, no Lagrangian multipliers appear by which the terms originating from the variation of the overlap integrals λ_i, induced by the variation of the orbitals, could be combined. This is the very reason why Ladik and Martino lost a condition in their derivations.

In our opinion (I. Mayer and G. Biczó, unpublished) the problem may be solved in the following way. One defines the formal expression obtained for the energy by using the auxiliary conditions as a new functional which has, strictly speaking, no immediate physical meaning, but is *equal* to the energy *when* (and only when) the auxiliary conditions are fulfilled. Then one has to search for the stationary values of this auxiliary quantity, subjected to the same conditions. Within the class of functions satisfying the auxiliary conditions, this quantity and the energy have, of course, the same stationary values realized for the same wave functions; as we have, from a mathematical point of view, two different variational problems, the Lagrangian multipliers will, of course, be different. In the EHF case one must perform an optimization also according to the λ_i values appearing in the energy formula as formal scalar parameters and introduce the additional constraints that the overlap integrals $\langle a_i | b_i \rangle$ must be equal to these (optimized) λ_i values. If one treats the EHF variational problem by using the Lagrangian multipliers in this manner (I. Mayer, unpublished), then one obtains the EHF equations exactly in the same form as they were derived by a systematic study of the different specified variations (see Mayer, 1974d, where the transformation to the form of equations described in Section III is also shown). We shall not, however, discuss in detail these alternative derivations in the present article.

---

[22] This is perhaps the case (e.g., G. Biczó, unpublished) for any Lagrangian multipliers which correspond to such conditions on the orbitals, which one can always satisfy without changing the many-electron wave function, i.e., which do not represent a true constraint in the variation problem.

III. The EHF Equations

A. Formulation of the Problem

As we have seen, the variational problem for the spin-projected determinant wave functions was investigated by a number of authors using quite different techniques and formalisms. This great variety of the methods was perhaps due to the complexity of the EHF variational problem, which makes it difficult to find a procedure that would be fully satisfactory in all respects. In this section we shall briefly summarize the approach used by the present author; he hopes that his work is a contribution to the progress in this field.

Although it leads to some repetitions, perhaps it would not be superfluous to start with a short formulation of the variational problem to be solved in the EHF method.

According to the discussion given in the previous sections, one uses in the EHF method a spin-projected determinant wave function

$$\Psi = \hat{O}^S \Phi \tag{32}$$

as trial function in the variation procedure. Here $\hat{O}^S$ is the spin projection operator (5) and Φ is a single DODS determinat wave function built up of the spatial orbitals a_i and b_i filled with electrons of spin α and β, respectively. Accordingly, in the case of an even number of electrons ($N = 2n$) and $S_z = 0$, the determinant Φ can be given as ($\hat{\mathscr{A}}$ is the antisymmetrizer):

$$\begin{aligned}\Phi = \hat{\mathscr{A}}[a_1(1)\alpha(1)b_1(2)\beta(2) \cdots a_k(2k-1)\alpha(2k-1)b_k(2k)\beta(2k) \\ \cdots a_n(2n-1)\alpha(2n-1)b_n(2n)\beta(2n)].\end{aligned} \tag{33}$$

As the operator $\hat{O}^S$ annihilates all the unwanted components of Φ, leaving unchanged the component corresponding to the desired multiplicity, $\Psi = \hat{O}^S\Phi$ is an eigenfunction of the operators $\hat{S}_z$ and $\hat{S}^2$, but is, in general, not normalized. As any projection operator, $\hat{O}^S$ is Hermitian and idempotent (e.g., Löwdin, 1960, 1966):

$$\hat{O}^S\hat{O}^S = \hat{O}^S = \hat{O}^{S+}. \tag{34}$$

Furthermore, $\hat{O}^S$ commutes with the antisymmetrizer and with any spinless operator, and, thus, also with the usual nonrelativistic Born–Oppenheimer Hamiltonian

$$\hat{H} = \sum_{i=1}^{N} \hat{H}^N(i) + \sum_{i<j} (r_{ij})^{-1} \tag{35}$$

Here $\hat{H}^N$ is the one-electron part of the Hamiltonian, containing the operators of the kinetic energy and nuclear attraction.

Now, the EHF variational problem can be formulated as follows:

determine the one-electron orbitals a_i and b_i in such a way as to obtain a stationary (minimum) energy value for the spin-projected many-electron wave function $\Psi = \hat{O}^S\Phi$.

According to the discussion given in Section II,D, one can always subject the orbitals to such transformations which ensure that they become orthonormalized (19a) and paired[23] (19b). This means that in the EHF case the doubly filled orbitals of the RHF method are in some sense replaced by pairs of orbitals which are not orthogonal to each other but are orthogonal to all the other orbitals.

Similarly to the RHF or UHF equations, the EHF equations discussed in this section are related to the one-electron orbitals and give the conditions for the many-electron wave function built up of the orbitals satisfying the equations to have a minimum (or, at least, a stationary) energy value.

B. A Proof of the Generalized Brillouin Theorem

The consideration of the EHF variational problem can be most generally based on the generalized Brillouin theorem (Section II,J) because (as will be seen) it is fully equivalent to the variational principle for the spin-projected determinant wave functions. For this reason we give here a general proof of this theorem (Mayer *et al.*, 1973). In this proof, in contrast to that of Kaldor (1968a), we shall not assume any orthogonality and pairing properties for the orbitals. This fact is of great importance from the point of view of the generality of the derivations discussed in the next sections. (In order to obtain manageable expressions, we shall utilize, of course, the orthogonality and pairing to the extent as they can be assumed for each given case.)

Let us consider the variational condition

$$\delta[\langle\Psi|\hat{H}|\Psi\rangle/\langle\Psi|\Psi\rangle] = 0 \tag{36}$$

for a wave function Ψ, having a stationary energy. By performing the variation of the fraction in Eq. (36) we obtain

$$\frac{\langle\delta\Psi|\hat{H}|\Psi\rangle\langle\Psi|\Psi\rangle - \langle\delta\Psi|\Psi\rangle\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle^2} + \text{c.c.} = 0. \tag{37}$$

Since $\delta\Psi$ contains an arbitrary phase factor, Eq. (37) is fulfilled for arbitrary variations $\delta\Psi$ if and only if both the quantity written down explicitly in Eq. (37) and its complex conjugate (c.c.) are separately equal to zero. As we may assume that $\langle\Psi|\Psi\rangle \neq 0$, we obtain from Eq. (37), by

---

[23] In Section III,G we shall consider also a case in which only limited pairing properties can be assumed for the orbitals.

introducing the energy $E = \langle\Psi|\hat{H}|\Psi\rangle/\langle\Psi|\Psi\rangle$, a very useful form of the variational principle which is completely equivalent to Eq. (36):

$$\langle\delta\Psi|\hat{H} - E|\Psi\rangle = 0. \tag{38}$$

If no restriction is put on the trial wave function Ψ, then its variation $\delta\Psi$ is fully arbitrary and Eq. (38) leads immediately to the Schrödinger equation. If, however, Ψ is required to belong to a given class of wave functions, then only those $\delta\Psi$-s are permitted which can be obtained by arbitrary variations of Ψ *within* the given class. Thus, in the case of a linear variational problem, Eq. (38) gives the secular equation, while for the nonlinear problems one can obtain from Eq. (38) the different (generalized) Brillouin-type theorems.

In our case $\Psi = \hat{O}^S\Phi$ is a spin-projected determinant. In order to vary this wave function, one has to perform the variation of the one-electron orbitals:

$$d_i \rightarrow d_i + \delta\, d_i = d_i + \eta c_i, \tag{39}$$

where $d = a$ or b, the c_i-s are arbitrary one-electron spatial orbitals and η is an arbitrary complex parameter of variation tending to zero. One obtains $\Psi + \delta\Psi$ for the most general variation if each orbital d_i is replaced in Eqs. (32) and (33) by $d_i + \delta d_i$. Thus $\Psi + \delta\Psi$ becomes a spin-projected determinant each column of which represents a sum of two terms d_i and $\delta d_i = \eta c_i$, respectively. According to the known properties of determinants, $\Psi + \delta C$ can be expressed as a sum of 2^N spin-projected determinants. However, most of these terms will contain the square or higher powers of η and can be neglected as compared with the terms of first order in η. Thus we may write

$$\Psi + \delta\Psi = \Psi + \eta \sum_{i=1}^{N} \Psi_1(d_i \rightarrow c_i). \tag{40}$$

Here $\Psi_1(d_i \rightarrow c_i) = \hat{O}^S\Phi_1(d_i \rightarrow c_i)$ is a "singly excited" spin-projected determinant which can be obtained from Ψ by replacing the orbital d_i by the arbitrary orbital c_i. Thus one obtains the most general first-order variation of Ψ as

$$\delta\Psi = \eta \sum_{i=1}^{N} \Psi_1(d_i \rightarrow c_i). \tag{41}$$

By substituting Eq. (41) into Eq. (38), interchanging the order of the summation and integration and dividing by η^* ($\eta \rightarrow 0$ but $\eta \neq 0$, or else $\delta\Psi \equiv 0$) we obtain:

$$\sum_{i=1}^{N} \langle\Psi_1(d_i \rightarrow c_i)|\hat{H} - E|\Psi\rangle = 0. \tag{42}$$

Since the orbitals c_i are arbitrary and independent of each other (we may set also $c_i \equiv 0$ for one or more c_i-s), we conclude that the necessary and sufficient condition of fulfilling the varational principle (36) for the spin-projected determinant wave function $\Psi = \hat{O}^S\Phi$ may be given as the generalized Brillouin theorem

$$\langle \Psi_1(d_i \rightarrow c_i)|\hat{H} - E|\Psi\rangle = 0, \tag{43}$$

which must hold for all d_i-s ($d = a$ or b) and arbitrary orbitals c_i.

One can show (Mayer, 1973b, 1977) that the generalized Brillouin theorem gives also a necessary condition which must be fulfilled if the energy of the spin-projected determinant is an absolute minimum. The proof is quite similar to that discussed in Section II,J for the ordinary Brillouin theorem of the UHF case, but is slightly more involved as one must take into account that the spin-projected determinants are usually unnormalized and, especially, that one may not require $\langle \Psi_1|\Psi\rangle = 0$ in the spin-projected case.[24]

C. The "Overall" Form of the EHF Equations

One obtains the EHF equations by expressing the generalized Brillouin theorem in terms of the one-electron orbitals and taking into account that Ψ_1 contains an arbitrary orbital (Mayer, 1971b, 1974c; Mayer *et al.*, 1973). To do this, one has to determine the integrals $\langle \Psi_1|\Psi\rangle$ and $\langle \Psi_1|\hat{H}|\Psi\rangle$ and the energy $E = \langle \Psi|\hat{H}|\Psi\rangle / \langle \Psi|\Psi\rangle$, i.e., different matrix elements taken between spin-projected determinants.

If $\Psi' = \hat{O}^S\Phi'$ and $\Psi = \hat{O}^S\Phi$ are spin-projected determinants with the same S and S_z values, then by making use the properties of the spin projection operator and of the expansion (6) we can write

$$\langle \Psi'|\Psi\rangle = \langle \hat{O}^S\Phi'|\hat{O}^S\Phi\rangle = \langle \Phi'|\hat{O}^S\Phi\rangle = \left\langle \Phi'\right| \sum_{m=0}^{n} c_m T_m \Big\rangle = \sum_{m=0}^{n} c_m\langle \Phi'|T_m\rangle \tag{44}$$

and similarly

---

[24] There is a consequence of mathematical character which follows from the comparison of these two proofs of different type: the absolute minimum of the energy is necessarily a stationary value also, i.e., the wave function with the lowest energy cannot be a "pathological" (e.g., boundary) point. The existence of such a wave function for which the energy reaches its exact lower limit looks evident physically, although it is not trivial to prove this mathematically (cf. Lieb and Simon, 1974, where the existence theorem for some UHF and RHF problems was proven). However, our proofs and considerations (similar to those given in Section 5 of Lieb and Simon, 1974) show that in the LCAO case (finite basis set expansion for the orbitals) such a wave function always exists and is an interior point, so the orbitals fulfill the (matrix) EHF equations derived from the generalized Brillouin theorem.

$$\langle\Psi'|\hat{H}|\Psi\rangle = \langle\hat{O}^S\Phi'|\hat{H}|\hat{O}^S\Phi\rangle = \left\langle\Phi'|\hat{H}|\sum_{m=0}^{n} c_m T_m\right\rangle = \sum_{m=0}^{n} c_m\langle\Phi'|\hat{H}|T_m\rangle. \tag{45}$$

Accordingly, the generalized Brillouin theorem can be rewritten as

$$\sum_{m=0}^{n} c_m\langle\Phi_1|\hat{H} - E|T_m\rangle = 0. \tag{46}$$

Thus one has to determine the matrix elements between Φ_1 and all the determinants occurring in the different T_m-s. This can be done by performing an appropriate classification of the orbitals, then a systematization of the determinants occurring in the different T_m-s and of the nonvanishing minors of the overlap matrices. (The latter appear in the derivations according to Löwdin's (1955a) formulas for the matrix elements between the determinants built up of not mutually orthogonal orbitals.) The values of these minors are to be determined in each case by utilizing as much as possible the orthogonality and pairing properties of the orbitals (cf. also Section II,I).

Turning to the derivation of the EHF equations, we shall at first consider the following specific case.

Let $\Psi = \hat{O}^S\Phi$ be a spin-projected determinant with a stationary energy, built up of the orbitals satisfying Eqs. (19a) and (19b), and let the "singly excited" spin-projected determinant $\Psi_1 = \hat{O}^S\Phi_1$, occurring in the generalized Brillouin theorem, differ from Ψ by replacing the orbital a_k with an arbitrary orbital c. At the moment, however, we subject the orbital c to the temporary restriction that it be orthogonal to all the filled orbitals a_i and b_i: $\langle c|a_i\rangle = \langle c|b_i\rangle = 0$. It is easy to see that in this case one obtains from Eq. (44) that $\langle\Psi_1|\Psi\rangle = 0$, and the generalized Brillouin theorem reduces to

$$\langle\Psi_1|\hat{H}|\Psi\rangle = \sum_{m=0}^{n} c_m\langle\Phi_1|\hat{H}|T_m\rangle = 0. \tag{47}$$

After a somewhat lengthy but simple derivation, Eq. (47) can be expressed as (Mayer, 1971b; Mayer *et al.*, 1973):

$$\int c^*[\hat{\mathscr{H}}^a(k)a_k + \lambda_k^*\hat{h}^b(k)b_k]dv = 0, \tag{48}$$

where the operators $\hat{\mathscr{H}}^a(k)$ and $\hat{h}^b(k)$ are defined as[25]

$$\begin{aligned}\hat{\mathscr{H}}^a(k) = {} & A_0^1(k)\hat{H}^N + A_0^1(k)\hat{J}_k^{bb} + A_1^1(k)\hat{K}_k^{bb} + \sum_{\substack{j=1\\(j\neq k)}}^{n} \{A_0^2(j,k)(\hat{J}_j^{aa} - \hat{K}_j^{aa} + \hat{J}_j^{bb}) \\ & + A_1^2(j, k)[\hat{K}_j^{bb} + \lambda_j(\hat{J}_j^{ba} - \hat{K}_j^{ba}) + \lambda_j^*(\hat{J}_j^{ab} - \hat{K}_j^{ab})]\},\end{aligned} \tag{49}$$

---

[25] In Mayer (1971b) some notations were different from those used here. For the relations between the notations see the footnote in Mayer *et al.* (1973).

$$\hat{\mathscr{h}}^{b}(k) = A_1^1(k)\hat{H}^N + \sum_{\substack{j=1\\(j\neq k)}}^{n} \{A_1^2(j, k)(\hat{J}_j^{aa} + \hat{J}_j^{bb} - \hat{K}_j^{aa} - \hat{K}_j^{bb} + \lambda_j\hat{K}_j^{ba})$$
$$+ A_2^2(j, k)[\lambda_j^*(\hat{J}_j^{ab} - \hat{K}_j^{ab}) + \lambda_j\hat{J}_j^{ba}]\}. \tag{50}$$

The generalized Coulomb and exchange operators occurring in Eqs. (49) and (50) are defined as

$$\hat{J}_j^{de}g(1) = \int \frac{d_j^*(2)e_j(2)}{r_{12}}\,dv_2 g(1), \qquad \hat{K}_j^{de}g(1) = \int \frac{d_j^*(2)g(2)}{r_{12}}\,dv_2 e_j(1), \tag{51}$$

where d, e, and g, respectively, stand for a or b. We have the following relationships for the adjoints of the operators $\hat{J}$ and $\hat{K}$:

$$\hat{J}_j^{cd+} = \hat{J}_j^{dc}; \qquad \hat{K}_j^{cd+} = \hat{K}_j^{dc}. \tag{52}$$

As has been already discussed in Section II, the quantities A_b^a playing an important role in the theory are functions of the Sanibel coefficients c_m and the overlap integrals λ_i between the corresponding orbitals. They are defined as

$$A_b^a(k_1, k_2, \ldots, k_a) = \sum_{m=b}^{n-a+b} (-1)^m c_m B_{m-b}^a(k_1, k_2, \ldots, k_a), \tag{53a}$$

where

$$B_r^a(k_1, k_2, \ldots, k_a) = \sum_{\substack{1\leqq i_1<i_2<\cdots<i_r\leqq n\\ i_p\neq k_1,k_2,\cdots,k_a\\ (p=1,2,\cdots,r)}} \prod_{q=1}^{r} |\lambda_{i_q}|^2 \tag{53b}$$

and $B_0^a = 1$ ($A_b^a = 0$ if $a > n$ or $b > a$). The basic recursion relation between the A_b^a-s is (Mayer *et al.*, 1973):

$$A_b^a(k_1, k_2, \ldots, k_a) = A_b^{a+1}(k_1, k_2, \ldots, k_{a+1})$$
$$+ |\lambda_{k_{a+1}}|^2 A_{b+1}^{a+1}(k_1, k_2, \ldots, k_{a+1}) \tag{54}$$

from which one can obtain a number of other useful recursion formulas (Mayer, 1978).

Because c is an arbitrary orbital which is, however, orthogonal to all the orbitals a_i and b_i filled in Φ, it follows from Eq. (48) that the function $\mathscr{H}^a(k)a_k + \lambda_k^*\hat{\mathscr{h}}^b(k)b_k$ in brackets can have only components proportional to the filled orbitals. This means that this function can be written as a linear combination of the filled orbitals. Thus we obtain the "overall" form of the EHF equations (Mayer, 1971b; Mayer *et al.*, 1973):

$$\hat{\mathscr{H}}^a(k)a_k + \lambda_k^*\hat{\mathscr{h}}^b(k)b_k = \sum_{i=1}^{n} (\varepsilon_{ki}^{aa}a_i + \varepsilon_{ki}^{ab}b_i) \qquad (k = 1, 2, \ldots, n) \tag{55}$$

It is easy to see that the similar equations for the orbitals b_k may be obtained from Eq. (55) if one interchanges everywhere a with b and λ with λ^*.

D. The Matrices $\boldsymbol{\varepsilon}$

The matrices $\boldsymbol{\varepsilon}$ occurring in the equations remained fully undetermined up to this point, as we have considered only the case when $\langle c|a_i\rangle = \langle c|b_i\rangle = 0$ for all i-s. Therefore Eq. (55) alone do not give all the conditions of the stationary energy. In order to get all these conditions, one must consider also the case when the orbital c is not orthogonal to the filled ones, especially to the b_i-s (Mayer, 1971b; Mayer *et al.*, 1973). As will be seen, in this case one obtains some explicit expressions for the elements of the matrices $\boldsymbol{\varepsilon}$, and this is necessary to complete the derivation of the EHF equations.

Let us first consider the case when the orbital c is not orthogonal to one of the orbitals b_l $(l \neq k)$ but is still orthogonal to all the other orbitals b_i and to all the orbitals a_i. One finds that $\langle \Psi_1|\Psi\rangle = 0$ holds also in this case, so one must again expand Eq. (47) only. After a lengthy derivation one obtains Eq. (48) but with an additional term on the left-hand side, proportional to $\langle c|b_l\rangle$. We denote this term temporarily as $\langle c|b_l\rangle Q$ (Q is a sum of a number of one- and two-electron integrals over the filled orbitals; its explicit form will be utilized later). We may combine the term $\langle c|b_l\rangle Q$ with the integral in Eq. (48) and obtain the expression

$$\int c^*[\hat{\mathscr{H}}^a(k)a_k + \lambda_k^*\hat{h}^b(k)b_k + Qb_l]\,dv = 0. \tag{56}$$

Now, we may apply considerations similar to those which led from Eq. (48) to Eq. (55), but in the present case c is not orthogonal to b_l, consequently the term containing b_l should be dropped from the linear combination occurring at the right-hand side. Thus we obtain

$$\hat{\mathscr{H}}^a(k)a_k + \lambda_k^*\hat{h}^b(k)b_k + Qb_l = \sum_{i=1}^{n} \varepsilon_{ki}^{aa}a_i + \sum_{\substack{i=1\\(i\neq l)}}^{n} \varepsilon_{ki}^{ab}b_i. \tag{57}$$

Comparing Eqs. (57) and (55) one finds

$$\varepsilon_{kl}^{ab} = -Q. \tag{58}$$

By substituting the explicit form of Q into Eq. (58) we obtain the matrix element ε_{kl}^{ab} (Mayer *et al.*, 1973):

$$\begin{aligned}
\varepsilon_{kl}^{ab} = {} & A_1^2(k, l)\lambda_l^*\langle a_l|\hat{H}^N|a_k\rangle + A_1^2(k, l)\lambda_k^*\langle b_l|\hat{H}^N|b_k\rangle \\
& + A_2^2(k, l)\lambda_k^*\lambda_l^*\langle a_l|\hat{H}^N|b_k\rangle - A_1^2(k, l)\langle b_l|\hat{H}^N|a_k\rangle \\
& + \sum_{\substack{j=1\\(j\neq k,l)}}^{n} \{A_2^3(j, k, l)[\lambda_k^*\lambda_l^*([a_ja_l|a_jb_k] - [a_ja_l|b_ka_j] + [b_ja_l|b_jb_k] \\
& - [a_lb_j|b_jb_k]) + \lambda_k^*[b_la_j|a_jb_k] + \lambda_k^*\lambda_j^*([a_jb_l|b_jb_k] - [b_la_j|b_jb_k]) \\
& + \lambda_k^*\lambda_l^*\lambda_j[a_lb_j|a_jb_k] + \lambda_k^*\lambda_j([b_jb_l|a_jb_k] - [b_jb_l|b_ka_j]) \\
& + \lambda_l^*\lambda_j([a_lb_j|a_ka_j] - [a_lb_j|a_ja_k]) + \lambda_l^*[a_lb_j|b_ja_k] \\
& - \lambda_j^*[a_jb_l|b_ja_k] + \lambda_j([b_jb_l|a_ka_j] - [b_jb_l|a_ja_k])
\end{aligned}$$

$$
\begin{aligned}
&+ \lambda_j^*\lambda_l^*([a_la_j|a_kb_j] - [a_la_j|b_ja_k])] \\
&- A_1^3(j, k, l)[[a_jb_l|a_ja_k] + [b_jb_l|b_ja_k] - [b_jb_l|a_kb_j] \\
&- [b_la_j|a_ja_k] + \lambda_k^*([b_jb_l|b_kb_j] - [b_jb_l|b_jb_k] - [a_jb_l|a_jb_k]) \\
&+ \lambda_l^*([a_la_j|a_ja_k] - [a_la_j|a_ka_j] - [a_lb_j|a_kb_j]) + \lambda_j^*[b_la_j|b_ja_k]] \\
&+ A_3^3(j, k, l)[\lambda_k^*\lambda_l^*\lambda_j[b_ja_l|a_jb_k] + \lambda_k^*\lambda_l^*\lambda_j^*([a_ja_l|b_jb_k] \\
&- [a_ja_l|b_kb_j])]\} + A_2^2(k, l)(\lambda_l^*[b_ka_l|a_kb_k] + \lambda_k^*[a_lb_l|b_ka_l]) \\
&- A_1^2(k, l)([b_kb_l|b_ka_k] - [b_kb_l|a_kb_k] + [a_lb_l|a_la_k] \\
&- [a_lb_l|a_ka_l] - \lambda_k^*[a_lb_l|a_lb_k] - \lambda_l^*[a_lb_k|a_kb_k]).
\end{aligned}
\tag{59}
$$

The two-electron integrals over the spatial orbital are defined as follows

$$
[d_ie_j|f_kg_l] = \iint d_i^*(1)e_j^*(2)(r_{12})^{-1}f_k(1)g_l(2)\,dv_1\,dv_2 \tag{60}
$$

(d, e, f, and g, respectively, stand for a or b).

Considering in a similar way the case when the orbital c is not orthogonal to the orbital b_k but is orthogonal to the other filled orbitals, we obtain the expression for the diagonal element ε_{kk}^{ab} of the matrix $\boldsymbol{\varepsilon}^{ab}$. In this case, however, $\langle\Psi_1|\Psi\rangle \neq 0$ and one has to use the full form (46) of the generalized Brillouin theorem. As a consequence, the total many-electron energy E also appears explicitly in the expression of ε_{kk}^{ab}:

$$
\begin{aligned}
\varepsilon_{kk}^{ab} = {}& -A_1^1(k)\langle b_k|\hat{H}^N|a_k\rangle + \sum_{\substack{j=1\\(j\neq k)}}^{n} \{A_1^2(j, k)([a_jb_k|a_ka_j] \\
&-\lambda_j^*[a_jb_k|a_kb_j] - [b_ka_j|a_ka_j] - [b_kb_j|a_kb_j] + [b_kb_j|b_ja_k]) \\
&+ A_2^2(j, k)[\lambda_j([b_jb_k|a_ka_j] - [b_jb_k|a_ja_k]) - \lambda_j^*[a_jb_k|b_ja_k]]\} \\
&+ A_1^1(k)\lambda_k^*E - \lambda_k^*q_k.
\end{aligned}
\tag{61}
$$

The total spin-projected many-electron energy is given by

$$
\begin{aligned}
E = (A_0^0)^{-1} \Bigg\{ & \sum_{j=1}^{n} [A_0^1(j)(\langle a_j|\hat{H}^N|a_j\rangle + \langle b_j|\hat{H}^N|b_j\rangle) \\
&+ A_1^1(j)(\lambda_j\langle b_j|\hat{\mathrm{H}}^N|a_j\rangle + \lambda_j^*\langle a_j|\hat{H}^N|b_j\rangle) + A_0^1(j)[a_jb_j|a_jb_j] \\
&+ A_1^1(j)[a_jb_j|b_ja_j]] + \sum_{j<k} \{A_0^2(j, k)([a_ja_k|a_ja_k] \\
&- [a_ja_k|a_ka_j] + [b_jb_k|b_jb_k] - [b_jb_k|b_kb_j]) \\
&+ A_2^2(j, k)[\lambda_j^*\lambda_k^*([a_ja_k|b_jb_k] - [a_ja_k|b_kb_j]) \\
&+ \lambda_j\lambda_k([b_jb_k|a_ja_k] - [b_kb_j|a_ja_k])]\} \\
&\qquad + \sum_{\substack{j,k\\(j\neq k)}} \{A_0^2(j, k)[a_jb_k|a_jb_k] + A_1^2(j, k)[[a_kb_j|b_ja_k] \\
&+ \lambda_k([a_jb_k|a_ja_k] + [b_jb_k|b_ja_k] - [a_jb_k|a_ka_j] - [b_jb_k|a_kb_j]) \\
&+ \lambda_k^*([a_ja_k|a_jb_k] + [b_ja_k|b_jb_k] - [a_ka_j|a_jb_k] - [a_kb_j|b_jb_k]) \\
&+ \lambda_k^*\lambda_j[a_kb_j|a_jb_k]] + A_2^2(j, k)\lambda_j^*\lambda_k[a_jb_k|b_ja_k]\} \Bigg\}.
\end{aligned}
\tag{62}
$$

In Eq. (61) q_k is the following real quantity

$$
\begin{aligned}
q_k = \sum_{\substack{j=1 \\ (j\neq k)}}^{n} & [A_1^2(j,k)(\langle a_j|\hat{H}^N|a_j\rangle + \langle b_j|\hat{H}^N|b_j\rangle + [a_jb_j|a_jb_j]) \\
& + A_2^2(j,k)(\lambda_j^*\langle a_j|\hat{H}^N|b_j\rangle + \lambda_j\langle b_j|\hat{H}^N|a_j\rangle + [a_jb_j|b_ja_j])] \\
& + \sum_{\substack{j<l \\ (j,l\neq k)}} \{A_1^3(j,k,l)([a_ja_l|a_ja_l] - [a_ja_l|a_la_j] + [b_jb_l|b_jb_l] \\
& - [b_jb_l|b_lb_j]) + A_3^3(j,k,l)[\lambda_j\lambda_l([b_jb_l|a_ja_l] - [b_jb_l|a_la_j]) \\
& + \lambda_j^*\lambda_l^*([a_ja_l|b_jb_l] - [a_ja_l|b_lb_j])]\} \\
& - \sum_{\substack{j,l \\ (j\neq l;\, j,l\neq k)}} \{-A_1^3(j,k,l)[a_jb_l|a_jb_l] + A_2^3(j,k,l)[\lambda_j([a_lb_j|a_ja_l] \\
& + [b_jb_l|b_la_j] - [b_ja_l|a_ja_l] - [b_jb_l|a_jb_l]) \\
& + \lambda_j^*([a_jb_l|b_lb_j] + [a_ja_l|a_lb_j] - [a_ja_l|b_ja_l] - [a_jb_l|b_jb_l]) \\
& - [a_lb_j|b_ja_l] - \lambda_j\lambda_l^*[a_lb_j|a_jb_l]] - A_3^3(j,k,l)\lambda_j\lambda_l^*[b_ja_l|a_jb_l]\}. \qquad (63)
\end{aligned}
$$

The elements of the matrix ε^{ba} occurring in the equations for the orbitals of type b_k can be obtained from these expressions by interchanging a with b and λ with λ^* everywhere. (The quantity q_k is invariant under these interchanges.)

One can obtain the expressions for the elements of the matrix ε^{aa} in a quite similar manner by considering the cases when the orbital c is not orthogonal to the different orbitals a_i. We shall not quote here explicitly these expressions (they may be found in Mayer *et al.*, 1973) because, as discussed by Mayer *et al.* (1973; Mayer, 1974b,d), it follows from the properties of the determinants that they do not give any independent condition for the stationary energy. (This is a consequence of the fact that a variation δa_k of the orbital a_k, which can be given as a linear combination of the filled orbitals a_i does not, in fact, change the determinant or spin-projected determinant wave function except, possibly, the normalization; consequently such a variation does not lead to a variation of the energy.) Accordingly, the elements of the matrix ε^{aa} (and ε^{bb}) can be considered as undetermined quantities, i.e., one can obtain their expressions directly from Eqs. (55), after Eqs. (59) and (61) for the elements of the matrix ε^{ab} have been substituted. To do this, one has to multiply (55) by a_l^* $(l \neq k)$ or a_k^* and integrate over dv by taking into account the orthogonality and pairing of the orbitals. The expressions for the elements of the matrix ε^{aa} (or ε^{bb}) obtained in this manner can be transformed to the form which one gets directly from the generalized Brillouin theorem by performing some algebraic manipulations making use of the identity (54).

It follows from the independence of the first-order variations (see also Mayer *et al.*, 1973, for a somewhat different reasoning) that the same expressions would result for the elements of the matrices ε also in the case when the arbitrary orbital c would be not orthogonal even to all the filled orbitals a_i and b_i simultaneously. Consequently, the EHF equations (55) together with Eqs. (59) and (61) for the elements of $\boldsymbol{\varepsilon}^{ba}$ are equivalent to the generalized Brillouin theorem and, therefore, give the necessary and sufficient conditions of stationary energy for the spin-projected determinants. (It may be emphasized once again that these equations are related to the case of orthonormalized and paired orbitals; one may, however, always ensure the fulfillment of these conditions.)

The properties of the matrices $\boldsymbol{\varepsilon}$ and some other forms of presentation for their elements were extensively discussed by Mayer *et al.* (1973; Mayer, 1974d). Here we shall briefly mention only two of the results of this analysis. The first is that the matrices $\boldsymbol{\varepsilon}^{aa}$ and $\boldsymbol{\varepsilon}^{bb}$ are Hermitian while the matrix $\boldsymbol{\varepsilon}^{ba}$ is the adjoint of $\boldsymbol{\varepsilon}^{ab}$ (Mayer *et al.*, 1973; Mayer, 1974d). The second result (it was obtained in an alternative derivation of the EHF equations, based on a systematic study of different specified variations of the orbitals) shows that the diagonal matrix element ε_{kk}^{ab} is immediately connected with the fact that the total energy (62) depends on the overlap integrals λ_k between the corresponding orbitals (Mayer, 1974d):

$$\varepsilon_{kk}^{ab} = -A_0^0(\partial E/\partial\lambda_k)\big|_{\lambda_k^*=\text{const}}. \tag{64}$$

The derivative at the right-hand side of Eq. (64) contains terms of three types: those resulting from the differentiation of the λ-s occurring explicitly in Eq. (62), those resulting from the differentiation of the A_b^a-s in the numerator of Eq. (62), and that resulting from the differentiation of A_0^0 (the normalization integral for the spin-projected determinant) in the denominator of Eq. (62). In this approach the last term is that which leads to the explicit appearance of the total many-electron energy E in the EHF equations, while the terms of the second type can be collected to form the quantity q_k. The relationship

$$\left.\frac{\partial A_b^a(k_1, k_2, \ldots, k_a)}{\partial\lambda_k}\right|_{\lambda_k^*=\text{const}} = \begin{cases} 0, & k \in \{k_1, k_2, \ldots, k_a\} \\ \lambda_k^* A_{b+1}^{a+1}(k_1, k_2, \ldots, k_a, k), & k \notin \{k_1, k_2, \ldots, k_a\} \end{cases} \tag{65}$$

which one can easily derive based on the definition (53) of the A_b^a-s, explains the similarity in the structure of Eq. (63) for q_k to that of the energy formula (62).

E. Some Transformations of the EHF Equations

The EHF equations derived in the previous sections have a form quite different from that of other SCF-type equations. It seems, therefore, reasonable to discuss how they can be transformed into some more familiar forms. From now on we shall use Dirac's bra and ket vectors which permit presentation of the expressions in a more concise form.

Let us first consider the second term on the left-hand side of Eq. (55), substituting $\lambda_k^* = \langle b_k | a_k \rangle$. It can be transformed as follows (Mayer *et al.*, 1973):

$$\lambda_k^* \hat{h}^b(k) | b_k \rangle = \langle b_k | a_k \rangle \hat{h}^b(k) | b_k \rangle = \hat{h}^b(k) | b_k \rangle \langle b_k | a_k \rangle = \hat{h}^b(k) \hat{P}_k^b | a_k \rangle, \quad (66)$$

where we have introduced the notation

$$\hat{P}_i^d = | d_i \rangle \langle d_i | \quad (67)$$

for the operator of the projection into the one-dimensional subspace spanned by the orbital $|d_i\rangle$ ($d = a$ or b).

While in Eq. (55) this term was considered as the operator $\lambda_k^* \hat{h}^b(k)$ acting on the orbital $|b_k\rangle$, in Eq. (66) we have transformed it into the form in which the operator $\hat{h}^b(k)\hat{P}_k^b$ acts on the orbital $|a_k\rangle$. This gives us the possibility to combine the two terms on the left-hand side of Eq. (55) into a single operator acting on $|a_k\rangle$. Note that the operator $\hat{\mathcal{H}}^a(k)$ is Hermitian as a consequence of Eqs. (52) but neither $\lambda_k^* \hat{h}^b(k)$ nor $\hat{h}^b(k)\hat{P}_k^b$ are Hermitian.

Furthermore, one can see upon inspection that the orbital a_k occurs as a "ket" in each term of Eqs. (59) and (61) for ε_{ki}^{ab} and ε_{kk}^{ab}, respectively, either in the coefficient $\lambda_k^* = \langle b_k | a_k \rangle$ or in the one- and two-electron integrals over the one-electron orbitals. One can, therefore, perform manipulations analogous to those shown in Eq. (66). The first terms of $\varepsilon_{kk}^{ab} | b_k \rangle$ lead just to the adjoint of the operator $\hat{h}^b(k)\hat{P}_k^b$, which can be seen by making use of Eq. (52). Thus one obtains (Mayer *et al.*, 1973):

$$\varepsilon_{kk}^{ab} | b_k \rangle = -\hat{P}_k^b \hat{h}^b(k)^\dagger | a_k \rangle - q_k \hat{P}_k^b | a_k \rangle + A_1^1(k) E \hat{P}_k^b | a_k \rangle \quad (68)$$

and

$$\varepsilon_{ki}^{ab} | b_i \rangle = -\hat{E}_{ki}^{ab} | a_k \rangle \qquad (i \neq k) \quad (69)$$

where the operator $\hat{E}_{ki}^{ab}$ is defined (Mayer and Kondász, 1975) as[26]:

$$\begin{aligned}\hat{E}_{ki}^{ab} = -\, & | b_i \rangle \langle a_i | \, \lambda_i^* \Big[A_1^2(i,k) \hat{H}^N + A_2^2(i,k) \hat{K}_k^{bb} + A_1^2(i,k) \hat{J}_k^{bb} \\ & + \sum_{\substack{j=1 \\ (j \neq i,k)}}^{n} \{ A_2^3(i,j,k) [\lambda_j (\hat{J}_j^{ba} - \hat{K}_j^{ba}) + \hat{K}_j^{bb} + \lambda_j^* (\hat{J}_j^{ab} - \hat{K}_j^{ab})] \\ & - A_1^3(i,j,k) (\hat{K}_j^{aa} - \hat{J}_j^{aa} - \hat{J}_j^{bb}) \} \Big]\end{aligned}$$

---

[26] In Mayer *et al.* (1973) a different, but fully equivalent definition of $\hat{E}_{ki}^{ab}$ was given; a form closely related to (70)–(71) was, however, used in discussing the LCAO case.

$$
\begin{aligned}
&- \hat{P}_i^b \Bigg[-A_1^2(i,\, k)(\hat{H}^N + \hat{J}_k^{bb} - \hat{K}_k^{bb} + \hat{J}_i^{aa} - \hat{K}_i^{aa}) \\
&+ \sum_{\substack{j=1 \\ (j \neq i,k)}}^{n} \{A_2^3(i, j, k)[\lambda_j(\hat{K}_j^{ba} - \hat{J}_j^{ba}) - \lambda_j^* \hat{J}_j^{ab}] \\
&- A_1^3(i,\, j,\, k)(\hat{J}_j^{aa} + \hat{J}_j^{bb} - \hat{K}_j^{aa} - \hat{K}_j^{bb} + \lambda_j^* \hat{K}_j^{ab})\} \Bigg] - p_{ki}^{ab} |b_i\rangle\langle b_k|.
\end{aligned}
\tag{70}
$$

The definition of the coefficients p_{ki}^{ab} is given by

$$
\begin{aligned}
p_{ki}^{ab} = {} & A_1^2(i,\, k)\langle b_i|\, \hat{H}^N\, |b_k\rangle + A_2^2(i,\, k)\lambda_i^* \langle a_i|\, \hat{H}^N\, |b_k\rangle \\
& + A_2^2(i,\, k)[a_i b_i \mid b_k a_i] + A_1^2(i,\, k)[a_i b_i \mid a_i b_k] \\
& + \sum_{\substack{j=1 \\ (j \neq i,k)}}^{n} \{A_2^3(i, j, k)[\lambda_i^*([a_j a_i \mid a_j b_k] - [a_j a_i \mid b_k a_j] + [b_j a_i \mid b_j b_k] \\
& -[a_i b_j \mid b_j b_k]) + [b_i a_j \mid a_j b_k] + \lambda_j^*([a_j b_i \mid b_j b_k] - [b_i a_j \mid b_j b_k]) \\
& + \lambda_i^* \lambda_j [a_i b_j \mid a_j b_k] + \lambda_j([b_j b_i \mid a_j b_k] - [b_j b_i \mid b_k a_j])] \\
& + A_3^3(i, j, k)[\lambda_i^* \lambda_j [b_j a_i \mid a_j b_k] + \lambda_i^* \lambda_j^*([a_j a_i \mid b_j b_k] - [a_j a_i \mid b_k b_j])] \\
& - A_1^3(i, j, k)([b_j b_i \mid b_k b_j] - [b_j b_i \mid b_j b_k] - [a_j b_i \mid a_j b_k])\}.
\end{aligned}
\tag{71}
$$

Performing these transformations for all terms containing the elements of the matrix $\boldsymbol{\varepsilon}^{ab}$, bringing them to the left-hand side, and combining into a single operator all the resulting operators acting on the orbital $|a_k\rangle$, we obtain a convenient form of the EHF equations (Mayer *et al.*, 1973):

$$
\hat{F}^{\alpha}(k)\, |a_k\rangle = \sum_{l=1}^{n} \varepsilon_{kl}^{aa}\, |a_l\rangle \qquad (k = 1, 2, \ldots, n),
\tag{72}
$$

where the operator $\hat{F}^a(k)$ is defined as

$$
\hat{F}^a(k) = \mathscr{H}^a(k) + \hat{h}^b(k)\hat{P}_k^b + \hat{P}_k^b \hat{h}^b(k)^{\dagger} + [q_k - A_1^1(k)E]\hat{P}_k^b + \sum_{\substack{l=1 \\ (l \neq k)}}^{n} \hat{E}_{kl}^{ab}.
\tag{73}
$$

As discussed already in the previous section, the elements of the matrix $\boldsymbol{\varepsilon}^{aa}$ can be considered as undetermined quantities. As in the previous cases, we obtain again the analogous set of n equations for the orbitals $|b_k\rangle$ by the interchanges $a \leftrightarrow b$; $\lambda \leftrightarrow \lambda^*$.

The operator $\hat{F}^a(k)$ defined in Eq. (73) is not Hermitian, because the operators $\hat{E}_{kl}^{ab}$ are not so. (The other terms are either Hermitian or occur together with their adjoints.) However, it is easy to see, that due to the pairing (19b) of the orbitals, one has for the adjoint of the operator $\hat{E}_{kl}^{ab}$ (Mayer *et al.*, 1973):

$$\hat{E}_{kl}^{ab+}|a_k\rangle = 0. \tag{74}$$

This means that the EHF equations (72) remain valid if one replaces $\hat{F}^a(k)$ by the Hermitian operator

$$\hat{F}^a(k)' = \hat{F}^a(k) + \sum_{\substack{l=1 \\ (l\neq k)}}^{n} \hat{E}_{kl}^{ab+} \tag{75}$$

and the elements of the matrix $\boldsymbol{\varepsilon}^{aa}$ also do not change. (We may note that in the equations discussed in Section III,G the operators $\hat{E}_{kl}^{ab+}$ appear without performing any such artificial transformations.)

Mayer *et al.* (1973; Mayer, 1977) transformed the EHF equations also to some other different (but, of course, equivalent) forms as (pseudo)-eigenvalue equations of different non-Hermitian and Hermiticized operators and forms in which the "k-dependence" of the equations was, at least formally, excluded by using the "coupling operator formalism." We shall not discuss here these further possible presentations of the EHF equations; they may be either found in the article by Mayer *et al.* (1973) or obtained from Eqs. (72)–(75) in a quite straightforward manner.

F. The Odd-Electron Case

In the odd-electron case the EHF equations can be derived from the generalized Brillouin theorem exactly in the same manner as was discussed previously for an even number of electrons. The structure of the equations is also very similar. However, in the odd-electron case one must consider explicitly three different types of equations: two for the orbitals a_i and b_i forming pairs of corresponding orbitals and one for the "unpaired" orbital. (None of them can be obtained from another simply by interchanging some symbols, as was the case in the previous sections for an even number of electrons.)

Rather than quote these lengthy equations here (they are given in Mayer, 1974c), we shall only discuss briefly some interrelations which hold between the EHF problems for systems differing by one electron (Mayer, 1977).

Let us consider the doublet state of a system containing $2\nu + 1$ electrons[27] ($S = S_z = \frac{1}{2}$; the number of electrons with spin α and β is

---

[27] In Mayer (1974c) the number of electrons was denoted as $2n + 1$. Here we use the notation ν instead of n, while n is defined as $\nu + 1$, in order to conserve the notations used above for the case of an even number ($2n$) of electrons.

$n = \nu + 1$ and ν, respectively) and the singlet state of another system with an even $(2\nu + 2 = 2n)$ number of electrons $(S = S_z = 0$; the number of electrons with both spins α and β, respectively, is n). In order to avoid any mistake we shall denote here the Sanibel coefficients for the odd-electron case as d_m $(m = 0, 1, 2, \ldots, \nu)$, while the notation c_m $(m = 0, 1, 2, \ldots, n)$ will be conserved for the case of an even number of electrons. Similarly, the quantities A_b^a and B_r^a defined in Eq. (53) will be denoted in the odd-electron case as D_b^a and E_r^a, respectively. (In the odd-electron case n must be replaced by ν in these definitions.) It follows from Eq. (16) that the Sanibel coefficients in these two cases differ only by a factor of 2 (the number of electrons with spin α is in both cases $\mu = n$):

$$d_m = 2c_m \qquad (m = 1, 2, \ldots, \nu). \tag{76}$$

Let us now assume that the wave function in both cases is built up of the same orbitals; of course, in the odd-electron case the number of orbitals filled with spin β is one less. Let the orbital absent in the odd-electron case be b_k. The orbital a_k which was paired with b_k becomes now the "unpaired" one; it will correspond to the orbital a_s in the article by Mayer (1974c).

In the odd-electron case there are $\nu = n - 1$ pairs of orbitals a_i–b_i and, accordingly, one has the same number of the quantities λ_i. For the subscript i we could assume that its value runs from 1 to ν. However, it will facilitate the comparisons if we conserve the notations used in the even-electron case. Thus we assume: $i = 1, 2, \ldots, n$, but $i \neq k$. Accordingly, if we have to perform a summation over the values of i (or an analogous subscript), then the summation running from 1 to ν must be replaced by that running from 1 to n, but excluding the value k. For instance, by applying this consideration to the definition (53b), we trivially find

$$E_r^a(k_1, k_2, \ldots, k_a) = B_r^{a+1}(k_1, k_2, \ldots, k_a, k), \tag{77}$$

where B_r^{a+1} and E_r^a correspond to our two systems with the even and odd number of electrons, respectively. Based on Eqs. (76) and (77) one obtains from definition (53a) (the A_b^a-s are denoted as D_b^a in the odd-electron case) the similar relationship

$$D_b^a(k_1, k_2, \ldots, k_a) = 2A_b^{a+1}(k_1, k_2, \ldots, k_a, k). \tag{78}$$

One may rewrite the energy formula for the odd-electron case [Eq. (9) in Mayer, 1974c] by taking into account that the number of electrons is now $2\nu + 1$ (instead of $2n + 1$), and perform manipulations analogous to those discussed above for all the summation indices. Then, by making use of Eq. (78), the energy of the odd-electron system can be rewritten by using only quantities corresponding to the even-electron case. By comparing the resulting expression with that of the diagonal matrix element ε_{kk}^{bb} corresponding to the latter system with $2n$ electrons,[28] one obtains (Mayer, 1977):

$$\varepsilon_{kk}^{bb} = A_0^1(k)[E_{2n} - E_{2n-1}(b_k)]. \tag{79}$$

Here E_{2n} is the total singlet spin-projected energy for the system with $2n$ electrons, while $E_{2n-1}(b_k)$ is the spin-projected energy for the doublet state of the system obtained by removing the electron from the orbital b_k.

As the energy for a spin-free Hamiltonian evidently does not depend on whether the system has a resulting S_z value of $+\frac{1}{2}$ or $-\frac{1}{2}$, an analogous expression holds when the electron is removed from the orbital a_k:

$$\varepsilon_{kk}^{aa} = A_0^1(k)[E_{2n} - E_{2n-1}(a_k)]. \tag{80}$$

If one transforms Eq. (72) to the form of an eigenvalue equation, then ε_{kk}^{aa} will be its eigenvalue (Mayer *et al.*, 1973). Accordingly, one may consider Eqs. (79) and (80) as a generalization of the Koopmans theorem for the EHF case. (An analogous generalization was obtained also by Goddard for his orbitals). According to definition (53), the coefficient $A_0^1(k)$ is independent of the orbitals a_k or b_k in question. This means that one could divide the equation for the orbital a_k by $A_0^1(k)$, then the resulting new quantity $\varepsilon_{kk}^{aa'} = \varepsilon_{kk}^{aa}/A_0^1(k)$ would correspond even more to the usual notion of the orbital energy. (For an application of this result see Mayer and Kertész, 1976.)

The interrelations between the case of odd and even numbers of electrons have also an important consequence from the practical point of view: they permit calculations for the doublet states of odd-electron systems by using a computer program set up for the case of an even number of electrons. (Thus one may avoid the explicit programming of the three different types of equations appearing in the odd-electron case.) One has only to add to the odd-electron system an additional electron located on a dummy center at infinity and perform the EHF calculation for the singlet state of the resulting system with an even number of electrons. In fact, the energy which is necessary to remove the electron from the dummy orbital is always equal to the ionization potential I_{dummy} of the latter and does not depend on the orbitals corresponding to the original system. Accordingly

$$E_{2n} = E_{2n-1} + I_{\text{dummy}}. \tag{81}$$

---

[28] This expression can be obtained from Eq. (31) of Mayer *et al.* (1973) by performing the interchanges $a \leftrightarrow b$, $\lambda \leftrightarrow \lambda^*$ already discussed in the foregoing sections.

Therefore, if one minimizes the energy E_{2n} of the system containing the additional electron at infinite distance, then this involves also the minimizing of the energy E_{2n-1} of the original system.[29]

It was meaningful to transform the expression for the spin density (Ukrainsky, 1972; Harris, 1966) to a form containing only quantities corresponding to the system with the additional electron. Assuming that this electron is on the orbital b_n, one obtains after a lengthy but simple algebra (I. Mayer and T. Pollák, unpublished):

$$\begin{aligned}\rho_s(\mathbf{r}) = {} & (3A_0^0)^{-1}\{[A_0^0 + 2A_1^1(n)]|a_n(\mathbf{r})|^2 \\ & + \sum_{j=1}^{n-1} ([A_0^2(j, n) + 2A_1^2(j, n)]|a_j(\mathbf{r})|^2 - A_0^1(j)|b_j(\mathbf{r})|^2 \\ & - 2A_1^2(j, n) \operatorname{Re} [\lambda_j^* b_j(\mathbf{r}) a_j(\mathbf{r})])\}. \qquad (82)\end{aligned}$$

G. The EHF Equations for the Successive Optimization of the Orbitals

The EHF equations discussed so far (we may call them "basic" EHF equations) are related to the final (converged) orbitals and give the necessary and sufficient conditions for the energy to be stationary. Their solutions are not evident, however. For instance, the simple iterative scheme of solution based on Eq. (72) transformed to an eigenvalue problem could diverge even when started from the closest vicinity of the solution (Mayer, 1974a). We needed, therefore, an algorithm of solution with reliable convergence properties. For this reason we worked out the EHF equations necessary to realize a procedure of the successive optimization of the individual orbitals, which is necessarily convergent.[30] In this method one recalculates the orbitals cyclically, one at a time, going through all the orbitals in each cycle. When the given orbital is updated, all the others are kept fixed and one must determine the orbital in question in such a way as to obtain a minimum possible total energy (within the limitation given by the other fixed orbitals). In this manner one obtains a decreasing (or at

---

[29] One can obtain (Mayer, 1977) the relationship (81) by performing an analysis of the energy formula for the system with the additional electron, taking into account that all the interactions with the dummy orbital are zero, as is the corresponding λ value. In practice—in accordance with the correct dissociation properties of the EHF method—it is not necessary to put the dummy center strictly into the infinity. It is enough, if it is on a sufficiently large distance so that all the interactions, except the long-range Coulomb ones, are negligible. The latter can then exactly be compensated by attributing a nuclear charge of unity to the dummy center.

[30] This approach to solving different SCF-type problems was repeatedly rediscovered in the literature. For recent discussions, see, e.g., Lunell (1970, 1972b, 1973), Mayer (1973a, 1974a), and Mayer and Kondász (1975).

least, not increasing) sequence of the energy values. This sequence is, in principle, infinite; it is bounded from below (as follows from the variational principle), therefore, it does converge to a definite minimum energy value.

Even in the UHF case the method of successive optimization of the orbitals requires the use of equations having a form slightly different from the usual one (Mayer, 1973a). In the EHF case a number of new terms appear in the equations which are absent from the "basic" EHF equations (and vanish, accordingly, for the converged paired solutions). These new terms are connected with the following aspects of pairing of the orbitals. If one has obtained the orbitals ensuring a stationary energy, then one can always subject them to the transformations providing the orthogonality and pairing, and thus investigate the conditions of stationary energy for these orthonormalized and paired orbitals. (These conditions are given by the "basic" EHF equations discussed in the previous sections.) If, however, one does not yet have a convergent solution, fixes all the orbitals but one, and optimizes this individual orbital according to the scheme discussed above, then one usually obtains an orbital which is not paired. It follows from the properties of the determinants, that this orbital can be assumed (with no loss of generality) to be orthogonal to all the other orbitals filled with the same spin. The pairing can be performed, however, only after the new orbital has been actually obtained. The pairing will affect, in general, all the orbitals of both sets a_i and b_i; this, however, does not influence the convergence properties of the algorithm. (The pairing must be performed in each step to conserve the possibility of using manageable equations.) As the convergence is approached, the new orbital obtained in the given step will be closer and closer to the old paired one, and finally we get paired convergent solutions.

Thus, in order to realize the algorithm discussed above, we shall solve a specific variational problem in which the energy is required to be stationary under the variations only of the orbital a_k which is actually optimized. We may assume the fulfilment of the orthogonality conditions (19a), but (19b) must be replaced by the weaker conditions

$$\langle a_i | b_j \rangle = \lambda_i \, \delta_{ij} \qquad (i \neq k). \tag{83}$$

In the spirit of the derivation given in Section III,B, one easily obtains that this variational problem can be formulated in the specific form of the generalized Brillouin theorem:

$$\langle \Psi_1(a_k \rightarrow c) | \hat{H} - E | \Psi \rangle = 0, \tag{84}$$

i.e., only the single excitations from a_k have to be considered.

Expanding (84) is terms of one-electron orbitals by taking into account (19a) and (83), we obtain the EHF equation for the orbital a_k in the same "overall" form as described in Section III,C; the elements of matrices ε are, of course, different from those discussed in Section III,D, since they contain a number of terms proportional to the integrals $\langle b_i|a_k\rangle$, which now are nonvanishing. We shall give the form of this equation transformed to an eigenvalue problem of an Hermitian operator (Mayer and Kondász, 1975) which we use in the actual calculations:

$$\left(1 - \sum_{\substack{i=1\\(i\neq k)}}^{n} \hat{P}_i^a\right) \hat{F}^a(k) \left(1 - \sum_{\substack{i=1\\(i\neq k)}}^{n} \hat{P}_i^a\right) |a_k\rangle = \varepsilon_{kk}^{aa}|a_k\rangle. \tag{85}$$

Here the operator $\hat{F}^a(k)$ is defined by

$$\begin{aligned}\hat{F}^a(k) = {}& \hat{\mathscr{H}}^{a}(k) + \hat{\hbar}^{b}(k)\hat{P}_k^b + \hat{P}_k^b\hat{\hbar}^{b}(k)^+ + [q_k - A_1^1(k)E]\hat{P}_k^b \\ & + \sum_{\substack{i=1\\(i\neq k)}}^{n} \{\hat{E}_{ki}^{ab} + \hat{E}_{ki}^{ab+} + [r_{ki}^{ab} - A_1^2(i, k)E]\hat{P}_i^b\} \\ & + \sum_{\substack{i,j=1\\(i\neq j;\, i,j\neq k)}}^{n} s_{kij}^{abb}|b_i\rangle\langle b_j|,\end{aligned} \tag{86}$$

whereas the coefficients r_{ki}^{ab} and s_{kij}^{abb} are given by

$$\begin{aligned}r_{ki}^{ab} = {}& A_1^2(i, k)(\langle a_i|\hat{H}^N|a_i\rangle + \langle b_k|\hat{H}^N|b_k\rangle + [a_ib_k|a_ib_k]) \\ & + A_2^2(i, k)[a_ib_k|b_ka_i] + \sum_{\substack{j=1\\(j\neq i,k)}}^{n} \{A_1^3(i, j, k)(\langle a_j|\hat{H}^N|a_j\rangle \\ & + \langle b_j|\hat{H}^N|b_j\rangle + [a_jb_j|a_jb_j] + [a_jb_k|a_jb_k] - [a_ja_i|a_ia_j] \\ & + [a_ja_i|a_ja_i] + [a_ib_j|a_ib_j] + [b_jb_k|b_jb_k] - [b_jb_k|b_kb_j]) \\ & + A_2^3(i, j, k)[\lambda_j^*\langle a_j|\hat{H}^N|b_j\rangle + \lambda_j\langle b_j|\hat{H}^N|a_j\rangle + \lambda_j^*([a_ja_i|b_ja_i] \\ & - [a_ja_i|a_ib_j] + [a_jb_k|b_jb_k] - [a_jb_k|b_kb_i]) + \lambda_j([a_ib_j|a_ia_j] \\ & - [b_ja_i|a_ia_j] + [b_kb_j|b_ka_j] - [b_kb_j|a_jb_k]) + [a_jb_k|b_ka_j] \\ & + [b_ja_i|a_ib_j] + [a_jb_j|b_ja_j]]\} + \sum_{\substack{j<l\\(j,l\neq i,k)}} \{A_1^4(i, j, k, l)([a_ja_l|a_ja_l] \\ & - [a_ja_l|a_la_j] + [b_jb_l|b_jb_l] - [b_jb_l|b_lb_j]) \\ & + A_3^4(i, j, k, l)[\lambda_j\lambda_l([b_jb_l|a_ja_l] - [b_jb_l|a_la_j]) + \lambda_j^*\lambda_l^*([a_ja_l|b_jb_l] \\ & - [a_ja_l|b_lb_j])]\} + \sum_{\substack{j,l\\(j\neq l;\, j,l\neq i,k)}} \{A_1^4(i, j, k, l)[a_jb_l|a_jb_l] \\ & + A_2^4(i, j, k, l)[\lambda_l([b_jb_l|b_ja_l] - [b_jb_l|a_lb_j] + [a_jb_l|a_ja_l] \\ & - [a_jb_l|a_la_j]) + \lambda_l^*([a_lb_j|b_lb_j] - [a_lb_j|b_jb_l] + [a_ja_l|a_jb_l] \\ & - [a_ja_l|b_la_j]) + [a_lb_j|b_ja_l] + \lambda_j\lambda_l^*[a_lb_j|a_jb_l]] \\ & + A_3^4(i, j, k, l)\lambda_l\lambda_j^*[b_la_j|a_lb_j]\},\end{aligned} \tag{87}$$

$$\begin{aligned}
s_{kij}^{abb} = {} & A_1^3(i, j, k)(-\langle b_i|\hat{H}^N|b_j\rangle - [a_ib_i|a_ib_j] - [a_jb_i|a_jb_j] - [b_ib_k|b_jb_k] \\
& + [b_ib_k|b_kb_j]) + A_2^3(i, j, k)[\lambda_i^*\lambda_j\langle a_i|\hat{H}^N|a_j\rangle - \lambda_i^*\langle a_i|\hat{H}^N|b_j\rangle \\
& - \lambda_j\langle b_i|\hat{H}^N|a_j\rangle + \lambda_i^*([a_ib_k|b_kb_j] + [a_ja_i|b_ja_j] - [a_ja_i|a_jb_j] \\
& - [b_ka_i|b_kb_j]) + \lambda_j([b_ib_k|b_ka_j] + [a_ib_i|a_ja_i] - [a_ib_i|a_ia_j] \\
& - [b_ib_k|a_jb_k]) + \lambda_j\lambda_i^*[b_ka_i|b_ka_j] - [a_ib_i|b_ja_i] - [a_jb_i|b_ja_j]] \\
& + A_3^3(i, j, k)\lambda_j\lambda_i^*[a_ib_k|b_ka_j] + \sum_{\substack{l=1 \\ (l\neq i,j,k)}}^{n} \{A_1^4(i, j, k, l)(-[a_lb_i|a_lb_j] \\
& + [b_ib_l|b_lb_j] - [b_ib_l|b_jb_l]) + A_2^4(i, j, k, l)[\lambda_j([a_lb_i|a_ja_l] \\
& - [a_lb_i|a_la_j] + [b_ib_l|b_la_j] - [b_ib_l|a_jb_l]) + \lambda_i^*([a_la_i|b_ja_l] \\
& - [a_la_i|a_lb_j] + [b_la_i|b_jb_l] - [b_la_i|b_lb_j]) + \lambda_l^*([a_lb_i|b_jb_l] \\
& - [a_lb_i|b_lb_j]) + \lambda_l([b_ib_l|a_lb_j] - [b_ib_l|b_ja_l]) - [a_lb_i|b_ja_l] \\
& - \lambda_l^*\lambda_j[a_lb_i|a_jb_l] - \lambda_l\lambda_i^*[b_la_i|b_ja_l] + \lambda_i^*\lambda_j([b_la_i|b_la_j] \\
& - [a_la_i|a_ja_l] + [a_la_i|a_la_j])] + A_3^4(i, j, k, l)[\lambda_j\lambda_l\lambda_i^*([b_la_i|a_la_j] \\
& - [b_la_i|a_ja_l]) + \lambda_j\lambda_l^*\lambda_i^*([a_la_i|b_la_j] - [a_la_i|a_jb_l]) \\
& + \lambda_l\lambda_j([b_ib_l|a_la_j] - [b_ib_l|a_ja_l]) + \lambda_i^*\lambda_l^*([a_la_i|b_jb_l] \\
& - [a_la_i|b_lb_j]) - \lambda_l\lambda_i^*[b_la_i|a_lb_j] - \lambda_j\lambda_l^*[b_ia_l|a_jb_l] + \lambda_j\lambda_i^*[b_la_i|a_jb_l]]\}.
\end{aligned} \tag{88}$$

The other terms were defined in the previous sections. The equations for the orbital b_k can again be obtained by the interchanges $a \leftrightarrow b$, $\lambda \leftrightarrow \lambda^*$.

In accordance with the previous discussion, the terms

$$\sum_{\substack{i=1 \\ (i\neq k)}}^{n} \{\hat{E}_{ki}^{ab+} + [r_{ki}^{ab} - A_1^2(i, k)E]\hat{P}_i^b\} + \sum_{\substack{i,j=1 \\ (i\neq j;\, i,j\neq k)}}^{n} s_{kij}^{abb}|b_i\rangle\langle b_j|, \tag{89}$$

which are present in Eq. (86) but are absent in Eq. (73), vanish for the paired convergent solution.

It could prove to be troublesome that the EHF equation for the orbital a_k contains the total many-electron energy E, which is, however, not known until the solution has actually been obtained, since it depends also on this orbital. It follows, however, from a slightly involved analysis done by the present author (Mayer, 1974a) that one may freely substitute into the EHF equation the value of the energy calculated by the orbitals of the previous iteration step, because this will not destroy the good convergence properties of the algorithm. Moreover, some further simplifications (Mayer and Kertész, 1975; Mayer, 1978) may often be introduced into the procedure without influencing the convergence (see Section IV,A).

The convergence properties of the method obtained in performing the EHF calculations for the π-electron system of butadiene are illustrated on Figs. 3 and 4.

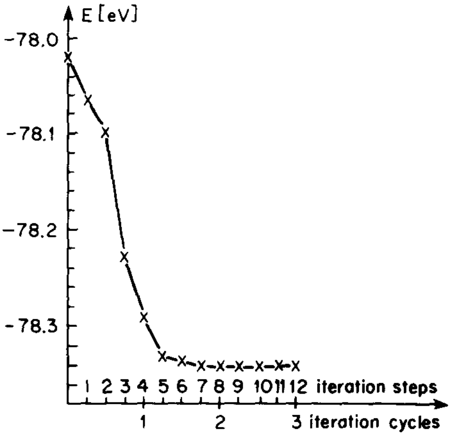


Fig. 3. Convergence of the EHF energy for the PPP model of butadiene. The EHF procedure was started from the UHF+SP wave function. (Each iteration cycle consists of four iteration steps in which a single orbital is optimized separately.) From Mayer (1974a).

IV. Some Applications of the EHF Method

A. Some Numerical Aspects and Results

Almost all authors dealing with the EHF variational problem (see Section II) performed actual calculations by using their approach (e.g., Shull and Löwdin 1956a,b, 1959; Hurst *et al.*, 1958; Kotani *et al.*, 1960; Gos-

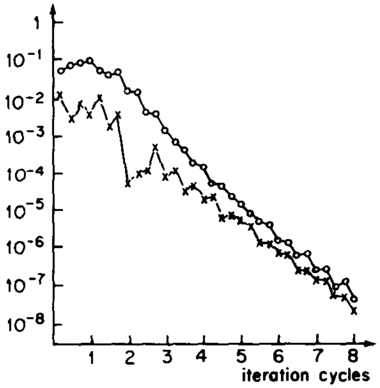


Fig. 4. Convergence of the EHF procedure for butadiene started from the UHF+SP wave function. (–o–o–) the maximum change of the *P* matrix elements in a whole iteration cycle (last four iteration steps); (–x–x–) the deviation from the pairing of the new orbital obtained in the given iteration step before the pairing of the orbitals is reconstructed. (Note the logarithmic scale.) From Mayer (1974a).

cinski and Calais, 1965; Chong, 1966; Lefebvre and Smeyers, 1967; Smeyers and Pinto Suarez, 1968; Smeyers and Doreste-Suarez, 1973; Smeyers and Delgado-Barrio, 1974; Kaldor, 1968a–d; Kaldor *et al.*, 1968; Harriman, 1967; Sando and Harriman, 1967; Harriman and Sando, 1968; Goddard, 1967c, 1968b,e, 1969; Gallup, 1969b; Burden, 1972; Gorlov and Ukrainsky, 1973, 1974; Brown and Larsson, 1973; Rosenberg and Martino, 1975; Klimo and Tiňo, 1976, 1977, 1978; V. Klimo and J. Tiňo, private communication, 1979; Tiňo *et al.*, 1979; Lunell, 1979; Mayer and Kertész, 1975, 1976; Mayer, 1974a, 1975, 1976, 1977, 1978) at the semiempirical or *ab initio* levels. As the systems studied and the bases or parametrizations were different, a detailed comparative analysis of the numerical results obtained is not always possible and would require a great amount of space anyway. For this reason we will not discuss this here. Instead, we will only mention the most important aspects connected with the computer realization of the EHF equations derived in Section III and discuss quite briefly the results of some calculations performed by the present author. This will give an idea of the possibilities and limitations of the method; our conclusions are completely in line with the findings of other authors.

1. *Computer Realization*

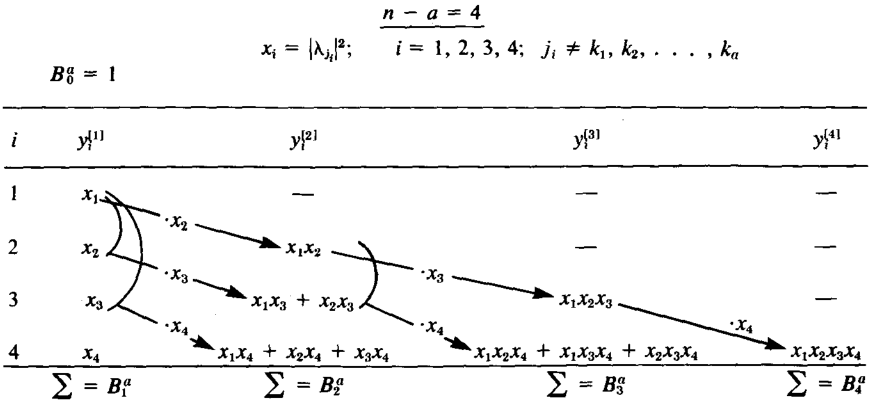
Based on the discussion given by Mayer *et al.* (1973; Mayer, 1974b) the EHF equations (Section III,G) can be transformed to the LCAO form in a quite straightforward manner. In programming it was convenient to deal directly with the matrices and vectors corresponding to the different operators and orbitals, respectively, without rewriting the formulas explicitly in terms of matrix elements and individual LCAO components [cf. Eqs. (68) and (69) in Mayer *et al.*, 1973].

One first forms the matrices $\mathbf{P}_j^{de} = \mathbf{c}_j^d \mathbf{c}_j^{e+}$ ($d, e = a$ or b; $\mathbf{c}_j^d$ is the column vector containing the LCAO coefficients of the orbital d_j) and then uses them to build up the matrices $\mathbf{J}^{de}$ and $\mathbf{K}^{de}$ corresponding to the operators $\hat{J}^{de}$ and $\hat{K}^{de}$. These matrices (stored on the "background" devices) are then utilized in calculating the one- and two-electron integrals over the molecular orbitals a_i, b_i, occurring in the equations and in setting up the final matrices to be diagonalized. This organization (Mayer, 1978) permits one to use the list of two-electron integrals only once in a given iteration step.[31]

The fast and accurate calculation of the coefficients A_b^a defined in Eqs. (53) is of central importance in realizing the EHF method. In order to determine the A_b^a-s one has to calculate the quantities B_r^a, which are,

---

[31] In the case of semiempirical integral approximations using the ZDO assumption (PPP and CNDO methods) it was possible to use programs of much simpler organization.



$$A_b^a = \sum_{m=b}^{n-a+b} (-1)^m c_m B_{m-b}^a = \sum_{q=0}^{4} (-1)^{q+b} c_{q+b} B_q^a \qquad \text{if } n - a = 4$$

Fig. 5. Illustration of the scheme of calculating the coefficients A_b^a for the case of four values of $x_i = |\lambda_i|^2$. From Mayer (1977).

according to Eq. (53b) "symmetric sums" of the $|\lambda_i|^2$ values. One may find different recursion schemes for calculating the B_r^a-s (e.g., Harriman, 1963a); the present author worked out a simple method in which an auxiliary vector is cyclically recalculated[32] (Mayer, 1974d; Mayer and Kertész, 1975). This scheme is illustrated in Fig. 5 for the case of four $x_i = |\lambda_i|^2$ values. Moreover, it is sufficient to use even this fast algorithm to find only a limited number of A_b^a-s; the others can then be obtained by some very simple recursion formulas (Mayer, 1978) derived from Eq. (54).

The EHF equations for the successive optimization of the individual orbitals (Section III,G) are rather complicated. Two types of simplifications were found to be useful in reducing the amount of necessary computations. The first is based on the observation that not all of the additional terms (89), absent from the "basic" EHF equations, are of the same importance; one may usually drop the most complicated ones without significantly influencing the convergence. The calculation of the remaining terms, which were found to be of upmost importance, is rather easy (Mayer and Kertész, 1975). This simplification was found to be especially useful when a simpler programming scheme was used in performing semiempirical calculations.

---

[32] A proof by induction for this algorithm is given in the Appendix of Mayer (1974d). Unfortunately, Eq. (A7) there contains misprints, which, however, may be easily found upon inspection.

TABLE I

COMPARISON OF DIFFERENT METHODS FOR THE PPP MODEL OF THE π-ELECTRONS IN BUTADIENE[a]

Wave function	Total π-electron energy (eV)	Part of correlation energy taken into account (%)	Overlap with exact (full CI) wave function[b]	λ values for corresponding orbitals
RHF	−77.09700	0	0.92954	1; 1
UHF	−77.38789	21.4	0.80500	0.9284; 0.6611
EHF local minimum	−77.82041	53.1	0.93696	0.7181; −0.7181
UHF + SP	−78.01878	67.7	0.97953	(as for UHF)
EHF	−78.34163	91.4	0.99085	0.7749; −0.5114
Full CI	−78.45836	100	1	—

[a] From Mayer (1974a).

[b] The projected wave functions are renormalized to unity.

Another simplification is most useful in the *ab initio* case, since it permits much less frequent use of the list of two-electron integrals and recalculation of the auxiliary matrices and other quantities. In this case one determines the corrections to all the orbitals a_i (or b_i) simultaneously, i.e., does not take into account the corrections obtained for the first orbital when recalculating the second one, etc. If the starting vector is not extremely poor, this simplification almost does not increase the number of full iteration cycles (consisting in recalculating all the orbitals) while each cycle becomes much faster (Mayer, 1978). Often it is useful to combine the different versions of the algorithm in each calculation.

2. *Some π-Electron Calculations*

Table I summarizes some results obtained by different one-electron methods for the π-electron model of butadiene in the PPP integral approximation[33] (Mayer, 1974a). In these and some other (Mayer, 1975) four-electron calculations performed at the semiempirical (PPP, CNDO)

---

[33] The details (parametrizations, etc.) of the calculations shown on the tables were given in Mayer (1974a, 1975) and Mayer and Kertész (1975). For the constant e^2 in the Mataga–Nishimoto formula a value of 14.4117 was used, assuming the distances being measured in angstroms and the γ values in electron volts. A small deviation in this value was the source of some minor differences between our RHF energies and those obtained by Lengsfield and Schug (1978).

level we were able to compare the results given by the different one-electron approximations also with those of the full CI calculations giving the exact solutions for the model Hamiltonians considered.

For butadiene (Table I) the UHF method permits account to be taken of 21.4% of the correlation energy, but the UHF wave function may be considered much worse than the RHF one: it has a much lower overlap with the exact ground-state CI wave function. The UHF energy improvement with respect to RHF results from a compromise: the lowering of the weight of the highest CI eigenfunction in the determinantal wave function is obtained with the result that there appears a large overlap 0.5401 (29.2% weight) with the lowest (−76.634 eV)[34] triplet eigenfunction.

If one performs the spin projection of the UHF wave function, then the triplet and quintet components disappear and the UHF + SP method gives a great improvement with respect to UHF (and RHF). A further significant improvement is given by the EHF method, as the projected wave function itself is optimized variationally. Thus for this system the EHF method permits account to be taken of 91.4% of the correlation energy and gives a wave function which is very close to the exact full CI one.

In both UHF and EHF cases the one-electron orbitals a_i and b_i, forming pairs of corresponding orbitals, are the mirror images of each other with respect to the reflection at the center of the molecule (Mayer, 1974a). A similar situation was observed for all linear polyene molecules (Mayer and Kertész, 1975, 1976). Accordingly, the wave function corresponds to the second case discussed by Popov and Gerratt (see Section II,E). For butadiene we found also the EHF wave function corresponding to the first case discussed in Section II,E, for which not only the many-electron wave function but also the individual one-electron orbitals are symmetry-adapted, similarly to the RHF ones. For this EHF solution, however, the energy is much higher, as the number of effectively free variational parameters is lower, so it represents only a local minimum (Mayer, 1974a).

In Table II we compare the above results with those obtained by using a different parametrization (Mayer, 1975) for the off-diagonal core integrals. The qualitative tendencies are the same in both cases (cf. also with the results of Table VI in Section V for the Goeppert-Mayer–Sklar integral approximation). The differences in percentages, however, indicate that the energy improvement which can be obtained by the UHF and UHF + SP methods is rather system (parametrization)-dependent. At the same time the EHF results are not much affected by the changes in parametrization.

[34] This figure was misprinted in Mayer (1974a).

TABLE II

COMPARISON OF THE RESULTS OBTAINED FOR BUTADIENE BY USING TWO DIFFERENT PARAMETRIZATIONS IN THE PPP INTEGRAL APPROXIMATION SCHEME[a,b]

	Part of correlation energy taken into account (%)		Overlap with exact (full CI) wave function	
Method	Parametrization "A"[a]	Parametrization "B"[b]	Parametrization "A"[a]	Parametrization "B"[b]
RHF	0	0	0.930	0.941
UHF	21.4	8.3	0.805	0.856
UHF + SP	67.7	47.6	0.980	0.971
EHF	91.4	86.3	0.991	0.987

[a] From Mayer (1974a).
[b] From Mayer (1975).

We shall also discuss the case of cyclobutadiene, for which the performance of the EHF method is extremely good; see Table III (Mayer, 1975). Quite similar results were obtained with another parametrization, too (Mayer and Kertész, 1975). Figure 6 shows that for cyclobutadiene one has a well-pronounced symmetry effect: as the ratio between the sides of the rectangle approaches unity, the deviation of the EHF wave function from the full CI one (the value of $1 - \langle\Psi_{\mathrm{EHF}}|\Psi_{\mathrm{CI}}\rangle$) falls a few orders of magnitude. At the configuration of a regular square, which is suggested by

TABLE III

COMPARISON OF DIFFERENT METHODS FOR THE PPP MODEL OF THE π-ELECTRONS IN CYCLOBUTADIENE[a]

Method	Error of total π-electron energy (eV)	Overlap with exact (full CI) wave function
RHF[b]	2.389109	0.667
UHF	0.532576	0.683
UHF + SP	0.200127	0.993
EHF	0.000063	0.9999986
Full CI	0	1

[a] From Mayer (1975).
[b] Single-determinant wave function with doubly filled orbitals and no spatial symmetry; it corresponds to only one of the two "Kekulé structures" ("symmetry dilemma").

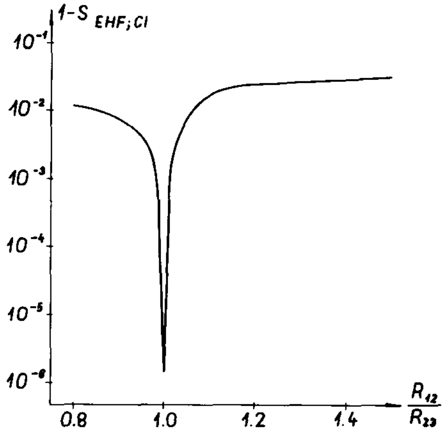
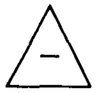


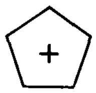
Fig. 6. Deviation from unity of the overlap between the EHF and full CI wave functions for the PPP model of cyclobutadiene as a function of the ratio between the sides of the rectangle. (Note the logarithmic scale.) From Mayer (1975).

recent experimental results, this measure of the error in the EHF wave function becomes only 1.4×10^{-6}, and it was not easy to establish (Mayer, 1975) that this small difference between the wave functions is not due only to the round-off errors.

A similar situation exists for the other simplest antiaromatics also with four π-electrons: cyclopropenyl anion (I) and cyclopentadienyl cation (II)



I



II

However, the exact π-electron full CI ground state of these systems is triplet, while the lowest full CI singlet eigenvalue is degenerated, indicating that the lowest singlet states of these ions perhaps have a deformed geometry. On the contrary, the full CI ground state of cyclobutadiene is singlet, though not total-symmetric: the wave function changes by a phase factor under the rotations around the fourfold axis.

The excellent results given by the EHF method for these antiaromatics (as compared with the very poor RHF ones) indicate that one cannot even qualitatively imagine the π-electron structure of these systems by using the picture of doubly filled orbitals. Owing to the antiferromagnetic character of the spin-coupling scheme inherent in the EHF wave function (see Section II,A) these molecules may be considered as the simplest systems of the antiferromagnetic type.

Table IV shows the comparison of the energy improvement per electron pair given by different DODS methods with respect to the RHF energy for the case of some π-electron molecules with 2–10 electrons (see Figs. 1 and 2 in Section II,F for longer polyene chains). We may again notice that the results given by the EHF method are not only better but also much less system-dependent than those obtained by the UHF and UHF + SP procedures. (Note the two orders of magnitude difference between the UHF results for benzene and decapentaene.) We will not discuss these results here in any detail (see Mayer and Kertész, 1975) and only note that in the case of the nonalternating fulvene molecule two EHF solutions were obtained: one with an appropriate spatial symmetry and another was a broken-symmetry solution with a slightly lower energy.

Concerning the radicals, we shall only mention a few results of our PPP calculations (I. Mayer and T. Pollák, unpublished). The π-electron spin densities calculated by the EHF method for the allyl radical are very close to those obtained by full CI, and agree very well also with the "experimental" values determined by using the well-known McConnell

TABLE IV

COMPARISON OF DIFFERENT DODS METHODS FOR SOME π-ELECTRON SYSTEMS IN THE PPP INTEGRAL APPROXIMATION[a]

	Energy improvement per electron pair with respect to RHF solution (eV)			
Molecule	UHF	UHF + SP	EHF	EHF $\lvert\lambda_i\rvert$ values
Ethylene	0.040	0.309	0.720[b]	0.581
Butadiene	0.146	0.461	0.623	0.775; 0.511
Hexatriene	0.189	0.459	0.557	0.837; 0.744; 0.454
Octatetraene	0.210	0.441	0.511	0.868; 0.822; 0.705; 0.406
Decapentaene	0.222	0.427	0.479	0.886; 0.860; 0.799; 0.667; 0.365
Benzene	0.002	0.072	0.406	0.860; 0.698; 0.698
Styrene	0.038	0.225	0.379	0.894; 0.809; 0.732; 0.618
Fulvene	0.047	0.189	0.366	0.864; 0.779; 0.556
	0.065	0.259	0.407	0.847; 0.786; 0.522
s-Triazine	0.045	0.308	0.524	0.835; 0.663; 0.663

[a] From Mayer and Kertész (1975).

[b] Full CI coincides with EHF.

equation. No such happy situation holds for larger systems. For the benzyl radical the EHF method takes into account about 50% of the correlation energy known from the literature, the spin densities, however (though they give a very good correlation with the full CI values), are strongly "overpolarized" (have too large absolute values), which is the general tendency observed in the literature for the EHF spin densities. For the polyene radicals, when the length of the chain increases, the ratio of the projected and unprojected spin densities quickly approaches the value $\frac{1}{3}$ predicted by the theory (Ukrainsky, 1972a) for the $N \rightarrow \infty$ limit.

3. *A Comparison of Potential Curves*

The potential curves obtained for the B–H molecule at the *ab initio* level by using the different one-electron methods are shown on Fig. 7 (Mayer, 1978). (A double zeta-type basis of Gaussian lobe functions was used.) These curves may well illustrate the discussion of different one-electron methods given in the Introduction (Section I,B).

At the interatomic distances lower than ~3.4 a.u. we have only two distinct (RHF and EHF) curves, as the UHF method gives no solutions differing from the RHF ones. The position of the minima on the two curves agree well with each other and with the experimental equilibrium distance. Thus the RHF method gives a reasonably good description of the potential curve at the equilibrium and smaller distances, but at the larger interatomic distances the RHF energy rapidly increases and from ~7 a.u. on the RHF procedure fails to converge.

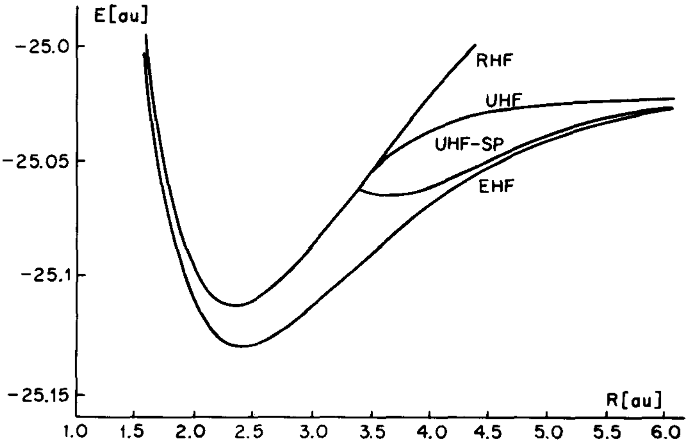


Fig. 7. Potential curves of BH molecule obtained by using different one-electron methods and a 4s/2p Gaussian lobe basis set. From Mayer, (1978).

From about 3.4 a.u. the UHF curve departs *quite smoothly* from the RHF one and then approaches a correct asymptotic energy value too quickly.[35]

The behavior of the UHF + SP method is rather pathological. After the separation of the UHF and RHF curves occurs, the UHF + SP energies, as expected, are between the UHF and EHF ones. The UHF + SP curve, however, departs sharply from the RHF one and exhibits spurious extrema. These artificial extrema on the UHF + SP curve are related to the limitations due to the nonvariational character of this procedure and show that the UHF + SP method is not adequate to describe even the qualitative features of potential curves. (Its asymptotic behavior seems, however, rather good. For a further discussion, see Mayer, 1978.)

The EHF energies are not only lower than those given by the other one-electron methods, but the EHF potential curve has a well-balanced shape and is correct, not only qualitatively but at least, semiquantitatively. The value of 67.5 kcal/mole obtained for the dissociation energy of BH (the depth of the minimum) is somewhat smaller than the most probable value 83 kcal/mole obtained by a combined experimental-theoretical estimation (Jones *et al.*, 1967). If one polarization function is added to the basis on each atom, then the calculated dissociation energy increases to 71 kcal/mole; one may assume that the use of a larger and/or optimized basis would lead to better (possibly quite good) agreement with the experiment.

A comparison of the results obtained by using some different basis sets for this molecule (Table V) shows that the EHF method is insensitive to the peculiarities of the basis, only its overall character is of importance. Though the absolute values of the energies were strongly different, we obtained almost exactly parallel potential curves for minimum basis sets of quite different quality, and similarly, for different basis sets which were of double zeta type at least in the valence shell.

As the interatomic distance increases, the potential curves given by the different DODS methods (UHF, UHF + SP, EHF) tend almost exactly to the same limiting value which is practically equal to the sum of the RHF energies of the two atoms. This shows that neither of these methods is able to describe the atomic correlation of the boron atom. This is in line with the general observations (e.g., Schaefer, 1972) and is immediately connected (Lunell, 1968; Mayer, 1978) with the spin-coupling scheme (cf. Section II,A) inherent in the EHF wave function: the resulting EHF multiplicity is obtained from coupling two subsystems of electrons

---

[35] According to an analytical study of a simple model problem, the UHF energy and its first derivative are continuous and coincide with the corresponding RHF values at the point where the departure of the UHF curve from the RHF one takes place. The second derivative of the UHF energy is, however, discontinuous at this point; the UHF + SP method has a discontinuity already for the first derivative of the energy.

TABLE V

BASIS-SET DEPENDENCE OF THE BH DISSOCIATION ENERGY CALCULATED BY THE EHF METHOD[a]

Basis set	Calculated BH dissociation energy (kcal/mole)
Minimum; STO/3G[b]	96.5
Minimum; STO/6G[b]	95.9
"Split shell" (4/3 + 1G)[b]	67.9
Dunning's 4s/2p	67.0
Gaussian lobe 4s/2p	67.5[c]

[a] I. Mayer, A. Ažman, and J. Koller, unpublished results.

[b] Internal basis sets generated by the "GAUSSIAN 70" program system.

[c] From Mayer (1978).

of spins α and β, respectively, each being in the state of maximum multiplicity. At the same time, e.g., two 1s electrons of an atom form a singlet subsystem to a good approximation. The effective singlet coupling for a given pair of electrons may be achieved in the EHF case either if their orbitals are not considerably "split" or if all the other orbitals are practically doubly filled. As a consequence, for systems which can be considered as consisting of such pairs of singlet-coupled electrons, one expects that only one orbital pair will be significantly split, while the other electrons will occupy almost identical pairs of orbitals.

According to the above results, the EHF method takes reasonably well into account the "extra" molecular correlation, whereas it does not describe at all that part of the correlation energy which has atomic character even in the molecule. This leads to a relatively good description of formation and breaking of a chemical bond. It is expected that this good behavior is not restricted to the case of such simple two-electron σ-bonds for which the EHF method may be considered as a generalization of Weinbaum's approach to the H_2 molecule.[36] We shall return to this question in Section V.

B. Application to the Superexchange

Magnetic ion pairs in a crystal lattice which are too far apart to have a considerable direct exchange interaction may still exhibit a correlation of usually antiferromagnetic type between their spin alignments due to the

---

[36] Specific problems may appear if there are more than two atoms (see, e.g., Rosenberg and Martino, 1975); in this case the existence of different EHF solutions may also become more dangerous.

interaction *via* the perturbation of the nonmagnetic environment. This phenomenon is called superexchange and represents an important field of experimental and theoretical investigations in solid-state physics and mangetochemistry. We cannot survey here the different theoretical approaches used to treat the problem but refer only to the basic work of Anderson (1963) and the recent reviews of Newman (1977) and van Kalkeren *et al.* (1979) (see also, e.g., Hay *et al.*, 1975; Harcourt, 1976; Block and Jansen, 1976).

Recently we proposed (Mayer and Angelov, 1978, 1980) to treat the problem of superexchange in insulators by using the EHF method. In insulators it is often sufficient, at least to a first approximation, to consider only a single cluster containing the given pair or ions, reducing in this way the problem to a molecular-type one.

The simplest model of superexchange used for qualitative or sometimes even semiquantitative considerations is the so-called "three-center, four-electron" model (Fig. 8) in which one considers explicitly only the magnetic ions and one ligand atom, and each center is represented by only one orbital. These usually are the appropriate 3d orbitals χ_1 and χ_3 of the magnetic ions and a 2p orbital χ_2 of the ligand atom. As in the zeroth-order approximation χ_1 and χ_3 are singly occupied and χ_2 contains two electrons, we have four electrons in the model.

The quantity which has to be determined is the singlet–triplet separation or the "exchange parameter" J in the (formal) spin Hamiltonian (Löwdin, 1962b)

$$\hat{H}_{\mathrm{sp}} = E_0 - 2J\hat{\vec{s}}_1\hat{\vec{s}}_2, \tag{90}$$

where

$$E_0 = \tfrac{1}{4} \cdot {}^1E + \tfrac{3}{4} \cdot {}^3E; \qquad J = \tfrac{1}{2}({}^1E - {}^3E); \tag{91}$$

$\vec{s}_1$ and $\vec{s}_2$ are the spin operators of the magnetic electrons, and ${}^1E$ and ${}^3E$ are the singlet and triplet energies, respectively.

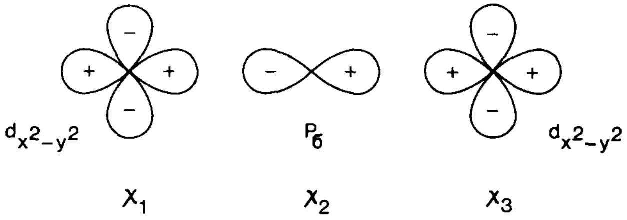


Fig. 8. The "three-center, four-electron" model of superexchange.

The model is usually treated by a CI-type approach, in which different "processes" (i.e., the admixture of different excited configurations) are considered. This leads to extremely complicated expressions (Huang and Orbach, 1967) as one has to take into account the overlap[37] between the orbitals χ. (Since the magnetic orbitals are essentially singly occupied, the usual RHF LCAO–MO theory is not well applicable even to the singlet state, because it strongly exaggerates the weight of the ionic terms.)

One of the evident advantages of applying the EHF method to the superexchange problem is that the singlet and triplet energies can be obtained in the framework of exactly the same formalism. Furthermore, it was established (Mayer and Angelov, 1978, 1980) that for the "three-center, four-electron" model the EHF wave function is identical with the full CI for the triplet and it is nearly so in the case of the singlet. This is due to the rather limited basis set and is not valid, of course, for larger models; the physical picture obtained by considering the EHF wave function may be easily generalized, however. The use of the EHF computational scheme is also, of course, not restricted to this simplest model.

Being a generalized one-particle approach, the EHF method immediately provides us with a well-visualizable picture of the one-electron MO-s, for which the different interactions cause each orbital to be more or less delocalized over the whole cluster. The analysis of the EHF wave function (Mayer and Angelov, 1978, 1980) also permitted a slightly different point of view of the problem, more closely related to the "unperturbed" physical model in which the orbital χ_2 is doubly filled and the orbitals χ_1 and χ_3 are singly occupied. The "direct" interactions between the magnetic electrons can be described by solving a Weinbaum-type "full CI" problem for the two electrons on χ_1 and χ_3, omitting from the explicit consideration or keeping doubly filled the ligand orbital χ_2. Now, the analysis of the EHF wave function in terms of different configurations showed (Mayer and Angelov, 1978, 1980) that the different "direct" and "indirect" interactions may be simultaneously taken into account in the following manner: the doubly filled ligand orbital χ_2 is replaced by a slightly delocalized doubly filled[38] MO; the admixture of the AO-s χ_1 and χ_3 to this orbital (the degree of which is optimized automatically in the course of the EHF variation procedure) reflects mainly the covalency effects in the metal–ligand bonding. The magnetic orbitals χ_1 and χ_3 are

---

[37] As the considerations usually remained at a semiqualitative level, the nonorthogonality problem was not treated, as usual in quantum chemistry, by performing the CI in an orthogonalized basis, since in this case one would lose the immediate connection with the original physical model.

[38] In the model discussed, one of the EHF orbitals must be doubly filled due to the small basis set.

also replaced by some new ones containing a small admixture of the ligand orbital χ_2. These orbitals may be immediately connected (Mayer and Angelov, 1980) with the "effective magnetic orbitals" in Anderson's (1963) "ligand field" or "kinetic exchange" model for superexchange. The CI for the two electrons on these new effective magnetic orbitals is also automatically included in the EHF variation procedure.

The use of the EHF method in treating the superexchange may become of specific importance if one considers problems in which each magnetic ion contains more than one d electron and is in a high spin state. In this case we have, besides the closed shells, two subsystems of electrons, within each of which the spins are coupled together, due to Hund's rule, to form the states of maximum multiplicity. We are essentially interested in the different states resulting from coupling of these subsystems. As the spin-coupling scheme (Section II,A) characteristic of the EHF wave function exactly corresponds to this situation, the EHF method is expected (Mayer and Angelov, 1980) to be especially adequate to treat problems of this type. One may propose also a straightforward and perhaps adequate simplified scheme: the orbitals corresponding to shells of physically closed character (for which the EHF method would give no significant orbital splitting anyway) may be kept doubly filled and one may apply the full EHF formalism only to the shells of physically open character, performing, however, a simultaneous variational optimization of the LCAO coefficients for the orbitals of both closed and open type.

V. Concluding Remarks

Since in 1955 when Löwdin published his classical work, the conceptual simplicity of the EHF method, on the one hand, and the complexity of solving the corresponding variational problem, on the other, have represented a great challenge for generations of quantum chemists. As we have seen, a number of different approaches were worked out; one of them is discussed in some detail in Section III, others are summarized in Section II. The development of the theory in this field (of course in permanent interaction with that in the other branches of quantum chemistry and related topics) led to a deeper understanding of different important aspects of the electronic structure of atoms, molecules, and the solid state. A number of actual calculations were also done by using the EHF method and the related approaches; these clarified the possibilities and limitations of the method. Accordingly, we shall close the present article devoted to the spin-projected EHF method by a brief summary of the conclusions which one may get at in this respect.

The great enthusiasm of the first period was based on the extremely good results which can be obtained for the two-electron systems by using the EHF method and, especially, the projected HF method in the more general sense (cf. Section II,I). Unfortunately, the limitations due to the EHF spin-coupling scheme did not permit obtainment of similar good results for larger systems. This essential limitation prevented the method from fulfilling the expectation that it might become a universal key for solving the correlation problem.[39] Its importance, however, is very large in conceptual analysis of problems like antiferromagnetism. As the EHF calculations are time-consuming, one must select those problems for which the characteristic features of the EHF method (not only the visualizability of the generalized one-particle scheme, but when possible, *also the spin-coupling scheme*) represent advantages with respect to the other approaches. Some fields will be discussed here in which, according to the opinion of the present author, the use of the EHF method may be especially fruitful.

The correlation in π-electron systems (especially antiaromatics). The "horizontal" correlation characteristic for such systems is described well by the EHF method. One may note in this connection that the more simple AMO method gives results which are in many cases almost as good as the EHF ones. The comparison of results for butadiene shown in Table VI may well illustrate this statement.[40] It is, however, this author's opinion that for larger systems it is practically easier to carry out a complete EHF calculation than a many-parameter AMO one.

Formation and breaking of a chemical bond. In this respect the EHF method may be most adequate for consideration of diatomics, especially in those cases in which other methods meet difficulties. Thus, the investigation of the potential curves for molecules like N_2 cannot be carried out for the interval of larger interatomic distances by using methods like the usual CI procedures for single and double substitutions or their equivalents. In fact, in order to obtain a correct asymptotic behavior, it would be necessary to take into account the higher excitations also (P. Pulay, private communication, 1978). At the same time the EHF wave function is

---

[39] The further generalizations of the method (e.g., the use of GSO-s) possibly could help to overcome these limitations. The corresponding variational problem, however, may be more complex than that in the EHF case by one or more orders of magnitude.

[40] As noted in Section II,I, the EHF method can be considered as an AMO using variationally optimized "starting" MO-s. These optimized MO-s can be obtained *a posteriori* by performing a natural orbital analysis of the EHF wave function. For butadiene these "starting" MO-s were found to be rather close to the canonical RHF MO-s, thus demonstrating the extremely ingenious character of the original AMO idea. (The mixing of the occupied and virtual RHF orbitals in the EHF natural orbitals is characterized by coefficients of value only 0.025 and 0.029. Similar results were obtained for the PPP parametrization also.)

TABLE VI

COMPARISON OF DIFFERENT METHODS FOR THE GOEPPERT-MAYER–SKLAR MODEL OF THE π-ELECTRONS IN BUTADIENE[a]

Method	Energy improvement with respect to RHF solution (eV)	Part of correlation energy taken into account (%)
UHF local minimum[b]	0.121	5.1
UHF local minimum[b] + SP	0.922	39.2
EHF local minimum[b]	1.555	66.2
UHF	0.641	27.3
UHF + SP	1.727	73.5
EHF	2.057	87.6
AMO (Δ-MO-s)[c,d]	1.956	83.2
AMO (RHF-MO-s)[c]	2.034	86.6
Full CI[c]	2.35	100

[a] I. Mayer, unpublished results.

[b] Solutions with spatially symmetry-adapted *individual* one-electron orbitals.

[c] Literature data taken from Pauncz's (1967) book.

[d] "Topological" starting MO-s (obtained by diagonalizing the overlap matrix).

quite adequate to describe these dissociation processes: at the larger interatomic distances the atoms are in states of high multiplicity *due to Hund's rule,* and their valence electrons form just two subsystems of maximum multiplicity which correspond to the EHF spin-coupling scheme. Accordingly, one may expect an especially good asymptotic behavior of the EHF potential curves. At intermediate distances the interaction between such atoms will be reflected by the delocalization of the EHF orbitals. (As an illustration, we show in Fig. 9 the RHF and EHF potential curves of the N_2 molecule calculated at the CNDO/2 level of integral approximation. From about 5 Å the EHF energy is equal to the full machine accuracy with the sum of the CNDO energies of two nitrogen atoms.)

A related useful application of the EHF method may be the calculation of potential curves for diatomic fragments necessary in calculating potential surfaces of reacting systems by using the "diatomics in molecules" method (L. Zülicke, private communication, 1978). The good asymptotic

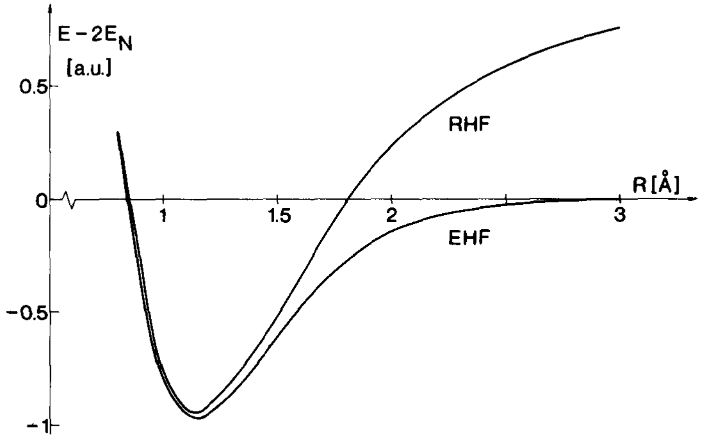


Fig. 9. Potential curves of the N_2 molecule calculated at the CNDO/2 level (I. Mayer, unpublished results).

behavior and the possibility of performing the calculations for different spin states (including "exotic" ones) in the framework of the same formalism may make the EHF method especially advantageous in this respect.

Superexchange and similar problems—see Section IV,B.

In all of the aforementioned cases the closed shells (including the σ-electrons of the conjugated molecules) may perhaps reasonably be treated by using the simplification discussed in Section IV,B.

It follows from the previous examples that (beyond the π-electron problems) the application of the EHF method may be advantageous for systems in which either there is a well-defined two-electron open-shell subsystem (like a bond which is broken) or in which the open-shell electrons form—due perhaps to *Hund's rule*—two subsystems of high multiplicity, coupled in an antiferromagnetic manner.

ACKNOWLEDGMENTS

The author feels himself to be indebted to Professor Per-Olov Löwdin, the very father of the EHF method, for the privilege of summarizing here his point of view on the subject.

The author also wishes to express his gratitude to all of his colleagues who contributed to his work on the EHF method: to those who worked together with him and were coauthors of some of his papers referred to in this chapter, and to those who helped him by their advice, useful discussions or correspondence, as well as by placing at his disposal computer programs or subprograms. It is only possible here to acknowledge the following people: S. A. Angelov, A. Ažman, G. Biczó, E. Kapuy, M. Kertész, V. Klimo, J. Koller, J. Kondász, V.

A. Kuprievich, J. Ladik, F. Martino, A. Neszmélyi, T. Pollák, P. Pulay, S. Suhai, J. Tiňo, I. I. Ukrainsky, and others.

The author is indebted to Professor G. Schay for reading the manuscript and for his useful comments.

Furthermore, the author should express his gratitude to the Director of the Central Research Institute for Chemistry of the Hungarian Academy of Sciences for making possible for many years the author's work on the subject. The financial support of the Boris Kidrič Fund (Yugoslavia) in performing the calculations presented in Table V is also gratefully acknowledged.

Last but not least, it is to be noted that Eqs. (33), (35)–(38), (48)–(51), (53)–(55), (59)–(63), (68), (70)–(73), (85)–(89) as well as Figs. 2–4 and 7 are reproduced by permission of the publisher John Wiley and Sons, Inc. from the earlier papers of the author, published originally in the *International Journal of Quantum Chemistry*. Similarly, Eqs. (28), (64), (65), and Fig. 6 are reproduced by permission of the Publishing House of the Hungarian Academy of Sciences from the author's papers which appeared in the *Acta Physica Academiae Scientiarum Hungaricae*. The generous permission to reproduce these materials from both publishers is gratefully acknowledged.

REFERENCES

Adams, W. H. (1963). *J. Chem. Phys.* **39**, 23.

Amos, A. T. (1962). *Mol. Phys.* **5**, 91.

Amos, A. T., and Hall, G. G. (1961). *Proc. R. Soc. London, Ser. A* **263**, 483.

Amos, A. T., and Snyder, L. C. (1964). *J. Chem. Phys.* **41**, 1773.

Anderson, P. W. (1963). *Solid State Phys.* **14**, 99.

Auffray, J. P., and Percus, J. (1962). *C. R. Hebd. Seances Acad. Sci.* **254**, 3170.

Beebe, N. H. F., and Lunell, S. (1975). *J. Phys. B* **8**, 2320.

Bendazzoli, G. L., Palmeri, P., and Zauli, C. (1971). *Chem. Phys. Lett.* **8**, 556.

Berencz, F. (1958). *Proc. Phys. Soc., London* **71**, 152.

Berencz, F. (1960). *Magy. Fiz. Foly.* **8**, 223.

Berencz, F. (1963). *Magy. Fiz. Foly.* **11**, 479.

Berencz, F., and Pauncz, R. (1958). *Proc. Phys. Soc., London* **71**, 145.

Berthier, G. (1954a). *C. R. Hebd. Seances Acad. Sci.* **238**, 91.

Berthier, G. (1954b). *J. Chim. Phys.* **51**, 363.

Birs, F. W., and Fraga, S. (1963). *J. Chem. Phys.* **38**, 2552.

Blint, R. J., and Goddard, W. A., III (1970). *Chem. Phys. Lett.* **5**, 302.

Block, R., and Jansen, L. (1976). *In* "Quantum Science. Methods and Structure" (J.-L. Calais *et al.*, eds.), p. 123. Plenum, New York.

Brandas, E. (1967). *Int. J. Quantum Chem.* **1**, 847.

Brandas, E. (1968). *Int. J. Quantum Chem.* **2**, 37, 391, 793.

Brigman, G. H., and Matsen, F. A. (1957). *J. Chem. Phys.* **27**, 829.

Brown, R. D., and Williams, G. R. (1973). *Mol. Phys.* **25**, 673.

Brown, R. E., and Larson, S. (1973). *J. Chem. Phys.* **59**, 4192.

Bunge, C. F. (1967). *Phys. Rev.* **154**, 70.

Bunge, C. F., and Bunge, A. (1971). *J. Comput. Phys.* **8**, 409.

Burden, F. R. (1972). *Int. J. Quantum Chem.* **6**, 647.

Burke, E. A., and Mulligan, J. F. (1958). *J. Chem. Phys.* **28**, 995.

Calais, J.-L. (1965). *Ark. Fys.* **28**, 479, 511.

Calais, J.-L. (1967). *Int. J. Quantum Chem.* **1S**, 661.

Calais, J.-L. (1979). "Mott and Peierls Gaps." Preprint No. 603. Quantum Chem. Group, Uppsala University.

Calais, J.-L., and Sperber, G. (1973). *Int. J. Quantum Chem.* **7**, 501.

Chong, D. P. (1966). *J. Chem. Phys.* **45**, 3317.

Claxton, T. A., and Smith, N. A. (1971). *Theor. Chim. Acta* **22**, 399.

Coulson, C. A. (1963). *Tetrahedron* **19**, Suppl. 2, 49.

Coulson, C. A. (1964). *Proc. Phys. Soc., London* **84**, 511.

Coulson, C. A. (1966). *In* "Quantum Theory of Atoms, Molecules and the Solid State" (P.-O. Löwdin, ed.), p. 97. Academic Press, New York.

Coulson, C. A. (1971). *Mol. Phys.* **20**, 687.

Coulson, C. A., and Fischer, I. (1949). *Philos. Mag.* [7] **40**, 386.

Coulson, C. A., and Hibbert, A. (1967). *Proc. Phys. Soc., London* **91**, 33.

Coulson, C. A., and Longuet-Higgins, H. C. (1947). *Proc. R. Soc., London, Ser. A* **192**, 16.

Coulson, C. A., and Rushbroke, G. S. (1940). *Proc. Cambridge Philos. Soc.* **36**, 193.

Cox, P. A., and Wood, M. H. (1976). *Theor. Chim. Acta* **41**, 269, 279.

Cremaschi, P., Gamba, A., Morosi, G., and Simonetta, M. (1976). *Theor. Chim. Acta* **41**, 177.

Dahl, J. P., Johanson, H., Truax, D. R., and Ziegler, T. (1970). *Chem. Phys. Lett.* **6**, 64.

Das, G., and Wahl, A. C. (1966). *J. Chem. Phys.* **44**, 87.

Davidson, E. R., and Jones, L. L. (1962). *J. Chem. Phys.* **37**, 2966.

Dearman, H. H., and Lefebvre, R. (1961). *J. Chem. Phys.* **34**, 72.

de Heer, J. (1962a). *J. Chem. Phys.* **37**, 2080.

de Heer, J. (1962b). *J. Phys. Chem.* **66**, 2288.

de Heer, J. (1963). *Rev. Mod. Phys.* **35**, 631.

de Heer, J., and Pauncz, R. (1963). *J. Chem. Phys.* **39**, 2314.

Delgado-Barrio, G., Tellez de Cepeda, M., and Smeyers, Y. G. (1977). *An. Fis.* **73**, 231.

Dermit, G. (1962). *Phys. Rev.* **127**, 1110.

Dugay, M., and Thomas, J. (1975). *Int. J. Quantum Chem.* **9**, 1087.

Eckart, C. (1930). *Phys. Rev.* **36**, 878.

Epstein, S. T. (1974). "The Variation Method in Quantum Chemistry." Academic Press, New York.

Fischer, H. (1962). *J. Chem. Phys.* **37**, 1094.

Froese, C. (1966). *Phys. Rev.* **150**, 1.

Fukutome, H. (1972). *Prog. Theor. Phys.* **47**, 1156.

Fukutome, H. (1973). *Prog. Theor. Phys.* **49**, 22.

Gallup, G. A. (1968). *J. Chem. Phys.* **48**, 1752.

Gallup, G. A. (1969a). *J. Chem. Phys.* **50**, 1206.

Gallup, G. A. (1969b). *J. Chem. Phys.* **50**, 1214.

Gallup, G. A. (1970). *J. Chem. Phys.* **52**, 893.

Gallup, G. A. (1973a). *J. Chem. Phys.* **58**, 407.

Gallup, G. A. (1973b). *J. Chem. Phys.* **58**, 2181.

Gallup, G. A. (1973c). *Adv. Quantum Chem.* **7**, 113.

Gerratt, J. (1971). *Adv. At. Mol. Phys.* **7**, 141.

Goddard, W. A., III (1967a). *Phys. Rev.* **157**, 73.

Goddard, W. A., III (1967b). *Phys. Rev.* **157**, 81.

Goddard, W. A., III (1967c). *Phys. Rev.* **157**, 93.

Goddard, W. A., III (1968a). *J. Chem. Phys.* **48**, 450.

Goddard, W. A., III (1968b). *J. Chem. Phys.* **48**, 1008.

Goddard, W. A., III (1968c). *J. Chem. Phys.* **48**, 5537.

Goddard, W. A., III (1968d). *Phys. Rev.* **169,** 120.

Goddard, W. A., III (1968e). *Phys. Rev.* **176,** 106.

Goddard, W. A., III (1969). *Phys. Rev.* **182,** 48.

Goddard, W. A., III (1970). *Int. J. Quantum Chem.* **3S,** 593.

Goddard, W. A., III (1972). *J. Chem. Phys.* **56,** 3348.

Goddard, W. A., III, Dunning, T. H., Jr., and Hunt, W. J. (1969). *Chem. Phys. Lett.* **4,** 231.

Goddard, W. A., III, Dunning, T. H., Jr., Hunt, W. J., and Hay, P. J. (1973). *Acc. Chem. Res.* **6,** 368.

Gorlov, Y. I., and Ukrainsky, I. I. (1973). Preprint No. ITP-73-138R. Inst. for Theoretical Physics, Ukrainian Academy of Sciences, Kiev.

Gorlov, Y. I., and Ukrainsky, I. I. (1974). Application of direct minimization of the energy functional for calculating the ground and excited states of molecules. Preprint No. ITP-74-69E. Inst. for Theoretical Physics, Ukrainian Academy of Sciences, Kiev.

Goscinski, O., and Calais, J.-L. (1965). *Ark. Fys.* **29,** 135.

Grein, F., and Chang, T. C.) (1971). *Chem. Phys. Lett.* **12,** 44.

Guberman, S. L., and Goddard, W. A., III (1970). *J. Chem. Phys.* **53,** 1802.

Hall, G. G., and Amos, A. T. (1965). *Adv. At. Mol. Phys.* **1,** 1.

Hall, G. G., and Amos, A. T. (1969). *In* "Physical Chemistry. An Advanced Treatise" (D. Henderson, ed.), Vol. 3, p. 447. Academic Press, New York.

Hameed, S., Hui, S., Musher, J. I., and Shullman, J. M. (1969). *J. Chem. Phys.* **51,** 502.

Harcourt, R. D. (1976). *In* "Quantum Science. Methods and Structure" (J.-L. Calais *et al.*, eds.), p. 105. Plenum, New York.

Hardisson, A., and Harriman, J. E. (1967). *J. Chem. Phys.* **46,** 3639.

Harriman, J. E. (1963a). The first-order density matrix for a projected single determinant. Preprint No. 97. Quantum Chem. Group, Uppsala University.

Harriman, J. E. (1964). The first-order density matrix for a spin-projected single determinant, II: Natural expansion. Preprint No. 104. Quantum Chem. Group, Uppsala University.

Harriman, J. E. (1964). *J. Chem. Phys.* **40,** 2827.

Harriman, J. E. (1967). *In* "La structure hyperfine magnètique des atomes et des molécules," (Colloques Internationaux du CNRS, No. 164), p. 139. CNRS, Paris.

Harriman, J. E., and Sando, K. M. (1968). *J. Chem. Phys.* **48,** 5138.

Harris, F. E. (1966). *Mol. Phys.* **11,** 243.

Harris, F. E. (1967a). *J. Chem. Phys.* **46,** 2769.

Harris, F. E. (1967b). *J. Chem. Phys.* **47,** 1047.

Harris, F. E. (1967c). *Adv. Quantum Chem.* **3,** 61.

Hay, P. J., Thibeault, J. C., and Hoffman, R. (1975). *J. Am. Chem. Soc.* **97,** 4884.

Heikes, L. G., and Gallup, G. A. (1970). *J. Chem. Phys.* **52,** 888.

Hibbert, A. (1967). *Proc. Phys. Soc., London* **91,** 819.

Hibbert, A. (1968). *J. Phys. B* **1,** 402.

Huang, N. L., and Orbach, R. (1967). *Phys. Rev.* **154,** 487.

Hurst, R. P., Gray, J. D., Brigman, G. H., and Matsen, F. A. (1958). *Mol. Phys.* **1,** 189.

Huzinaga, S. (1969). *J. Chem. Phys.* **51,** 3971.

Hylleraas, E. A. (1929). *Z. Phys.* **54,** 347.

Itoh, I., and Yoshizumi, H. (1955). *J. Phys. Soc. Jpn.* **10,** 201.

Johns, J. W. C., Grimm, F. A., and Porter, R. F. (1967). *J. Mol. Spectrosc.* **22,** 435.

Jucys, A. P., and Lazauskas, V. M. (1974). *Int. J. Quantum Chem.* **8,** 839.

Jug, K. (1973). *Theor. Chim. Acta* **30,** 231.

Kaldor, U. (1968a). *J. Chem. Phys.* **48,** 835.

Kaldor, U. (1968b). *J. Chem. Phys.* **49,** 6.

Kaldor, U. (1968c). *Phys. Rev.* **176,** 19.

Kaldor, U. (1968d). *Phys. Rev.* **176,** 71.

Kaldor, U. (1970a). *Phys. Rev. A* **1,** 1586.

Kaldor, U. (1970b). *Phys. Rev. A* **2,** 1267.

Kaldor, U. (1972). *Chem. Phys. Lett.* **14,** 124.

Kaldor, U., and Harris, F. E. (1969). *Phys. Rev.* **183,** 1.

Kaldor, U., Schaefer, H. F., and Harris, F. E. (1968). *Int. J. Quantum Chem.* **2S,** 13.

Kancerevičius, A. (1968). *Liet. Fiz. Rinkinys* **8,** 291.

Kertész, M. (1977). Thesis, Hungarian Academy of Sciences, Budapest.

Kertész, M., Koller, J., and Ažman, A. (1976). *Phys. Rev. B* **14,** 76.

Kertész, M., Koller, J., and Ažman, A. (1979). *Phys. Rev. B* **19,** 2034.

Kerwin, J., and Burke, E. A. (1962). *J. Chem. Phys.* **36,** 2987.

Klimo, V., and Tiňo, J. (1976). *Int. J. Quantum Chem.* **10,** 761.

Klimo, V., and Tiňo, J. (1977). *Chem. Phys. Lett.* **46,** 481.

Klimo, V., and Tiňo, J. (1978). *Mol. Phys.* **35,** 1777.

Kotani, M., Amemiya, A., Ishiguro, E., and Kimura, T. (1955). "Table of Molecular Integrals." Maruzen, Tokyo.

Kotani, M., Mizuno, Y., and Kayama, K. (1960). *Rev. Mod. Phys.* **32,** 266.

Kouba, J., and Öhrn, Y. (1969). *Int. J. Quantum Chem.* **3,** 513.

Koutecký, J. (1967). *J. Chem. Phys.* **46,** 2443.

Kruglyak, N. E., Gey E., and Kruglyak, Yu. A. (1973). Preprint No. ITP-73-132R. Inst. for Theoretical Physics, Ukrainian Academy of Sciences, Kiev.

Kruglyak, Yu. A., and Ukrainsky, I. I. (1970). *Int. J. Quantum Chem.* **4,** 57.

Kunik, D., and Kaldor, U. (1971). *J. Chem. Phys.* **55,** 3127.

Kunik, D., and Kaldor, U. (1972). *J. Chem. Phys.* **56,** 1741.

Ladik, J., and Martino, F. (1970). *J. Chem. Phys.* **52,** 2259.

Ladik, J., and Otto, P. (1975). *Chem. Phys. Lett.* **31,** 83.

Ladner, R. C., and Goddard W. A., III (1969). *J. Chem. Phys.* **51,** 1073.

Larson, E. G., and Thorson, W. R. (1966). *J. Chem. Phys.* **45,** 1539.

Laskowski, B., and Lunell, S. (1975). *Int. J. Quantum Chem.* **9S,** 175.

Laskowski, B., and van Leuven, P. (1971). *Int. J. Quantum Chem.* **5,** 507; erratum: **7,** 134 (1973).

Lefebvre, R. (1959). *Cah. Phys.* **13,** 369.

Lefebvre, R., and Prat, R. (1967). *Chem. Phys. Lett.* **1,** 388.

Lefebvre, R., and Prat, R. (1969). *Int. J. Quantum Chem.* **3,** 93.

Lefebvre, R., and Smeyers, Y. G. (1967). *Int. J. Quantum Chem.* **1,** 403.

Lefebvre, R., Dearman, H. H., and McConnell, H. M. (1960). *J. Chem. Phys.* **32,** 176.

Lengsfield, B. H., III, and Schug, J. C. (1978). *Chem. Phys. Lett.* **57,** 442.

Levy, B., and Berthier, G. (1968). *Int. J. Quantum Chem.* **2,** 307.

Lieb, E. H., and Simon, B. (1974). *J. Chem. Phys.* **61,** 735.

Lim, T. K. (1975). *Int. J. Quantum Chem.* **9,** 981.

Lindner, P., and Lunell, S. (1967). *Int. J. Quantum Chem.* **1,** 841.

Longuet-Higgins, H. C. (1966). *In* "Quantum Theory of Atoms, Molecules, and the Solid State" (P.-O. Löwdin ed.), p. 105. Academic Press, New York.

Löwdin, P.-O. (1955a). *Phys. Rev.* **97,** 1474.

Löwdin, P.-O. (1955b). *Phys. Rev.* **97,** 1490.

Löwdin, P.-O. (1955c). *Phys. Rev.* **97,** 1509.

Löwdin, P.-O. (1956). *Adv. Phys.* **5,** 1.

Löwdin, P.-O. (1958a). *Ann. Acad. Reg. Sci. Ups.* **2,** 127.

Löwdin, P.-O. (1958b). *In* "Calcul des fonctions d'onde moléculaire," (Colloques Internationaux du CNRS, No. 82), p. 23. CNRS, Paris.

Löwdin, P.-O. (1959). *Adv. Chem. Phys.* **2,** 207.

Löwdin, P.-O. (1960). *Rev. Mod. Phys.* **32,** 328.
Löwdin, P.-O. (1962a). *J. Appl. Phys.* **33,** Suppl. 251.
Löwdin, P.-O. (1962b). *Rev. Mod. Phys.* **34,** 80.
Löwdin, P.-O. (1962c). *Rev. Mod. Phys.* **34,** 520.
Löwdin, P.-O. (1963). *Rev. Mod. Phys.* **35,** 496.
Löwdin, P.-O. (1964). *Rev. Mod. Phys.* **36,** 966.
Löwdin, P.-O. (1966). *In* "Quantum Theory of Atoms, Molecules, and the Solid State" (P.-O. Löwdin, ed.), p. 601. Academic Press, New York.
Löwdin, P.-O. (1969). *Adv. Chem. Phys.* **14,** 283.
Löwdin, P.-O. (1976). "Molecular structure calculations." Preprint No. 498. Quantum Chem. Group, Uppsala University.
Löwdin, P.-O., and Goscinski, O. (1970). *Int. J. Quantum Chem.* **3S,** 533.
Lunell, S. (1968). *Phys. Rev.* **173,** 85.
Lunell, S. (1970). *Phys. Rev. A* **1,** 360.
Lunell, S. (1972a). *Chem. Phys. Lett.* **13,** 93.
Lunell, S. (1972b). *Chem. Phys. Lett.* **15,** 27.
Lunell, S. (1973). *Phys. Rev. A* **7,** 1229.
Lunell, S. (1979). *Int. J. Quantum Chem.* **15,** 97.
Lunell, S., and Lindner, P. (1968). *J. Chem. Phys.* **48,** 2752.
McWeeny, R., and Sutcliffe, B. T. (1969). "Methods of Molecular Quantum Mechanics." Academic Press, New York.
Manne, R. (1966). *Theor. Chim. Acta* **6,** 116.
Mano, K. (1970). *J. Chem. Phys.* **52,** 2785.
Martino, F., and Ladik, J. (1970). *J. Chem. Phys.* **52,** 2262; erratum: **54,** 1432 (1971).
Martino, F., and Ladik, J. (1971). *Phys. Rev. A* **3,** 862.
Matsen, F. A. (1964). *Adv. Quantum Chem.* **1,** 59.
Matsen, F. A., and Cantu, A. A. (1969). *J. Phys. Chem.* **73,** 2488.
Mayer, I. (1971a). *Acta Phys. Acad. Sci. Hung.* **30,** 373; correction of a misprint: **31,** 403 (1972).
Mayer, I. (1971b). *Chem. Phys. Lett.* **11,** 397.
Mayer, I. (1973a). *Acta Phys. Acad. Sci. Hung.* **34,** 83.
Mayer, I. (1973b). *Acta Phys. Acad. Sci. Hung.* **34,** 305.
Mayer, I. (1974a). *Int. J. Quantum Chem.* **8,** 363.
Mayer, I. (1974b). *Acta Phys. Acad. Sci. Hung.* **36,** 11.
Mayer, I. (1974c). *Int. J. Quantum Chem.* **8,** 893.
Mayer, I. (1974d). *Acta Phys. Acad. Sci. Hung.* **37,** 39.
Mayer, I. (1975). *Acta Phys. Acad. Sci. Hung.* **39,** 133.
Mayer, I. (1976). *Kém. Közl.* **46,** 397.
Mayer, I. (1977). Thesis, Hungarian Academy of Sciences, Budapest.
Mayer, I. (1978). *Int. J. Quantum Chem.* **14,** 29.
Mayer, I., and Angelov, S. A. (1978). *Physica* (*Utrecht*) **95B,** 93.
Mayer, I., and Angelov, S. A. (1980). *Int. J. Quantum Chem.* (in press).
Mayer, I., and Kertész, M. (1975). *Int. J. Quantum Chem.* **9,** 527.
Mayer, I., and Kertész, M. (1976). *Int. J. Quantum Chem.* **10,** 961.
Mayer, I., and Kondász, J. (1975). *Int. J. Quantum Chem.* **9,** 517; correction of a misprint in Mayer (1978).
Mayer, I., Ladik, J., and Biczó, G. (1973). *Int. J. Quantum Chem.* **7,** 583.
Mestechkin, M. M. (1967). *Int. J. Quantum Chem.* **1,** 421.
Mestechkin, M. M. (1973). *Int. J. Quantum Chem.* **7,** 425.
Mestechkin, M. M. (1977). "Metod Matrici Plotnosti v Teorii Molekul." Naukova Dumka, Kiev.

Mestechkin, M. M., and Whyman, G. E. (1974). *Int. J. Quantum Chem.* **8,** 45.

Misurkin, I. A., and Ovchinikov, A. A. (1972). *Teor. Eksp. Khim.* **8,** 291.

Morikawa, T. (1974). *Chem. Phys. Lett.* **32,** 521.

Moskowitz, J. (1963). *J. Chem. Phys.* **38,** 677.

Musher, J. I. (1969). *J. Chem. Phys.* **50,** 3741.

Musher, J. I. (1970). *Chem. Phys. Lett.* **7,** 397.

Nesbet, R. K. (1961). *Rev. Mod. Phys.* **33,** 28.

Newman, D. J. (1977). *Physica (Utrecht)* **86B–88B,** 1018.

Pauncz, R. (1962). *J. Chem. Phys.* **37,** 2739.

Pauncz, R. (1963). *Tetrahedron* **19,** Suppl. 2, 43.

Pauncz, R. (1964). *In* "Molecular Orbitals in Chemistry, Physics, and Biology" (P.-O. Löwdin and B. Pullman, eds.), p. 433. Academic Press, New York.

Pauncz, R. (1967). "Alternant Molecular Orbital Method." Saunders, Philadelphia, Pennsylvania.

Pauncz, R. (1969). *In* "Physical Chemistry. An Advanced Treatise" (D. Henderson, ed.), Vol. 3, p. 185. Academic Press, New York.

Pauncz, R., de Heer, J., and Löwdin, P.-O. (1962a). *J. Chem. Phys.* **36,** 2247.

Pauncz, R., de Heer, J., and Löwdin, P.-O. (1962b). *J. Chem. Phys.* **36,** 2257.

Percus, J. K., and Rotenberg, A. (1962). *J. Math. Phys.* **3,** 928.

Philips, D. H., and Schug, J. C. (1974a). *J. Chem. Phys.* **60,** 1597.

Philips, D. H., and Schug, J. C. (1974b). *J. Chem. Phys.* **61,** 1031.

Pople, J. A. (1953). *Trans. Faraday Soc.* **49,** 1375.

Pople, J. A., and Nesbet, R. K. (1954). *J. Chem. Phys.* **22,** 571.

Popov, N. A. (1970). *Zh. Strukt. Khim.* **11,** 727.

Poshuta, R. D., and Kramling, R. W. (1968). *Phys. Rev.* **167,** 139.

Pratt, G. W., Jr. (1953). *Phys. Rev.* **92,** 278.

Pratt, G. W., Jr. (1956). *Phys. Rev.* **102,** 1303.

Rosenberg, M. (1975). *Diss. Abstr. Int. B* **36,** 1785.

Rosenberg, M., and Martino, F. (1975). *J. Chem. Phys.* **63,** 5345.

Ruedenberg, K., and Poshuta, R. D. (1972). *Adv. Quantum Chem.* **6,** 268.

Salotto, A. W., and Burnelle, R. (1970). *J. Chem. Phys.* **52,** 2936.

Sando, K. M., and Harriman, J. E. (1967). *J. Chem. Phys.* **47,** 180.

Sarma, C. R. (1975). *Int. J. Quantum Chem.* **9,** 1109.

Sasaki, F., and Ohno, K. (1963). *J. Math. Phys.* **4,** 1140.

Schaefer, H. F., III (1972). "The Electronic Structure of Atoms and Molecules. A Survey of Rigorous Quantum Mechanical Results." Addison-Wesley, Menlo Park, California.

Shull, H., and Löwdin, P.-O. (1956a). *Phys. Rev.* **101,** 1730.

Shull, H., and Löwdin, P.-O. (1956b). *J. Chem. Phys.* **25,** 1035.

Shull, H., and Löwdin, P.-O. (1959). *J. Chem. Phys.* **30,** 617.

Silberman, Z., Gershgorn, Z., and Pauncz, R. (1968). *Int. J. Quantum Chem.* **2,** 453.

Simons, J., and Harriman, J. E. (1969). *J. Chem. Phys.* **51,** 296.

Slater, J. C. (1930). *Phys. Rev.* **35,** 509.

Slater, J. C. (1951). *Phys. Rev.* **82,** 538.

Smeyers, Y. G. (1971). *An. Fis.* **67,** 17.

Smeyers, Y. G., and Bruceña, A. M. (1978). *Gazz. Chim. Ital.* **108,** 171.

Smeyers, Y. G., and Delgado-Barrio, G. (1974). *Int. J. Quantum Chem.* **8,** 733.

Smeyers, Y. G., and Delgado-Barrio, G. (1976). *Int. J. Quantum Chem.* **10,** 461.

Smeyers, Y. G., and Delgado-Barrio, G. (1977). *Phys. Rev. A* **15,** 1359.

Smeyers, Y. G., and Delgado-Barrio, G. (1978). *Int. J. Quantum Chem.* **14,** 641.

Smeyers, Y. G., and Doreste-Suarez, L. (1973). *Int. J. Quantum Chem.* **7,** 687.

Smeyers, Y. G., and Pinto Suarez, I. (1968). *Afinidad* **25,** 197.

Smith, V. H., Jr. (1964). *J. Chem. Phys.* **41,** 277.

Smith, V. H., Jr., and Harris, F. E. (1969). *J. Math. Phys.* **10,** 771.

Snook, I. K. (1972). *Chem. Phys. Lett.* **13,** 229.

Snyder, L. C., and Amos, A. T. (1964). *J. Am. Chem. Soc.* **86,** 1647.

Snyder, L. C., and Amos, A. T. (1965). *J. Chem. Phys.* **42,** 3670.

Sperber, G. (1973). *Int. J. Quantum Chem.* **7,** 537.

Sperber, G., and Calais, J.-L. (1973). *Int. J. Quantum Chem.* **7,** 521.

Stösser, R., Lück, R., Sauer, J., Jung, C., and Henning, H. (1975). *Z. Phys. Chem.* **256,** 574.

Sullivan, J. J. (1968). *J. Math. Phys.* **9,** 1369.

Sullivan, J. J. (1972). *Phys. Rev. A* **5,** 29.

Swalem, J. D., and de Heer, J. (1964). *J. Chem. Phys.* **40,** 378.

Szondy, T., and Kapuy, E. (1976). *Theor. Chim. Acta* **43,** 161.

Tamir, I., and Pauncz, R. (1968). *Int. J. Quantum Chem.* **2,** 433.

Tiňo, J., and Klimo, V. (1976a). *Collect. Czech. Chem. Commun.* **41,** 960.

Tiňo, J., and Klimo, V. (1976b). *Collect. Czech. Chem. Commun.* **41,** 3516.

Tiňo, J., Kühnel, V., and Klimo, V. (1978). *Collect. Czech. Chem. Commun.* **43,** 30.

Tiňo, J., Klimo, V., Claxton, T. A., and Burton, B. (1979). *J. Chem. Soc., Faraday Trans. 2* **75,** 1307.

Tyutyulkov, N. N. (1975). *Int. J. Quantum Chem.* **9,** 683.

Tyutyulkov, N. N. (1978). *Int. J. Quantum Chem.* **13,** 443.

Ukrainsky, I. I. (1971). Thesis, Ukrainian Academy of Sciences, Kiev.

Ukrainsky, I. I. (1972). *Int. J. Quantum Chem.* **6,** 473.

Ukrainsky, I. I., and Kruglyak, Yu. A. (1970). *Ukr. Fiz. Zh.* **15,** 1068.

Ukrainsky, I. I., and Kruglyak, Yu. A. (1973). Preprint No. ITP-73-3R. Inst. for Theoretical Physics, Ukrainian Academy of Sciences, Kiev.

Ukrainsky, I. I., Kruglyak, Yu. A., Preuss, H., and Janoschek, R. (1972). *Teor. Eksp. Khim.* **8,** 299.

van Kalkeren, G., Schmidt, W. W., and Block, R. (1979). *Physica* (*Utrecht*) **97B,** 315.

van Leuven, P. (1969). *Physica* (*Utrecht*) **45,** 86.

Weinbaum, S. (1933). *J. Chem. Phys.* **1,** 593.

Yonezawa, T., Nakatsuji, H., Kawamura, T., and Kato, H. (1969). *J. Chem. Phys.* **51,** 669.

Yoshizumi, H., and Itoh, I. (1955). *J. Chem. Phys.* **23,** 412.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 12

Molecular Structure Calculations*

PER-OLOV LÖWDIN
Department of Quantum Chemistry
Uppsala University
Uppsala, Sweden

and

Quantum Theory Project
University of Florida
Gainesville, Florida

---

* Work supported in part by the U.S. National Science Foundation.

ISBN 0-12-034812-8

I. Importance of the Schrödinger Equation

Since the discovery of modern quantum theory in 1925 independently by Schrödinger, Heisenberg, and Dirac, it has generally been believed that the structure of the stationary or scattering states of any molecular system could be described by finding the eigenfunctions Ψ to the time-independent Schrödinger equation

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

where H is the Hamiltonian operator

$$H = \sum_k \frac{p_k^2}{2m_k} + \sum_{k<l} \frac{e_k e_l}{r_{kl}}, \tag{2}$$

having each classical momentum p_k replaced by the differential operator

$$p_k = \frac{h}{2\pi i}\left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial y_k}, \frac{\partial}{\partial z_k}\right). \tag{3}$$

For stationary states, the complex wave function Ψ is assumed to be absolutely quadratically integrable, i.e., it belongs to a Hilbert space of type L^2, whereas for scattering states it is assumed to be finite at the boundaries of the system.

It should be observed that, in classical quantum mechanics, it had been impossible to treat molecular systems from first principles, and that it was essentially the discovery of the exchange phenomenon by Heisenberg (*1*) in 1926 in his study of the helium atom that opened the way to the treatment of two- and many-electron systems. Using this idea in a study of the ground state of the hydrogen molecule consisting of the two atoms a and b, Heitler and London (*2*) could show in 1927 that a wave function

$$\Phi(1, 2) = \tfrac{1}{2}[a(1)b(2) + b(1)a(2)], \tag{4}$$

which is invariant under the exchange of the two electrons 1 and 2, seems to correspond to the "covalent bond" in chemistry. This paper gave "birth" to the new field of *quantum chemistry,* which celebrated its fiftieth anniversary in 1977.

The first two papers on exchange were certainly more qualitative than quantitative—expressing a new idea in physics and chemistry. The final test of the validity of any theory depends on how well its numerical results compare with the experimental experience and whether one obtains complete agreement in all figures known, when all factors are taken into consideration. The first proof that the Schrödinger equation (1) in a configuration space with two electrons would give theoretical results in excellent agreement with experience was given by Hylleraas (*3*) in 1929 in his study of the helium atom. Similar evidence for the hydrogen molecule was given by James and Coolidge (*4*) in 1933. These two numerical treatments started a long series of papers on the helium atom and the hydrogen molecule which culminated in the works by Pekeris (*5*) and by Kolos and Wolniewicz (*6*), respectively, and which is still going on.

However, in the treatment of many-electron systems, the Schrödinger equation (1) is not enough—one must also consider Pauli's famous exclusion principle for the electrons. Since the total Hamiltonian H in Eq. (2) is spin free and invariant under all permutations P_r of the electronic coordinates $(r_1, r_2, \ldots, r_N)$ so that $P_rH = HP_r$, one may formulate the exclusion principle in a *spin-free theory* in terms of the properties of the irreducible representations of the symmetric group (*7*).

Following Slater, one may instead introduce the spin coordinates $\zeta_1, \zeta_2, \ldots, \zeta_n$ of the electrons explicitly and assume that the total wave function Ψ depends on the combined coordinates $X = (x_1, x_2, \ldots, x_n)$, where $x_k = (r_k, \zeta_k)$, so that $\Psi = \Psi(x_1, x_2, \ldots, x_N)$. In forming the binary product

$$\langle \Psi_1 | \Psi_2 \rangle = \int \Psi_1^* \Psi_2 \, dX, \tag{5}$$

where $dX = dx_1 \, dx_2 \cdots dx_N$, one should further integrate over the ordinary coordinates $r_1, r_2, \ldots, r_N$, and sum over the spins $\zeta_1, \zeta_2, \ldots, \zeta_N$. One may then formulate the exclusion principle as an *antisymmetry condition,* which has to be fulfilled by an electronic wave function Ψ:

$$P\Psi = (-1)^p \Psi, \tag{6}$$

where p is the parity of the permutation $P = P_x$ of the coordinates $X = (x_1, x_2, \ldots, x_N)$. If one has found a solution $\Phi = \Phi(x_1, x_2, \ldots, x_N)$ to the Schrödinger equation (1) that is not fully antisymmetric, one may always consider its antisymmetric component $\Phi_{\mathrm{AS}} = \mathcal{O}_{\mathrm{AS}}\Phi$, which is selected by the projector for the antisymmetric representation:

$$\mathcal{O}_{\mathrm{AS}} = (N!)^{-1} \sum_P (-1)^p P. \tag{7}$$

It satisfies the relations $\mathcal{O}_{\mathrm{AS}}^2 = \mathcal{O}_{\mathrm{AS}}$, $P\mathcal{O}_{\mathrm{AS}} = (-1)^p \mathcal{O}_{\mathrm{AS}}$ and $\mathcal{O}_{\mathrm{AS}}^\dagger = \mathcal{O}_{\mathrm{AS}}$, where the adjoint $T^\dagger$ of an arbitrary linear operator T is defined by the relation $\langle T\Psi_1 | \Psi_2 \rangle = \langle \Psi_1 | T^\dagger \Psi_2 \rangle$. Using this projector, one may replace the antisymmetry condition (6) by the single relation

$$\mathcal{O}_{\mathrm{AS}} \Psi = \Psi. \tag{8}$$

It is hence evident that the solutions to the Schrödinger equation (1) have to satisfy certain *auxiliary conditions* associated with the Pauli exclusion principle in order to represent many-electron systems.

In retrospect over the last 50 years, one can now say that the original belief that the Schrödinger equation (1) should present a firm basis for the nonrelativistic theory of atomic and molecular structure has turned out to be valid, and that Dirac's famous prediction (*8*) of 1928 has become true:

> The general theory of quantum mechanics is now almost complete, the imperfections that still remain being in connection with the exact fitting in of the theory with relativity ideas. These give rise to difficulties only when high-speed particles are involved, and are therefore of no importance in the consideration of atomic and molecular structure and ordinary chemical reactions. . . . The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. . . .

At the same time, it is also true that our knowledge of the proper relativistic corrections and radiation effects are highly uncertain—at least in principle—and that a great deal remains to be done in developing a fully relativistic theory for many-electron systems and to study its implications for our understanding of atomic and molecular structure.

II. Calculation of Approximate Eigenfunctions

A. The Born–Oppenheimer Approximation

Since one can solve the Schrödinger equation (1) exactly in closed form only for a few exceptional cases, one has instead to rely on approximate methods in the study of molecular structure such that, in principle, they will give any numerical accuracy desired. Some methods of this type are discussed in this section.

Most molecular studies start from the Born–Oppenheimer (*9*) scheme, in which the atomic nuclei are assumed to be fixed in a first approximation. Denoting the electrons by the indices i and j, respectively, one finds that the Hamiltonian (2) reduces to the form

$$H_{\mathrm{el}} = e^2 \sum_{g<h} \frac{Z_g Z_h}{R_{gh}} + \sum_{i=1}^{\mathcal{N}} \frac{p_i^2}{2m} - e^2 \sum_i \sum_g \frac{Z_g}{r_{ig}} + \sum_{i<j} \frac{e^2}{r_{ij}}, \qquad (9)$$

where the first term $H_{(0)}$ is the internuclear Coulomb repulsion, the second term T the electronic kinetic energy, the third term L the Coulomb attraction between the electrons and the atomic nuclei, and the fourth term C the interelectronic Coulomb repulsion. After solving the electronic Schrödinger equation for various values of the nuclear coordinates R_g, one gets an energy map $E = E(R_g)$ from which one can determine the various energy minima corresponding to the different conformations of the molecular system and the associated bond lengths and bond angles. From the shape of the various local minima, one can further determine the constants characteristic of the harmonic and anharmonic vibrations of the different nuclei, the force constants of the various bonds, etc. Reintroducing the kinetic energy of the atomic nuclei into the energy considerations and going over to normal coordinates, one obtains finally a basis for the theory of molecular vibrations and rotations.

In principle, this approach seems simple and straightforward, but there are unfortunately numerous theoretical and computational complications, some of which are further discussed below. In adding the nuclear kinetic energy

$$T_N = \sum_g p_g^2/2M_g, \tag{10}$$

to the "unperturbed" Hamiltonian (8), one may consider the constants $\lambda_g = 1/M_g$ as perturbational parameters and include the term T_N by means of modern perturbation theory. Even if this approach has been fairly successful in various parts of solid-state theory, a great deal remains to be done in the theory of molecular structure.

It is hence clear that, even if the Born–Oppenheimer scheme works excellently for heavy atomic nuclei, it may not necessarily be the best starting point for the treatment of light nuclei and particularly protons. This fact may influence the future development of the theory of base–acid reactions in quantum chemistry as well as our understanding of the hydrogen bond.

B. The Variation Principle

A molecular Hamiltonian of type (9) is self-adjoint and bounded from below. It has an eigenvalue spectrum $\{E\}$, which usually consists of a point spectrum and a continuum. The spectrum starts with a series of discrete eigenvalues E_0, E_1, E_2, . . . in order from below, associated with the eigenfunctions Ψ_0, Ψ_1, Ψ_2, . . . corresponding to stationary states, followed by a continuum associated with scattering states. According to Rayleigh and Ritz, the solution of the Schrödinger equation $H\Psi = E\Psi$ for stationary states is equivalent to the variation principle:

$$\delta\langle H\rangle_{\text{AV}} = 0, \tag{11}$$

for arbitrary variations of the trial wave function in the expectation value

$$\langle H\rangle_{\text{AV}} \equiv \langle\Phi|H|\Phi\rangle/\langle\Phi|\Phi\rangle, \tag{12}$$

where Φ belongs to the Hilbert space $\mathcal{H}$ associated with the binary product (5).

Since H is bounded from below, one has in particular $\langle H\rangle_{\text{AV}} \geq E_0$. If the trial function for the ground state $\Phi = \Phi(\alpha_1, \alpha_2, \ldots)$ depends on a set of parameters $\{\alpha_k\}$, one tries to vary these quantities so that $\langle H \rangle_{AV}$ becomes a minimum. From Eckart's criterion (*10*) it follows that the closer the value of $\langle H \rangle_{AV}$ is to E_0, the closer the trial function Φ is to the ground-state eigenfunction Ψ_0.

For all trial functions Φ orthogonal to Ψ_0, the variation principle takes the form $\langle H \rangle_{AV} \geq E_1$, whereas for all Φ orthogonal to both Ψ_0 and Ψ_1, one has $\langle H \rangle_{AV} \geq E_2$, etc.

If one spans the Hilbert space $\mathcal{H}$ by means of a complete basis, one may expand all the eigenfunctions Ψ exactly in terms of this basis. However, in most practical applications, one will only be able to consider a finite set $\Phi = \{\Phi_1, \Phi_2, \Phi_3, \ldots, \Phi_n\}$ as an approximation for such a basis. An eigenfunction Ψ may then be approximately expanded in the form

$$\Psi = \sum_k \Phi_k C_k = \mathbf{\Phi C}, \tag{13}$$

where the coefficients are arranged in a column vector $\mathbf{C} = \{C_k\}$. The best coefficients are then derived from the variation principle (11), which gives the equation system

$$\sum_l \langle \Phi_k | H - \mathscr{E} \cdot 1 | \Phi_l \rangle C_l = 0, \tag{14}$$

for $k = 1, 2, 3, \ldots, n$. Introducing the Hamiltonian matrix $\mathbf{H}$ and the metric matrix $\mathbf{\Delta}$ having the elements

$$H_{kl} = \langle \Phi_k | H | \Phi_l \rangle, \qquad \Delta_{kl} = \langle \Phi_k | \Phi_l \rangle, \tag{15}$$

respectively, one may write (14) in matrix form:

$$(\mathbf{H} - \mathscr{E}\mathbf{\Delta})\mathbf{C} = \mathbf{0}. \tag{16}$$

Such a homogenous system of linear equations has a nontrivial solution if and only if the determinant of the coefficient vanishes, i.e.,

$$|\mathbf{H} - \mathscr{E}\mathbf{\Delta}| = 0. \tag{17}$$

This so-called secular equation has n roots $\mathscr{E}_0, \mathscr{E}_1, \mathscr{E}_2, \ldots$ and, using the variation principle, it is easily shown that they are *upper bounds* to the true eigenvalues $E_0, E_1, E_2, \ldots$ in order. The accuracy of the corresponding eigenfunctions is determined by generalizations (*11*) of the Eckart criterion.

Since the metric matrix $\mathbf{\Delta} = \langle \mathbf{\Phi} | \mathbf{\Phi} \rangle$ has the property $\mathbf{a}^\dagger \mathbf{\Delta a} = \langle \mathbf{\Phi a} | \mathbf{\Phi a} \rangle \geq 0$, it is positive semidefinite, which implies that its eigenvalues μ_k can never be negative. If one of the eigenvalues μ is vanishing, it indicates that there is a linear dependency $\mathbf{\Phi}\boldsymbol{\alpha} = 0$, and, if q eigenvalues μ_k are vanishing, there are q linear dependencies, and the set $\mathbf{\Phi}$ contains only $p = n - q$ linearly independent elements. In such a case, secular equation (17) is identically vanishing for all values of the variable $\mathscr{E}$, and the eigenvalue problem has to be reformulated. By considering the matrix $\mathbf{K} = \mathbf{H} - \mathscr{E}\mathbf{\Delta}$ and the associated "super secular equation" (*12*) $|\mathbf{K} - Z \cdot \mathbf{1}| = 0$, it is easily shown that (17) has to be replaced by the condition that the sum of all the principle minors of order p of the matrix $\mathbf{K} = \mathbf{H} - \mathscr{E}\mathbf{\Delta}$ should be vanishing. Exact linear dependencies are hence not hard to handle, at least not in principle.

If the set $\mathbf{\Phi}$ is linearly independent, one would anticipate that this problem would have disappeared entirely. In such a case, all of the eigenvalues $\mu_1 \geq \mu_2 \geq \mu_3 \geq \cdots \geq \mu_n > 0$ are necessarily positive, and the smallest one μ_n is usually referred to as the "measure of linear independency." If this number is smaller than the accuracy used in the computations, i.e., $\mu_n < 10^{-m}$, the set $\mathbf{\Phi}$ is linearly independent for all practical purposes, and one speaks of an "approximate linear dependency." The occurrence of this phenomenon seems to be one of the main difficulties of modern computations, (*13*) which is not so easily avoided.

The questions associated with the metric matrix $\mathbf{\Delta} = \mathbf{1} + \mathbf{S}$ are often referred to as the "nonorthogonality problem," and they are certainly essential in molecular structure calculations as well as in solid-state theory. It seems perhaps that this problem could be completely avoided by going over to an orthonormal basis $\boldsymbol{\varphi}$ having the property $\langle\boldsymbol{\varphi}|\boldsymbol{\varphi}\rangle = \mathbf{1}$ by means of a linear transformation $\boldsymbol{\varphi} = \mathbf{\Phi}\mathbf{A}$. For this purpose, one may use one of the standard procedures available: successive, symmetric, or canonical orthonormalization (*14*). However, if the original set $\mathbf{\Phi}$ has some approximate linear dependencies, the associated difficulties will nevertheless show up in the computations, since the coefficients in the orthonormal set $\boldsymbol{\varphi} = \mathbf{\Phi}\mathbf{A}$ will "blow up" and cause a corresponding loss of significant figures in the elements of the matrix $\langle\boldsymbol{\varphi}|H|\boldsymbol{\varphi}\rangle$. Even if the orthonormalization procedures are of essential theoretical importance, they will hence not automatically solve the numerical problems involved.

An evaluation of an approximate eigenfunction Ψ in a molecular structure calculation by means of the variation principle will hence involve the following steps:

(1) choice of a proper finite basis $\mathbf{\Phi}$ and elimination of exact and approximate linear dependencies;

(2) evaluation of the elements of the matrices $\mathbf{H} = \langle\mathbf{\Phi}|H|\mathbf{\Phi}\rangle$ and $\mathbf{\Delta} = \langle\mathbf{\Phi}|\mathbf{\Phi}\rangle$;

(3) evaluation of the approximate eigenvalues $\mathscr{E}_0$, $\mathscr{E}_1$, $\mathscr{E}_2$, . . . by solving secular equation (17) or the corresponding problem involving a sum of principal minors of order p;

(4) evaluation of the vectors $\mathbf{C}_0$, $\mathbf{C}_1$, $\mathbf{C}_2$, . . . corresponding to the different approximate eigenfunctions;

(5) analysis and interpretation of the results.

The details of some of these steps are further discussed below.

C. Use of Constants of Motion and Symmetry Properties to Simplify the Eigenvalue Problem

If there is a constant of motion Λ satisfying the relation $H\Lambda = \Lambda H$, the Schrödinger equation $H\Psi = E\Psi$ may often be essentially simplified. Let us assume that Λ is a normal operator, so that $\Lambda\Lambda^\dagger = \Lambda^\dagger\Lambda$, which has all its eigenvalues gathered in m different points in the complex plane: λ_1, λ_2, . . . , λ_m. These eigenvalues λ_k are often referred to as "quantum numbers," and they are used to classify the energy levels of the system. It is easily shown that, if Ψ is an exact eigenfunction to H it is also an exact eigenfunction to Λ or, in the case of a degenerate level E, may be chosen in that way. To each eigenvalue λ_k, there is associated a projector (*15*):

$$\mathcal{O}_k = \prod_{l \neq k} (\Lambda - \lambda_l \cdot 1)/(\lambda_k - \lambda_l), \tag{18}$$

which has the fundamental properties $\mathcal{O}_k^2 = \mathcal{O}_k$, $\mathcal{O}_k\mathcal{O}_l = 0$ for $k \neq l$, $1 = \sum_k \mathcal{O}_k$, $\Lambda\mathcal{O}_k = \lambda_k\mathcal{O}_k$, $\Lambda = \sum_k \lambda_k\mathcal{O}_k$, and $\mathcal{O}_k^\dagger = \mathcal{O}_k$. Since $\mathcal{O}_k$ is a polynomial in Λ, it is also a constant of motion satisfying the relation $H\mathcal{O}_k = \mathcal{O}_k H$. By means of the projectors $\mathcal{O}_1$, $\mathcal{O}_2$, . . . , $\mathcal{O}_m$, it is now possible to split the Hilbert space $\mathscr{H}$ into subspaces $\mathscr{H}_1$, $\mathscr{H}_2$, . . . , $\mathscr{H}_m$, which are mutually orthogonal and noninteracting with respect to H:

$$\mathcal{O}_k^\dagger\mathcal{O}_l = 0, \qquad \mathcal{O}_k^\dagger H\mathcal{O}_l = 0. \tag{19}$$

These projectors are as useful in the approximate calculations based on the variation principle. Let us assume that the finite basis $\mathbf{\Phi}$ is stable under Λ, so that $\Lambda\mathbf{\Phi}$ belongs to the space spanned by the set $\mathbf{\Phi}$. The subsets $\mathbf{\Phi}_k = \mathcal{O}_k\mathbf{\Phi}$ are then mutually orthogonal and noninteracting with respect to H for $k \neq l$, according to (18). Using the properties of the projectors, one gets finally

$$\langle\mathbf{\Phi}_k|H - \mathscr{E}\cdot 1|\mathbf{\Phi}_l\rangle = \delta_{kl}\langle\mathbf{\Phi}|(H - \mathscr{E}\cdot 1)\mathcal{O}_l|\mathbf{\Phi}\rangle, \tag{20}$$

which indicates that the original secular matrix has been split into block-diagonal form with each block characterized by a quantum number λ_k. This approach is particularly convenient in treating constants of motion of type spin and angular momenta (*16*).

Let us next consider the treatment of an atomic or molecular system having certain symmetry properties. All unitary constants of motion g satisfying the relations $Hg = gH$, $g^{\dagger}g = gg^{\dagger} = 1$, form together a group $G = \{g\}$, which is often referred to as the "unitary group of the Hamiltonian" and which describes the symmetry properties of the system. Such a group of order $|G|$ is characterized by its irreducible representations Γ^{α} of order f^{α}, and one has the completeness relations $\Sigma_{\alpha}(f^{\alpha})^2 = |G|$. The various symmetry functions are then automatically generated by the well-known projection and shift operators (*17*):

$$P^{\alpha}_{kl} = (f^{\alpha}/|G|) \sum_{g} \Gamma^{\alpha}_{kl}(g)g^{-1}, \tag{21}$$

which satisfy the relations $(P^{\alpha}_{kl})^{\dagger} = P^{\alpha}_{lk}$ and

$$P^{\alpha}_{kl}P^{\beta}_{mn} = \delta^{\alpha\beta}\ \delta_{kn}P^{\beta}_{ml}. \tag{22}$$

The quantities P^{α}_{kk} correspond to the projectors $\mathcal{O}_k$ treated above, and one has particularly the relation $e = \Sigma_{\alpha}\Sigma_k P^{\alpha}_{kk}$ instead of the resolution of the identity $1 = \Sigma_k \mathcal{O}_k$. By using this relation, it is easily shown that the exact eigenfunctions Ψ to H are either automatically symmetry-adapted or may be chosen in that way.

In an approximate calculation based on the finite set $\mathbf{\Phi}$, it is convenient to assume that this set is stable under the group $G = \{g\}$. Introducing the subsets $\mathbf{\Phi}^{\alpha}_{kl} = P^{\alpha}_{kl}\mathbf{\Phi}$ and using (21), one gets directly the result

$$\begin{aligned} &\langle \mathbf{\Phi}^{\alpha}_{kl} | H - \mathscr{E}\cdot 1 | \mathbf{\Phi}^{\beta}_{mn}\rangle \\ &\quad = \delta^{\alpha\beta}\ \delta_{ln}\langle \mathbf{\Phi} | H - \mathscr{E}\cdot 1 | \mathbf{\Phi}^{\beta}_{mk}\rangle \\ &\quad = \delta^{\alpha\beta}\ \delta_{ln}(f^{\alpha}/|G|)\sum_{g} \Gamma^{\alpha}_{mk}(g)\langle \mathbf{\Phi} | H - \mathscr{E}\cdot 1 | g^{-1}\mathbf{\Phi}\rangle, \end{aligned} \tag{23}$$

which shows that the original secular matrix splits into block-diagonal form after the various irreducible representations and after columns within each representation. The proper use of symmetry properties is hence essential to simplify the eigenvalue problem in every calculation of molecular structure.

D. The Configuration Space Constructed from One-Electron Functions; the Hartree–Fock Scheme and the Method of Superposition of Configurations

In treating a system of N electrons, it is often convenient to start from a one-electron Hilbert space that is spanned by a complete set $\{\psi_1, \psi_2, \ldots\}$ of one-electron functions or spin–orbitals $\psi_k = \psi_k(x)$, which are functions of a single electronic coordinate $x = (\mathbf{r}, \zeta)$. In analyzing a wave function Ψ of N coordinates $X = (x_1, x_2, \ldots, x_N)$, one may then expand Ψ in terms of $\{\psi_k\}$ after one coordinate at a time, which gives

$$\begin{aligned}\Psi &= \Psi(x_1, x_2, \ldots, x_N)\\ &= \sum_{k_1, k_2, \ldots, k_N} \psi_{k_1}(x_1)\psi_{k_2}(x_2) \cdots \psi_{k_N}(x_N) C_{k_1, k_2, \ldots, k_N}\end{aligned} \tag{24}$$

where each of the indices $k_1, k_2, \ldots, k_n$ runs over the set of integers $k = 1, 2, 3, \ldots$ independently of each other. However, an electronic wave function Ψ satisfies the relation $\Psi = \mathcal{O}_{\mathrm{AS}}\Psi$ according to (8), and, since the projector (7) shifts each spin–orbital product $\psi_{k_1}\psi_{k_2} \cdots \psi_{k_N}$ into a determinant, one obtains

$$\Psi = 1/N! \sum_{k_1, k_2, \ldots, k_N} |\psi_{k_1}\psi_{k_2} \cdots \psi_{k_N}| C_{k_1, k_2, \ldots, k_N} \tag{25}$$

where

$$|\psi_{k_1}\psi_{k_2} \cdots \psi_{k_N}| = \begin{vmatrix} \psi_{k_1}(x_1) & \psi_{k_2}(x_1) & \cdots & \psi_{k_N}(x_1) \\ \psi_{k_1}(x_2) & \psi_{k_2}(x_2) & \cdots & \psi_{k_N}(x_2) \\ & & \vdots & \\ \psi_{k_1}(x_N) & \psi_{k_2}(x_N) & \cdots & \psi_{k_N}(x_N) \end{vmatrix} \tag{25'}$$

is a so-called Slater determinant. Only the determinants having all indices $k_1, k_2, \ldots, k_n$ different are nonvanishing and, since each combination occurs $N!$ times with essentially the same result, one obtains finally

$$\Psi = \sum_K D_K C_K, \tag{26}$$

where $D_k = |\psi_{k_1}\psi_{k_2} \cdots \psi_{k_N}|$, and one should sum only over the "ordered configurations" K having $k_1 < k_2 < \cdots < k_N$. The result implies that the set of determinants $\{D_K\}$ is complete with respect to the N- electron space.

1. *The Hartree–Fock Scheme*

In the independent-particle model (IPM) developed originally by Bohr (*18*), each electron in a many-electron system moves in the field of the nuclei and the "average field" of all the other electrons. Somewhat later Hartree combined this idea with the one-electron Schrödinger equation, which he solved iteratively for each of the electrons in its motion in an estimated field, until the average field created by all of them agreed with the starting field, and the model became self-consistent. The exclusion principle was taken into account by permitting maximum two electrons with opposite spin in each orbital. Since Hartree focused his interest on the field involved, this approach became known as the self-consistent-field (SCF) scheme.

A much more profound derivation of the independent-particle model was given by Slater (*19*) and by Fock (*20*), who assumed that the wave function Ψ could be approximated by a single determinant D of type (25) built up from N spin orbitals or one-electron functions $\boldsymbol{\psi} = \{\psi_1, \psi_2, \ldots, \psi_N\}$ so that

$$\Psi \approx D = (N!)^{-1/2}|\psi_k(x_i)|. \tag{27}$$

Starting from the many-electron Hamiltonian (9) and using the variation principle (11), they could then derive the conditions for the best spin orbitals ψ_k.

For a nonsingular linear transformation $\psi_k = \Sigma_l \varphi_l \alpha_{lk}$, one obtains for the associated determinant $|\psi_k(x_i)| = |\varphi_l(x_i)||\alpha_{lk}|$, which implies that the determinant D is essentially invariant under such transformations. Without any loss of generality, one may hence choose the set $\boldsymbol{\psi}$ orthonormal, so that $\langle \boldsymbol{\psi}|\boldsymbol{\psi}\rangle = \mathbf{1}$ or

$$\langle \psi_k|\psi_l\rangle = \delta_{kl}. \tag{28}$$

Since $\langle D|D\rangle = 1$, one obtains then for the expectation value (12) of the many-electron Hamiltonian (9)

$$\begin{aligned}\langle H\rangle_{\mathrm{AV}} &= \langle D|H|D\rangle \\ &= H_{(0)} + \sum \left\langle \psi_i(1) \,\middle|\, \frac{p_1^2}{2m} - e^2 \sum_g \frac{Z_g}{r_{1g}} \,\middle|\, \psi_i(1)\right\rangle \\ &\quad + \frac{1}{2}\sum_{i,j} \left\langle \psi_i(1)\psi_j(2) \,\middle|\, \frac{e^2}{r_{12}}(1 - P_{12}) \,\middle|\, \psi_i(1)\psi_j(2)\right\rangle. \end{aligned} \tag{29}$$

This expression may be simplified if one introduces the Fock–Dirac density matrix (*20, 21*):

$$\rho(x_1,x_2) = \sum_k \psi_k(x_1)\psi_k^*(x_2), \tag{30}$$

which gives

$$\begin{aligned}\langle H\rangle_{\mathrm{AV}} &= e^2 \sum_{g<h} \frac{Z_g Z_h}{R_{gh}} + \frac{1}{2m}\int p_1^2 \rho(x_1, x_1')\Big|_{x_1'=x_1} dx_1 \\ &\quad - e^2 \sum_g Z_g \int \frac{\rho(x_1, x_1)}{r_{1g}}\, dx_1 \\ &\quad + \frac{e^2}{2}\int \frac{\rho(x_1, x_1)\rho(x_2, x_2) - \rho(x_1, x_2)\rho(x_2, x_1)}{r_{12}}\, dx_1\, dx_2 \end{aligned} \tag{31}$$

In the second term, we have used the convention that the operator p_1^2 works only on the unprimed coordinate x_1 and that one puts $x_1' = x_1$ before the integration.

In applying the variation principle $\delta\langle H\rangle_{\mathrm{AV}} = 0$, one observes that

$$\begin{aligned}\delta\langle D|H|D\rangle = \langle \delta D|H|D\rangle + \langle D|H|\delta D\rangle &= \langle \delta D|HD\rangle + \langle HD|\delta D\rangle\\ &= \langle \delta D|HD\rangle + \langle \delta D|HD\rangle^*\\ &= 2\,\mathrm{Re}\langle \delta D|H|D\rangle.\end{aligned}$$

Keeping all the spin orbitals $\psi_1, \psi_2, \ldots, \psi_N$ fixed except for ψ_k, which is given the small variation $\delta\psi_k = \epsilon_k\bar{\psi}_k$, where ϵ is a small complex number and the spin orbital $\bar{\psi}_k$ is chosen *orthogonal* to all the original spin orbitals, one obtains $\delta D = (N!)^{-1/2}|\psi_1,\psi_2, \ldots, \delta\psi_k, \ldots, \psi_N| = \epsilon(N!)^{-1/2}|\psi_1,\psi_2, \ldots, \bar{\psi}_k, \ldots, \psi_N| = \epsilon D_{\mathrm{se}}$, where the symbol D_{se} means that the original determinant D has been "singly excited" by changing ψ_k into $\bar{\psi}_k$. Choosing the phase of ϵ properly, one gets

$$\langle D_{\mathrm{se}}|H|D\rangle = 0, \tag{32}$$

the so-called Brillouin theorem (*22*). It is easily shown that it is not only a necessary but also sufficient condition for the validity of the relation $\delta\langle H\rangle = 0$ in the determinantal scheme. Varying all the spin orbitals ψ_k subject to the auxiliary conditions (28) with the Lagrangian multipliers λ_{kl}, one obtains finally the Hartree–Fock equations

$$H_{\mathrm{eff}}(1)\psi_k(1) = \sum_l \psi_l(1)\lambda_{lk}, \tag{33}$$

where the "effective Hamiltonian" or Fock operator has the form

$$H_{\mathrm{eff}}(1) = \frac{p_1^2}{2m} - e^2\sum_g \frac{Z_g}{r_{1g}} + e^2\int dx_2\,\frac{\rho(x_2, x_2) - \rho(x_2, x_1)P_{12}}{r_{12}} \tag{34}$$

and P_{12} is the "permutation operator" having the property $P_{12}f(x_1) = f(x_2)$, etc. Since $H_{\mathrm{eff}}^\dagger = H_{\mathrm{eff}}$, one has also the property $\boldsymbol{\lambda}^\dagger = \boldsymbol{\lambda}$ for the matrix $\boldsymbol{\lambda} = \{\lambda_{kl}\}$. Hence there is a unitary transformation $\mathbf{U}$, which brings the matrix $\boldsymbol{\lambda}$ to diagonal form $\boldsymbol{\epsilon} = \{\epsilon_k\delta_{kl}\}$, so that

$$\mathbf{U}^\dagger\boldsymbol{\lambda}\mathbf{U} = \boldsymbol{\epsilon}, \qquad \mathbf{U}^\dagger\mathbf{U} = \mathbf{U}\mathbf{U}^\dagger = 1. \tag{35}$$

Writing (33) in the form $H_{\mathrm{eff}}\boldsymbol{\psi} = \boldsymbol{\psi}\boldsymbol{\lambda} = \boldsymbol{\psi}\mathbf{U}\boldsymbol{\epsilon}\mathbf{U}^\dagger$, and introducing the substitution $\boldsymbol{\psi}' = \boldsymbol{\psi}\mathbf{U}$, one obtains $H_{\mathrm{eff}}\boldsymbol{\psi}' = \boldsymbol{\psi}'\boldsymbol{\epsilon}$, i.e.,

$$H_{\mathrm{eff}}(1)\psi_k'(1) = \epsilon_k\psi_k'(1), \tag{36}$$

which is closely analogous to the one-electron Schrödinger equation originally used by Hartree but also characterized by the exchange operator containing the permutation operator P_{12}. In this relation, the eigenvalues ϵ_k may be interpreted as spin-orbital energies related to the ionization potentials by means of Koopmans' theorem (*23*).

It is evident that the density matrix ρ is the fundamental quantity in the Hartree–Fock scheme, and Eqs. (36) are usually solved iteratively starting from an estimate of the wave functions ψ'_k and using the cycle

$$\begin{tikzcd} \psi'_k \arrow[r] & \rho \arrow[r] & H_{\text{eff}} \arrow[r] & \psi'_k \arrow[lll, bend left] \end{tikzcd} \qquad (37)$$

until the process becomes self-consistent, i.e., there is no change in the significant figures of the wave functions ψ'_k. Even if this process is convergent for most atomic and molecular systems, there are systems for which it is clearly divergent, and one must then use the general theory of iterative processes to determine the solution. A mathematical proof for the existence of solutions to the Hartree–Fock equations (36) was not given until quite recently, and then only in the case of general spin orbitals:

$$\psi(x) = \begin{Bmatrix} \psi_+(\mathbf{r}) \\ \psi_-(\mathbf{r}) \end{Bmatrix} = \psi_+(\mathbf{r})\alpha + \psi_-(\mathbf{r})\beta. \qquad (38)$$

The density matrix $\rho(x_1, x_2)$ is the kernel of an operator having the form

$$\rho = \sum_{k=1}^{N} |\psi_k\rangle\langle\psi_k| = |\psi\rangle\langle\psi|, \qquad (39)$$

and it is evident that it fulfills the fundamental relations

$$\rho^2 = \rho, \qquad \rho^\dagger = \rho, \qquad \mathrm{Tr}(\rho) = N. \qquad (40)$$

The projection operator associated with the Hartree–Fock manifold M_{HF} is the subspace of the one-electron Hilbert space spanned by the Hartree–Fock functions $\psi_1, \psi_2, \ldots, \psi_N$. The projector ρ has hence very simple properties.

2. *The MO–LCAO–SCF Approach*

In a molecular system, the solutions to the Hartree–Fock equations (33) are localized or delocalized depending on the conditions one puts on the matrix of the Lagrangian multipliers $\boldsymbol{\lambda} = \{\lambda_{kl}\}$. In the special case (36), the eigenfunctions ψ'_k are definitely extended over the entire molecule and have the character of *molecular orbitals* in the sense of Hund and Mulliken. In the MO–LCAO scheme, these molecular orbitals (MO) are built up by linear combinations (LC) from atomic orbitals (AO).

Let us start from a system of M atomic spin orbitals $\boldsymbol{\phi} = \{\phi_1, \phi_2, \ldots, \phi_M\}$, which are naturally associated with the atoms of the system and which have the metric matrix $\boldsymbol{\Delta}_\phi = \langle \boldsymbol{\phi} | \boldsymbol{\phi} \rangle$ in the one-electron space. We assume that $M \geq N$ and that there are no linear dependencies. Even if the set $\boldsymbol{\phi}$ is not complete, we assume that the approximate eigenfunctions ψ_k may be expanded in the form

$$\psi_k = \sum_{\mu=1}^{M} \phi_\mu C_{\mu k}, \tag{41}$$

and that $\boldsymbol{\psi} = \boldsymbol{\phi}\mathbf{c}$, where $\mathbf{c} = \{c_{\mu k}\}$ is an $M \times N$ rectangular matrix having the $N \times M$ adjoint matrix $\mathbf{c}^\dagger = \{c^*_{k\mu}\}$. The orthonormality condition $\langle \boldsymbol{\psi} | \boldsymbol{\psi} \rangle = \langle \boldsymbol{\phi}\mathbf{c} | \boldsymbol{\phi}\mathbf{c} \rangle = \mathbf{c}^\dagger \langle \boldsymbol{\phi} | \boldsymbol{\phi} \rangle \mathbf{c} = \mathbf{c}^\dagger \boldsymbol{\Delta} \mathbf{c} = \mathbf{1}$ gives the relation

$$\mathbf{c}^\dagger\, \boldsymbol{\Delta}\mathbf{c} = \mathbf{1}, \tag{42}$$

and the projector ρ may be expressed in the form

$$\rho = |\boldsymbol{\psi}\rangle\langle\boldsymbol{\psi}| = |\boldsymbol{\phi}\mathbf{c}\rangle\langle\boldsymbol{\phi}\mathbf{c}| = |\boldsymbol{\phi}\rangle\mathbf{c}\mathbf{c}^\dagger\langle\boldsymbol{\phi}| = |\boldsymbol{\phi}\rangle\mathbf{R}\langle\boldsymbol{\phi}|,$$

where

$$\mathbf{R} = \mathbf{c}\mathbf{c}^\dagger \tag{43}$$

is an $M \times M$ matrix having the elements $R_{\mu\nu} = \sum_k c_{\mu k} c^\dagger_{k\nu} = \sum_k c_{\mu k} c^*_{\nu k}$. The matrix $\mathbf{R}$ is called the charge and bond order matrix, and it was introduced by Coulson and Longuet-Higgins (*24*). The charge orders $R_{\mu\mu}$ and the bond orders $R_{\mu\nu}$ are important "molecular indices," which have been used, e.g., by the Pullmans (*25*) in their encyclopedic studies of the large molecules of importance in biochemistry. Using (42) and (43), one finds directly that $\mathbf{R}$ satisfies the fundamental relations

$$\mathbf{R}\,\boldsymbol{\Delta}\mathbf{R} = \mathbf{R}, \qquad \mathbf{R}^\dagger = \mathbf{R}, \qquad \mathrm{Tr}(\boldsymbol{\Delta}\mathbf{R}) = N, \tag{44}$$

which are analogous to (40). Relation (44) has in various forms been used for "population analysis" of the charge clouds of the N-electron system.

Substituting the expression $\rho = |\boldsymbol{\phi}\rangle\mathbf{R}\langle\boldsymbol{\phi}|$, i.e.,

$$\rho(x_1, x_2) = \sum_{\mu\nu} \phi_\mu(x_1) R_{\mu\nu} \phi^*_\nu(x_2), \tag{45}$$

into energy formula (31), one obtains

$$\begin{aligned}\langle H\rangle_{\mathrm{AV}} &= e^2 \sum_{g<h} \frac{Z_g Z_h}{R_{gh}} \\ &+ \sum_{\mu,\nu} R_{\mu\nu} \left\langle \phi_\nu(1) \,\middle|\, \frac{p_1^2}{2m} - e^2 \sum_g \frac{Z_g}{r_{1g}} \,\middle|\, \phi_\mu(1) \right\rangle \\ &+ \sum_{\mu,\nu} \sum_{\kappa,\lambda} R_{\mu\nu} R_{\kappa\lambda} \left\langle \phi_\nu(1)\phi_\kappa/(2) \,\middle|\, \frac{e^2}{r_{12}}(1 - P_{12}) \,\middle|\, \phi_\mu(1)\phi_\lambda(2) \right\rangle\end{aligned} \tag{46}$$

The problem is now to evaluate the best form of the rectangular matrix $\mathbf{c}$ in the quadratic matrix $\mathbf{R} = \mathbf{c}\mathbf{c}^\dagger$ subject to auxiliary condition (42), so that $\delta\langle H\rangle_{\text{AV}} = 0$. Using the same technique as in the general case, one obtains instead of (36) the homogeneous equation system

$$\sum_{\nu} \langle \phi_\mu(1) | H_{\text{eff}}^{(R)}(1) - \epsilon \cdot 1 | \phi_\nu(1) \rangle c_\nu = 0, \tag{47}$$

where the effective Hamiltonian is now given by

$$\begin{aligned} H_{\text{eff}}^{(R)}(1) &= (p_1^2/2m) - e^2 \sum_g (Z_g/r_{1g}) \\ &\quad + e^2 \sum_{\kappa,\lambda} R_{\kappa\lambda} \int dx_2 \, \frac{\phi_\lambda^*(2)\phi_\kappa(2) - \phi_\lambda^*(2)\phi_\kappa(1)P_{12}}{r_{12}}. \end{aligned} \tag{48}$$

Introducing the matrix elements

$$\begin{aligned} F_{\mu\nu} &= \langle \phi_\mu(1) | H_{\text{eff}}^{(R)}(1) | \phi_\nu(1) \rangle \\ &= \int \phi_\mu^*(1) \left\{ \frac{p_1^2}{2m} - e^2 \sum_g \frac{Z_g}{r_{1g}} \right\} \phi_\nu(1) \, dx_1 \\ &\quad + e^2 \sum_{\kappa,\lambda} R_{\kappa\lambda} \int\int dx_1 \, dx_2 \\ &\quad \times \frac{\phi_\mu^*(1)\phi_\lambda^*(2)\phi_\kappa(2)\phi_\nu(1) - \phi_\mu^*(1)\phi_\lambda^*(2)\phi_\kappa(1)\phi_\nu(2)}{r_{12}}, \end{aligned} \tag{49}$$

One may write Eqs. (47) in the form

$$\sum_{\nu} (F_{\mu\nu} - \epsilon \, \Delta_{\mu\nu}) c_\nu = 0. \tag{50}$$

The associated secular equation has M roots of which the N lowest correspond to occupied spin orbitals, whereas the M-N highest correspond to so-called virtual solutions. System (50) is usually solved by an iterative procedure based on an estimate of the rectangular matrix $\mathbf{c}$ and an iterative cycle of the type

$$\begin{tikzcd} \mathbf{c} \arrow[r] & \mathbf{R} = \mathbf{c}\mathbf{c}^\dagger \arrow[r] & \mathbf{F} \arrow[r] & \mathbf{c}, \arrow[lll, bend left] \end{tikzcd} \tag{51}$$

which is repeated until one obtains self-consistency. The first MO–LCAO scheme of this type was developed by Roothaan (*26*).

3. *The Symmetry Dilemma*

It has previously been pointed out that, if the many-electron Hamiltonian (9) has a normal constant of motion Λ or a symmetry group $G = \{g\}$, the eigenvalue problem of H may be essentially simplified by using the associated projectors $\mathcal{O}_k$ defined by (18) of P_{kk}^{α} defined by (21), respectively; for these projectors, we here introduce the common notation Q.

That implies that H has a new constant of motion, which satisfies the relations

$$HQ = QH, \qquad Q^2 = Q, \qquad Q^\dagger = Q. \tag{52}$$

Using the resolution of the identity, one can then easily prove that every exact eigenfunction Ψ to H either automatically satisfies the relation

$$Q\Psi = \Psi, \tag{53}$$

or in the case of a degeneracy may be chosen that way.

Since the variation principle (11) is equivalent with the eigenvalue problem (1), one usually takes it for granted that even approximate solutions Φ found by the variation principle should more or less automatically satisfy also the relation $Q\Phi = \Phi$. However, if variational freedom is restricted in some way, this is not necessarily the case, and the condition $Q\Phi = \Phi$ is then a "constraint" that will usually raise the variational energy $\langle H \rangle_{\mathrm{AV}}$.

This reasoning applies also to the Hartree–Fock scheme, in which the exact eigenfunction Ψ is approximated by a single Slater determinant D. It is remarkable that the condition $QD = D$ is self-consistent in the sense that, if the original estimate D fulfills this condition, the iterated solutions to the Hartree–Fock equations will also have this property. However, since the relation $QD = D$ is usually a constraint, it is by no means necessary for the determinant D associated with the absolute minimum of $\langle H \rangle_{\mathrm{AV}}$ to satisfy this relation—it may instead be a mixture of different symmetry types:

$$D = \sum_{\alpha} \sum_{k} D^{\alpha}_{kk}. \tag{54}$$

This situation may be characterized as the symmetry dilemma (*27*) of the Hartree–Fock (HF) scheme and related variational approaches: the absolute energy minimum corresponds to a mixture of various symmetry types, whereas a state of pure symmetry corresponds to a "local" minimum of higher energy. One should perhaps remember that this is a mathematical feature of the restricted variational principle, and that the "symmetry breakdown" of the approximate wave function associated with the absolute minimum hence has no deeper physical significance whatsoever.

If the condition $QD = D$ is superimposed on the solution D for a particular type of symmetry under consideration, one speaks of a restricted Hartree–Fock (RHF) scheme; the energy error in this approach in comparison to the exact energy is often referred to as the *correlation error:*

$$E_{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{HF}}, \tag{55}$$

and, for the ground state of an atomic or molecular system, it seems to be about 1% of the total energy, and hence an appreciable quantity. For a survey of the correlation problem, we refer elsewhere (*28*).

If the symmetry restrictions are removed, one gets successively lower and lower energies, until the absolute minimum is reached. In this unrestricted Hartree–Fock (UHF) scheme, one deals with general spin orbitals of type (38), and the final determinant D is a mixture of various symmetries according to (54). In such a scheme, one wonders whether one may restore the symmetry properties of the solution by considering the proper projection

$$D_{kk}^{\alpha} = P_{kk}^{\alpha} D. \tag{56}$$

It is evident that the wave function D_{kk}^{α} is not necessarily optimal with respect to the energy, since the projection is carried out after the variation. If one reverses the two procedures, one gets instead an alternative scheme.

4. *The Projected Hartree–Fock (PHF) Scheme*

In this approach, the total wave function Ψ is approximated by a projection Q of a single Slater determinant

$$\Psi = QD, \tag{57}$$

and this means that the relation $Q\Psi = \Psi$ is going to be automatically satisfied. Since any linear transformation $\boldsymbol{\psi} = \boldsymbol{\varphi\alpha}$ of the basic spin orbitals $\boldsymbol{\psi}$ gives the determinantal relation $|\boldsymbol{\psi}| = |\boldsymbol{\varphi}| \cdot |\boldsymbol{\alpha}|$, and further

$$Q|\boldsymbol{\psi}| = \{Q|\boldsymbol{\varphi}|\} \cdot |\boldsymbol{\alpha}|, \tag{58}$$

it is clear that wave function (57) is uniquely determined by the one-electron space spanned by the orbitals $\boldsymbol{\psi}$ and by its projector $\rho = |\boldsymbol{\psi}\rangle\langle\boldsymbol{\psi}|$. Hence the wave function $\Psi = QD$ is determined by a Fock–Dirac density matrix ρ satisfying relations (40), or by its matrix counterpart $\mathbf{R} = \mathbf{c}\mathbf{c}^{\dagger}$ in the MO–LCAO approach. To some extent, the PHF scheme is hence characterized by the same fundamental quantities as the ordinary HF scheme.

For the expectation value of the energy, one obtains directly by using (52)

$$\langle H \rangle_{\mathrm{AV}} = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\langle D | HQ | D \rangle}{\langle D | Q | D \rangle} = \mathscr{E}. \tag{59}$$

Putting $\langle H \rangle = A/B$, one gets for the variation that $\delta\langle H \rangle = (\delta A - \mathscr{E}\delta B)/B = 0$, i.e.,

$$\mathrm{Re}\langle \delta D | (H - \mathscr{E}) Q | D \rangle = 0. \tag{60}$$

Choosing the same type of variations of the basic spin orbitals $\boldsymbol{\Psi}$ as in deriving (32), one obtains

$$\langle D_{\text{se}}|(H - \mathscr{E})Q|D\rangle = 0, \tag{61}$$

which is a generalization of Brillouin's theorem to the PHF scheme. As in the ordinary HF scheme, one may then derive the generalized HF equations, if one observes that the many-electron Hamiltonian H has been replaced by the operator

$$\Omega = (H - \mathscr{E})Q, \tag{62}$$

which contains also "many-body interactions" introduced through the symmetry considerations (*29*).

In the PHF scheme, one should, of course, use general spin orbitals (GSO) of the type (38), since this approach renders two spatial functions—$\psi_+(\mathbf{r})$ and $\psi_-(\mathbf{r})$—per electron, i.e., $2N$ orbitals for an N-electron system, which means a great deal of variational freedom. So far, little work has been done in this area, and instead one has concentrated the efforts to some simplified versions of the PHF scheme using only N orbitals of either α or β type.

It is intuitively clear that, due to the Coulomb repulsion e^2/r_{12}, electrons should have a tendency to avoid each other. For electrons having parallel spins, this effect is partly covered by the Pauli exclusion principle and the associated "Fermi hole," and it is hence sufficient if the wave function satisfies antisymmetry condition (8). On the other hand, electrons having opposite spins may avoid each other simply by being in different parts of space, and this idea is basic for the scheme using "different orbitals for different spins" (DODS). So far, the strength of this approach has not yet been fully explored. In the "alternant-molecular orbital" (AMO) method, the DODS idea is utilized in a somewhat simplified form to study the correlation problem in alternant conjugated systems (*30*) with very encouraging results.

The most common argument against the PHF scheme is that the wave function is no longer a single Slater determinant but a sum of such determinants. It should be remembered, however, that a function $\Psi = QD$ is a *conceptual entity,* which is uniquely defined by a single determinant D and the associated Fock–Dirac density matrix ρ. The PHF scheme has hence preserved very much of the simplicity of the original independent-particle model.

5. *The MC–SCF Method*

In this scheme, one starts out from a set of M spin orbitals $\boldsymbol{\psi} = \{\psi_1, \psi_2, \ldots, \psi_M\}$, where $M > N$, and studies the $\binom{M}{N}$ configurations that span the subspace of the N-electron Hilbert space. In this approach, the wave function Ψ is approximated by a sum of the type

$$\Psi \approx \sum_K D_K C_K, \tag{63}$$

where the coefficients C_K are determined by the equations

$$\sum_L \langle D_K | H - \mathscr{E} \cdot 1 | D_L \rangle C_L = 0. \tag{64}$$

At the same time, one tries to lower the energy $\langle H \rangle_{\mathrm{Av}}$ by varying the M spin orbitals ψ_k until they become optimal (*31*). This approach leads to a multiconfigurational self-consistent-field (MC–SCF) scheme, which is a natural generalization of the HF method and is currently in frequent use also in molecular calculations. It should be observed also that this scheme is subject to the symmetry dilemma discussed above.

6. *The CI Method*

If the number M becomes larger and larger, the set $\boldsymbol{\psi} = \{\psi_1, \psi_2, \ldots, \psi_M\}$ becomes more and more complete with respect to the one-electron Hilbert space, and the optimization problem starts then to lose its meaning. Instead, expansion (63) becomes true "in the mean," and wave function Ψ is expressed as a "superposition of configurations." Using an early term from the theory of atomic and molecular spectroscopy, one speaks also of the "configurational interaction" (CI) method. In computational work, one is always using a truncated basis but, when M becomes large, the number of configurations

$$\binom{M}{N} = M!/N!(M - N)!, \tag{65}$$

becomes exceedingly large, and it becomes difficult to handle the equation system (64) and the associated secular equation $|H_{KL} - \mathscr{E}\delta_{KL}| = 0$ numerically. Even if the CI method is straightforward, it runs rather soon into computational troubles, which are further discussed below. The CI expansion

$$\Psi = \sum_K D_K C_K \tag{66}$$

converges in principle in the mean for any complete set $\boldsymbol{\psi}$ of one-electron functions but, if the set is not properly chosen—e.g., if one tries to describe the electronic structure of an atom situated in one place by means of a set of hydrogen-like orbitals situated in a different place—the convergence may be exceedingly slow. The convergence properties of (66) may be changed by going over to another complete set $\boldsymbol{\varphi}$ by means of a unitary transformation, $\psi = \varphi \mathbf{U}$. The question of which basis leads to the most rapid convergence of expansion (66) has led to the introduction of *natural spin orbitals,* and this and related problems are most conveniently discussed in connection with the question of how one interprets results of large-scale computations of wave functions in general.

E. Use of Density Matrices in Interpreting the Results—Natural Spin Orbitals

Quantum mechanics is a special case of quantum statistics and, in many cases, it is worthwhile to use concepts from this more general field in interpreting the results of quantum-mechanical calculations. In quantum statistics, a physical assembly is described by a density operator Γ and its kernel $\Gamma(X|X')$, which is often referred to as a density matrix. Here $X = (x_1, x_2, \ldots, x_N)$ and $X' = (x_1', x_2', \ldots, x_N')$ are different sets of electronic coordinates. The density operator is assumed to be self-adjoint and positive semidefinite, and it is usually normalized to unity so that

$$\operatorname{Tr} \Gamma = \int \Gamma(X|X)\, dX = 1,$$

where the symbol Tr (trace) indicates that one takes the diagonal sum of the associated matrix. If F is a general linear operator, one defines its expectation value with respect to the assembly through the relation

$$\langle F \rangle_{\mathrm{AV}} = \operatorname{Tr}\{F\Gamma\} = \int \overrightarrow{F_{\mathrm{op}}\Gamma(X}|X')|_{X'=X}\, dx, \tag{68}$$

where one has introduced the convention that F_{op} works only on the unprimed coordinate X in the density matrix $\Gamma(X|X')$, and that one puts $X' = X$ before the integration.

In treating operators of the type $F = \Sigma_i F_i$, $G = \Sigma_{i<j}\, G_{ij}$, etc., it is convenient to introduce the so-called reduced density matrices (*32*):

$$\gamma(x_1|x_1') = N \int \Gamma(x_1 x_2 x_3 \cdots |x_1' x_2 x_3 \cdots)\, dx_2\, dx_3 \cdots dx_N, \tag{69}$$

$$\Gamma(x_1 x_2|x_1' x_2') = \binom{N}{2} \int \Gamma(x_1 x_2 x_3 \cdots |x_1' x_2' x_3 \cdots)\, dx_3 \cdots dx_N, \tag{70}$$

of the first and second order, etc. For the expectation values of the one- and two-electron operators, one gets immediately

$$\left\langle \sum_i F_i \right\rangle_{\mathrm{AV}} = \int F_1\, \gamma(x_1|x_1')|_{x_1'=x_1}\, dx_1, \tag{71}$$

$$\left\langle \sum_{i<j} G_{ij} \right\rangle_{\mathrm{AV}} = \int G_{12} \Gamma(x_1 x_2|x_1' x_2')|_{x_i'=x_i}\, dx_1\, dx_2, \tag{72}$$

where one is using the same operator convention as in (68). The reduced-density matrices are hence excellent tools for evaluating expectation values, and we now use them to study the total energy.

1. *Evaluation of the Total Energy*

For the kinetic energy T of an electronic system, one gets directly

$$T = \left\langle \sum_i p_i^2/2m \right\rangle_{\mathrm{AV}} = (2m)^{-1} \int p_1^2\, \gamma(x_1|x_1')|_{x_1'=x_1}\, dx_1. \tag{73}$$

For the nuclear attraction energy L, one gets similarly

$$L = \left\langle -e^2 \sum_i \sum_g \frac{Z_g}{r_{ig}} \right\rangle_{\mathrm{AV}} = -e^2 \sum_g Z_g \int \frac{\gamma(x_1|x_1)}{r_{1g}}\, dx_1, \tag{74}$$

whereas the electronic Coulomb repulsion energy C is given by

$$C = \left\langle \sum_{i<j} \frac{e^2}{r_{ij}} \right\rangle_{\mathrm{AV}} = e^2 \iint \frac{\Gamma(x_1x_2|x_1x_2)}{r_{12}}\, dx_1\, dx_2. \tag{75}$$

The expectation value of the many-electron Hamiltonian (9) is hence given by

$$\begin{aligned}\langle H_{\mathrm{op}} \rangle_{\mathrm{av}} = e^2 \sum_{g<h} \frac{Z_g Z_h}{R_{gh}} &+ \frac{1}{2m} \int p_1^2 \gamma(x_1|x_1')\, dx_1 \\ &- e^2 \sum_g Z_g \int \frac{\gamma(x_1|x_1)}{r_{1g}}\, dx_1 + e^2 \iint \frac{\Gamma(x_1x_2|x_1x_2)}{r_{12}}\, dx_1\, dx_2.\end{aligned} \tag{76}$$

Here the first term—the nuclear repulsion energy—enters essentially as a constant. However, one has to observe that, for very large molecules, this term has to be carefully combined with the other terms in the energy expression, since otherwise one will obtain a result in the form of a difference between very large numbers associated with a loss in significant figures: in fact, for a polymer or a large biomolecule, there is a danger that the result may turn out to be ambiguous. This phenomenon is related to the fact that the so-called Madelung energy for an infinite ionic crystal is usually represented by a conditionally convergent sum.

The problem may be handled (*33*) by writing expression (76) as a sum of three terms $E = E_1 + E_2 + E_3$, where E_1 is the kinetic energy T given by (73), E_2 the correspondence to the "classical" Coulomb energy

$$\begin{aligned}E_2 = e^2 \sum_{g<h} \frac{Z_g Z_h}{R_{gh}} &- e^2 \sum_g Z_g \int \frac{\gamma(x_1)}{r_{1g}}\, dx_1 \\ &+ \frac{e^2}{2} \iint \frac{\gamma(x_1)\gamma(x_2)}{r_{12}}\, dx_1\, dx_2,\end{aligned} \tag{77}$$

where $\gamma(x_1) = \gamma(x_1|x_1)$ is the electron density, whereas E_3 is the quantum-mechanical part of the total Coulomb energy

$$E_3 = e^2 \iint \frac{\Gamma(x_1x_2) - \frac{1}{2}\gamma(x_1)\gamma(x_2)}{r_{12}}\, dx_1\, dx_2. \tag{78}$$

here $\Gamma(x_1x_2) = \Gamma(x_1x_2|x_1x_2)$ is the two-electron density. In the HF scheme, the E_3 term represents the so-called exchange energy, whereas in a more general scheme it contains both exchange and correlation effects.

2. *Natural Representations*

In order to study the density operators Γ in greater detail, we now introduce a basis $\mathbf{\Phi} = \{\Phi_k\}$ in the N-electron space, which is assumed to be orthonormal and complete, so that $\langle\Phi_k|\Phi_l\rangle = \delta_{kl}$ and $1 = \sum_k|\Phi_k\rangle\langle\Phi_k|$, i.e.,

$$\langle\mathbf{\Phi}|\mathbf{\Phi}\rangle = \mathbf{1}, \qquad 1 = |\mathbf{\Phi}\rangle\langle\mathbf{\Phi}|. \tag{79}$$

Using the preceding relation, one obtains

$$\Gamma = 1 \cdot \Gamma \cdot 1 = |\mathbf{\Phi}\rangle\langle\mathbf{\Phi}|\Gamma|\mathbf{\Phi}\rangle\langle\mathbf{\Phi}| = |\mathbf{\Phi}\rangle\mathbf{\Gamma}\langle\mathbf{\Phi}|,$$

i.e.,

$$\Gamma = \sum_{k,l} |\Phi_k\rangle\Gamma_{kl}\langle\Phi_l|, \qquad \Gamma_{kl} = \langle\Phi_k|\Gamma|\Phi_l\rangle. \tag{80}$$

Since the matrix $\mathbf{\Gamma} = \{\Gamma_{kl}\}$ is self-adjoint, there exists a unitary transformation $\mathbf{U}$ that brings it to diagonal form $\boldsymbol{\omega} = \{\omega_k\delta_{kl}\}$, so that

$$\mathbf{U}^\dagger\mathbf{\Gamma}\mathbf{U} = \boldsymbol{\omega}, \qquad \mathbf{\Gamma} = \mathbf{U}\boldsymbol{\omega}U^\dagger. \tag{81}$$

Introducing a new basis $\boldsymbol{\chi}$ through the unitary transformation $\boldsymbol{\chi} = \mathbf{\Phi}\mathbf{U}$, one gets directly that $\Gamma = |\mathbf{\Phi}\rangle\mathbf{\Gamma}\langle\mathbf{\Psi}| = |\mathbf{\Phi}\rangle\mathbf{U}\boldsymbol{\omega}\mathbf{U}^\dagger\langle\mathbf{\Phi}| = |\boldsymbol{\chi}\rangle\boldsymbol{\omega}\langle\boldsymbol{\chi}|$, i.e.,

$$\Gamma = \sum_k \omega_k|\chi_k\rangle\langle\chi_k| = \sum_k \omega_k\Gamma_k, \tag{82}$$

where the operators $\Gamma_k = |\chi_k\rangle\langle\chi_k|$ are projectors satisfying the relations $\Gamma_k^2 = \Gamma_k$, $\Gamma_k^\dagger = \Gamma_k$. Relation (82) is a *spectral resolution* of the density operator Γ and is called the natural representation of Γ; the basis $\boldsymbol{\chi} = \{\chi_k\}$ is referred to as the *natural basis* for Γ. Result (82) implies that the operator Γ may be considered as a superposition of the operators $\Gamma_k = |\chi_k\rangle\langle\chi_k|$ associated with the wave functions χ_k, with the weight factors ω_k satisfying the relations $0 \leq \omega_k \leq 1$ and $\sum_k \omega_k = 1$.

Treating the reduced-density matrices in a similar fashion, one finds that they have natural representations of the form

$$\gamma(x_1|x_1') = \sum_k n_k\chi_k(x_1)\, \chi_k^*(x_1'), \tag{83}$$

$$\Gamma(x_1x_2|x_1'x_2') = \sum_k m_k g_k(x_1x_2)\, g_k^*(x_1'x_2'), \tag{84}$$

where the one-electron functions $\chi_k(x_1)$ are called *natural spin orbitals*, whereas the two-electron functions $g_k(x_1x_2)$ are referred to as *natural geminals*. The latter play an important part in modern solid-state theory, particularly in the theory of superconductivity, but they have not yet been taken into more extensive use in molecular calculations, even if they are on the research programs of several groups.

3. *Quantum Mechanics as a Special Case*

If the sum (82) reduces to a single projector $\Gamma = |\chi\rangle\langle\chi|$, one speaks of a *homogeneous assembly*, which is characterized by a single wave function $\chi \equiv \Psi = \Psi(X)$. For the kernel of Γ, one obtains directly

$$\Gamma(X|X') = \Psi(X)\Psi^*(X'). \tag{85}$$

According to (69) and (70), the first- and second-order reduced-density matrices are then given by

$$\gamma(x_1|x_1') = N \int \Psi(x_1x_2x_3\cdots)\, \Psi^*(x_1'x_2x_3\cdots)\, dx_2\, dx_3\cdots dx_N, \tag{86}$$

$$\Gamma(x_1x_2\,|x_1'x_2') = \binom{N}{2} \int \Psi(x_1x_2x_3\cdots)\, \Psi^*(x_1'x_2'x_3\cdots)\, dx_3\cdots dx_N, \tag{87}$$

and they may hence be evaluated directly from the wave function Ψ. If this function is given in the form of a CI expansion (66), one can now show (*34*) that one obtains the *most rapid convergence* by introducing the "natural spin orbitals" χ_k, which diagonalize the first-order reduced-density matrix $\gamma(x_1|x_1')$. The corresponding eigenvalues n_k may be interpreted as "occupation numbers," and they satisfy the relations $0 \le n_k \le 1$ and $\Sigma_k n_k = N$. For two-electron systems, the natural spin orbitals lead also to considerable simplifications of the analytical form of the wave function (*35*).

Because of their importance for the convergence properties of the CI expansions, the natural spin orbitals are now a standard tool in molecular structure calculations; for a survey of the recent developments, we refer elsewhere (*36*).

4. *The Representability Problem*

It was early realized that, since one has the connection

$$\gamma(x_1|x_1') = \frac{2}{N-1} \int \Gamma(x_1x_2|x_1'x_2)\, dx_2, \tag{88}$$

the total energy $\langle H\rangle_{\mathrm{AV}}$ in expression (76) is essentially determined by the second-order density matrix $\Gamma(x_1x_2|x'_1x'_2)$. Using (76) and (88), one easily gets

$$\langle H_{\mathrm{T}}\rangle_{\mathrm{AV}} = e^2 \sum_{g<h} \frac{Z_gZ_h}{R_{gh}} + \int K_{12}\Gamma(x_1x_2|x'_1x'_2)\Big|_{x'_i=x_i} dx_1\, dx_2 \tag{89}$$

where

$$K_{12} = \frac{1}{N-1}\left\{\frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \sum_g Z_g\left(\frac{e^2}{r_{1g}} + \frac{e^2}{r_{2g}}\right)\right\} + \frac{e^2}{r_{12}}, \tag{90}$$

is a "reduced" two-electron Hamiltonian, which has been chosen symmetric in the indices 1 and 2. The question is now to what extent the variational principle $\delta\langle H\rangle_{\mathrm{AV}} = 0$ can be simplified by considering variations of the second-order density matrix $\Gamma(x_1x_2|x'_1x'_2)$ only.

A second-order density matrix $\Gamma(x_1x_2|x'_1x'_2)$ is said to be *Γ-representable* if it may be derived from a self-adjoint and positive-definite density operator Γ according to formula (70), and it is said to be *Ψ-representable* if it may be derived from a wave function Ψ according to formula (87). In both cases, one has a variation principle of the form

$$\langle H_{\mathrm{op}}\rangle_{\mathrm{AV}} \geq E_0, \tag{91}$$

but, if the second-order density matrix $\Gamma(x_1x_2|x'_1x'_2)$ is no longer representable, this inequality breaks down. The necessary and sufficient conditions for N-representability are currently being intensely studied (*37*), but since they are comparatively complicated this approach has so far had very little practical importance in atomic and molecular theory.

The main importance of the second-order density matrix $\Gamma(x_1x_2|x'_1x'_2)$ in the future will probably be in the "geminal description" of molecular systems, particularly since such an approach contains the one-electron scheme as a special case. It is also evident that the reduced-density matrices are excellent tools for interpreting the results of large-scale calculations leading to complicated wave functions.

III. Computational Problems in Molecular Studies

A. Calculation of the Energy Matrix; the Molecular Integral Problem

One of the key problems in molecular structure calculations is the evaluation of the total energy $\langle H\rangle_{\mathrm{AV}}$ according to (76). In the HF scheme, one starts instead from expression (31), but the problems are similar in nature. The kinetic energy T and the nuclear attraction energy L are given by formulas (73) and (74), respectively, and their evaluation represents usually no major computation problems. The difficulties are instead concentrated on the electronic Coulomb repulsion energy, given by

$$C = e^2 \int \int [\Gamma(x_1, x_2)/r_{12}]\, dx_1\, dx_2, \tag{92}$$

according to (75), and it is remarkable that so far no one has been able to utilize this simple formula for computational purposes. In the HF scheme, one has instead used the last term in expression (46):

$$C = \sum_{\mu,\nu} \sum_{\kappa,\lambda} R_{\mu\nu} R_{\kappa\lambda} \langle \phi_\nu(1)\phi_\kappa(2)|e^2/r_{12}(1 - P_{12})|\phi_\mu(1)\phi_\lambda(2)\rangle. \tag{93}$$

In the following, we use the notations

$$\begin{aligned} \langle \phi_\gamma(1)\phi_\kappa(2)|e^2/r_{12}|\phi_\mu(1)\phi_\lambda(2)\rangle &= \langle \gamma\kappa|e^2/r_{12}|\mu\lambda\rangle \\ &= e^2 \int \int \frac{\phi_\gamma^*(1)\phi_\kappa^*(2)\phi_\mu(1)\phi_\lambda(2)}{r_{12}}\, dx_1\, dx_2 \\ &= e^2 \int \int \frac{\phi_\gamma^*(1)\phi_\mu(1) \cdot \phi_\kappa^*(2)\phi_\lambda(2)}{r_{12}}\, dx_1\, dx_2 = (\gamma\mu|\kappa\lambda), \end{aligned} \tag{94}$$

where the last symbol $(\gamma\kappa|\eta\lambda)$, introduced by Mulliken, refers to the two charge densities involved in the integral. Substitution into (93) gives

$$C = \sum_{\mu,\gamma} \sum_{\kappa,\lambda} (R_{\mu\gamma} R_{\kappa\lambda} - R_{\lambda\gamma} R_{\kappa\mu})(\gamma\mu|\kappa\lambda). \tag{95}$$

In the evaluation of C, one encounters two problems: (1) the calculation of all the molecular integrals $(\nu\mu|\kappa\lambda)$, and (2) the summation over the four indices μ, ν, κ, and λ. It is clear that if one starts from M atomic spin orbitals ϕ_μ, the sum C is going to contain essentially M^4 different terms, even if the number may be slightly reduced by symmetry considerations, e.g., $(\mu\nu|\lambda\kappa) = (\nu\mu|\kappa\lambda)^*$. For real spin orbitals, one has further the symmetry relations $(\nu\mu|\kappa\lambda) = (\mu\nu|\kappa\lambda) = (\nu\mu|\lambda\kappa) = (\mu\nu|\lambda\kappa)$, etc. Since there are four atomic indices μ, ν, κ, and λ involved, the quantities $(\nu\mu|\kappa\lambda)$ may be classified as one-, two-, three-, and four-center integrals depending on the number of atomic centers occurring.

The first two-center exchange integral $(ab|ab)$ was evaluated by Sugiura (*38*) using the Neumann expansion of $1/r_{12}$ in elliptic coordinates. This approach dominated for a long time the calculation of all two-center integrals. The general problem was tackled by Kotani (*39*), who published extensive tables of auxiliary quantities to be used in the evaluation of molecular integrals.

If one starts from the Legendre expansion

$$1/r_{12} = \sum_{l=0}^{\infty} (r_{<}^{l}/r_{>}^{l+1})P_l(\cos\theta), \tag{96}$$

one may deduce explicit formulas for the calculation of one-, two-, three-, and four-center integrals (*40*), which involve fairly rapidly convergent infinite sums. The practical evaluation of the molecular integrals $(\nu\mu|\kappa\lambda)$ still turned out to be one of the most difficult and time-consuming problems in quantum-mechanical molecular studies.

The matrix elements H_{KL} of the Hamiltonian H or the effective Hamiltonian H_{eff} given by (34) or (48) are essentially calculated in the same way as the total energy. Once the matrices are found, one has also the problem of solving large homogeneous equation systems and the associated secular equations. Even if such problems can often be greatly simplified by the use of constant of motions and symmetry splitting by means of projectors, the remaining computational difficulties are still appreciable. Even if the mathematical methods for handling such problems were known in principle, the practical progress in molecular structure calculations has still been slow for more than two decades.

B. The Large-Scale Electronic Computers

The event that drastically changed development in the field of molecular calculations was the construction of large-scale electronic computers after the end of World War II. It was early realized by Mulliken and Slater in the United States and by Hartree and Boys in England that the new computers might provide a valuable tool for computational work in atomic, molecular, and solid-state theory. At the first international symposium at Shelter Island (*41*) in 1951, many of the participants were very enthusiastic about the new development, since the new machines based on punch cards, electronic vacuum tubes, and magnetic tapes offered such a tremendous increase in computational capacity. It is interesting to observe that today one considers those early machines highly primitive.

The electronic computers turned out to be very useful, even if their programming was cumbersome. Fundamental work on the bottleneck of molecular integrals $(\nu\mu|\kappa\lambda)$ was carried out in Mulliken's group in Chicago, particularly by Roothaan and Ruedenberg, and in Oxford by Barnett and Coulson to adapt the molecular problems to the new tool. In Cambridge, Boys solved the problem in his own way by introducing basic sets $\mathbf{\Phi}$ consisting of "Gaussians," for which the molecular orbitals $(\nu\mu|\kappa\lambda)$ may be evaluated analytically in closed form.

At the time of the Texas Conference (*42*) in 1955, quantum chemists were ready to recommend that every research group in the field have access to a large-scale electronic computer. In many countries, quantum chemists became pioneers in trying to obtain electronic computers as standard tools at universities and, in many places, for some time they took the general leadership of the overall computational effort. At the Boulder Conference (*43*) in 1958, it became evident that the new approach was exceedingly fruitful and that further development should be encouraged. Since a great deal of time and manpower went into the construction of computer programs, it seemed desirable to coordinate these efforts on an international basis, and—on the initiative of Harrison Shull—a special Quantum Chemistry Program Exchange (QCPE) was set up at the University of Indiana in Bloomington, which turned out to be highly valuable and which is still in full effect.

C. Some New Aspects of the Computational Work

With the development of electronic computers and their steadily increasing computational capacity, it became clear that one may have a completely new and different outlook on the numerical methods to be utilized in solving the given problems, and that simple schemes involving many computational steps may be preferable to complicated schemes involving only a few steps. For instance, in the beginning, it had been natural to assume that the atomic orbitals $\boldsymbol{\phi} = \{\phi_\mu\}$ were either atomic HF orbitals or analytic approximations to such orbitals in the form of Slater-type orbitals (STO), i.e., $r^{n-1}e^{-\alpha r}$ for $n = 1, 2, 3, \ldots$ multiplied by spherical harmonics $Y_{lm}(\theta, \varphi)$. However, Boys had pointed out that even a set of Gaussians

$$e^{-nr^2}, \qquad \text{for} \quad n = 1, 2, 3, \ldots, \tag{97}$$

is mathematically complete, that they could be successfully used in molecular calculations, and that one could approximate the atomic orbitals involved by proper spatial arrangements of Gaussians.

In numerical work, one soon learned also in practice that the infinite set of discrete hydrogen-like orbitals is by no means complete unless one includes also the continuum (*44*), and that it may be more practical to use the modified discrete set introduced already by Schrödinger (*45*).

One also learned about the existence of *approximate linear dependencies* (*46*), and some numerical studies showed that all basic sets based on powers, e.g., $x^{n-1}, r^{n-1}e^{-\alpha r}, e^{-nr}, e^{-nr^2}$, for $n = 1, 2, 3, \ldots$, very quickly become approximately linearly dependent, which means that, with increasing n, one has to increase the number of figures in the computations considerably in order to get a certain fixed number of significant figures in the final energy.

Computer methods for solving large-order secular equations were adapted also to molecular problems, and the classical Jacobi method for diagonalizing large matrices by means of repeated two-dimensional rotations was put into practical use—often in combination with Gieven's method for bringing a matrix to tridiagonal form (*47*). If only a few low-lying energy eigenvalues were needed, the partitioning technique (*48*) for handling matrices turned out to be of particular value.

In the years 1960–1975, one treated larger and larger molecules, and the results were reported in international periodicals associated with the field. Summaries and surveys were presented at a series of meetings: the Gordon Conferences (*49*), the Sanibel Symposia (*50*), and finally at the Congresses of the International Academy of Molecular Quantum Science (*51*). Even some of the molecular processes in biochemistry were treated quantum mechanically and, for surveys, we refer to the proceedings of the International Symposia in Quantum Biology (*52*) held under the auspices of the International Society of Quantum Biology.

IV. Some Trends in Current Molecular Structure Calculations

A. General Aspects

In this section, we briefly review some of the recent developments in large-scale molecular calculations as to molecular ground states and low-lying excited states, molecular geometry, and conformations, as well as intermolecular energy surfaces. There are essentially three different types of approaches, referred to as *ab initio*, semi-*ab initio*, and semiempirical calculations. The nomenclature is not entirely proper, since quantum mechanics represents the quintessence of about 150 years of experimental research, and so all calculations are essentially "empirical" in their origin.

The experimental discovery of the atomic numbers Z in the Rutherford atomic model of 1911 implies that one may discribe a molecule by means of the atomic numbers of its constituents. One can then study the many-electron Hamiltonian (9) in terms of the electronic charge e, the electronic mass m, and Planck's constant $\hbar$, and express all physical results in these quantities as units—a system often referred to as "atomic units." Since the Schrödinger equation $H\Psi = E\Psi$ in the Born–Oppenheimer approximation with fixed nuclei becomes a purely mathematical problem, one may speak of *ab initio* calculations, since no more experimental quantities have to be specified—until one introduces the nuclear kinetic energy (10), which requires knowledge of the nuclear masses M_g. Even some of the approximate methods of HF or MC–SCF type belong to this class, but it should be observed that, if one introduces a truncated basis, the particular choice of such a basis introduces a subjective element into the theory.

In the semi-*ab initio* calculations, it has usually been found convenient to introduce some mathematical approximations into the theory that depend on one or more parameters, which in turn are determined by other calculations. A typical example is the X_α method, in which the values of the parameter α in a molecular calculation may be determined from previous atomic calculations.

In the semiempirical methods, one starts again from the many-electron or one-electron Schrödinger equations but, instead of evaluating all the basic matrix elements in the theory, one may try to determine them from known experimental data in order to be able to predict other experimental results. This approach has essentially been used in treating very large molecules of conjugated type, in the Hückel or the Pariser–Parr–Pople (PPP) schemes, or in the CNDO, INDO, MINDO, etc., methods.

The molecular calculations of all three types are under continuous development as one gets better programs, new algorithms, and larger computers. There is hence a tendency for so-called standard programs to become obsolete, and it is always a good idea to check with QCPE and the leading research centers about the latest developments. It is also possible that some of the programs mentioned below will have become more or less obsolete when this text is published.

B. *Ab Initio* Calculations

In the *SCF methods* of the HF or the MC–SCF type, there are now standard programs available that will handle the electronic structure of molecules with 10–30 atomic centers and up to 200 basis functions. The molecular integrals are often evaluated by using Gaussian orbitals or by approximating given atomic functions by such orbitals. The most well-known programs based on Gaussians are the GAUSSIAN 70 developed by Pople and co-workers (*53*), IBMOL 6 developed by Clementi and Popkie (*54*), MUNICH developed by Diercksen and Kraemer (*55*), POLYATOM developed by Neumann *et al.* (*56*), ASTERIX developed by Veillard and co-workers (*57*), and MOLECULE developed by Almlöf (*58*).

Examples of recent SCF work of this type include molecular studies of the quanine–cytosine base pair and carbazole-2,4,7-trinitrofluorene by Clementi (*59*), of lithium halides in water clusters by Clementi and Popkie (*60*), of metal complexes by Veillard (*61*), of porphine by Almlöf (*62*), of TCNQ with ions by Johansen (*63*), of chlorpromazine and promazine by Popkie and Kaufman (*64*), and of various ions in water clusters by Lie, Clementi, and Yoshimine (*65*).

In the SCF treatment of the electronic structure of molecules, one has studied closed as well as open shells—the latter usually handled by a modification of Roothaan's decoupling scheme (*66*). A well-known program is ALCHEMY, which was developed by Bagus and co-workers (*67*) for linear molecules. For more general molecules, there is the special program MOLECULE-ALCHEMY, which is a combination of Almlöf's MOLECULE and Bagus' ALCHEMY put together by Bagus and Wahlgren (*68*). In some cases, one has had trouble with the convergence of the SCF schemes, and these questions have been studied by several authors (*69*).

There has also been considerable progress in the development of MC–SCF methods, particularly by Das and Wahl (*70*) and Hinze (*71*). A recent application of this approach to the NO_2 molecule has been carried out by Gillespie *et al.* (*72*).

A different type of SCF method has been developed by Goddard and co-workers (*73*), who have studied molecular valence bond (VB) functions, which are generalizations of the original Heitler–London function. Since they are many-determinantal functions, the VB–SCF scheme is related to the MC–SCF scheme, but it has its own difficulties associated with the use of nonorthogonal atomic orbitals. A recent application to the O_3 molecule may serve as a typical example (*74*).

In the *CI method* based on the use of a truncated set of M molecular orbitals $\boldsymbol{\psi} = \{\psi_i\}$, which in turn are formed by linear combinations of atomic orbitals $\boldsymbol{\phi} = \{\phi_\mu\}$, one is usually faced with four major problems:

(1) evaluation of molecular integrals $(\nu\mu|\kappa\lambda)$ over the atomic orbitals;

(2) transformation of the integrals to molecular orbital basis—the so-called four-index transformation $(\nu\mu|\kappa\lambda) \to (ij|kl)$;

(3) the evaluation of the elements H_{ij} of the Hamiltonian matrix $\mathbf{H}$;

(4) the diagonalization of $\mathbf{H}$ and the evaluation of its eigenvectors.

As in the SCF methods, the problem of the evaluation of the molecular integrals $(\nu\mu|\kappa\lambda)$ is now usually handled by using Gaussians, and it is no longer considered a bottleneck. A more serious problem is the second step, involving the four-index transformation, which has been treated and simplified by Yoshimine, by Diercksen, and by Shavitt (*75*), who have shown that, by convenient combination of terms, this M^8 process may be reduced to a $4M^5$ process. As to the diagonalization of $\mathbf{H}$, some new aspects have been developed (*76*), and, by this approach, one seems able to handle up to 10,000 configurations without too much difficulty.

In the direct CI method developed for molecular systems by Roos (*77*), one tries to eliminate step 3 and, in this way, one can handle up to 100,000 configurations.

As an example of recent molecular applications of the CI method, the treatment of C_6H_6 by Shavitt (*78*) involving 2636 configurations should be mentioned. A study of the He–He interaction by Liu and McLean (79) involves 8000 configurations. The investigations of the molecular complexes Li^+–H_2O, F^-–H_2O, and $(H_2O)_2$ by Diercksen, Kraemer, and Roos (*80*) involve 10,101, 36,204, and 56,268 configurations, respectively.

The technique for evaluating intermolecular potential energy surfaces has essentially followed the development of the general quantum-mechanical methods for treating molecular systems and for a survey of the field up to 1970, we refer to the proceedings from the Santa Cruz Conference (81). During the last few years, the CI method has been extensively used to tackle this problem. Some scattering surfaces of particular importance in astrophysics (HCN and HCO^+) have been studied by Diercksen and co-workers (*82*). At the 1977 Sanibel Conference, Diercksen mentioned some results on He + CO, Li + CO, CNH, N_2H^+, and HCO^+, and reported that an energy surface calculation involving 30–40 points would take about 60 hours computing time on the IBM 360/91 available in Munich.

In large-scale computations of this type, there is a great deal of experience in using various types of one-electron basis sets. For a study of the general properties of and the contractions of such sets, we refer to some reviews (*83*) and to the work by Dunning (*84*) and by Pople (*85*) *et al.* during the years 1971–1972. In using larger and larger basis sets, we have become increasingly aware of the existence of approximate linear dependencies mentioned earlier (*13*), and for a recent survey we refer to McLean (*86*). At the New Orleans Congress and the Sapporo Conference in 1976, Hall (*87*) reemphasized the importance of this problem for molecular calculations based also on relatively small basis sets of a certain type, and the danger of losing significant figures in the final results unless things are handled properly. In conclusion, it should be mentioned that the Sapporo group (*88*) has worked out a new method for treating the nonorthogonality problem, which so far has given very promising results.

For a more detailed review of the development of the molecular computer programs in the years 1972–1975, the reader is referred to a paper and a book by Schaefer (*89*), and as to the development in the years 1976–1979, to the references given in the "Note Added in Proof" on p. 311.

C. Semi-*ab Initio* Calculations

In the original HF scheme, the handling of the Coulomb operator in (34)

$$V_{\mathrm{op}}(x_1) = e^2 \int dx_2[\rho(x_2, x_2) - \rho(x_2, x_1)P_{12}]/r_{12} \tag{98}$$

represents a comparatively difficult problem because of the occurrence of the exchange operator P_{12}. In 1951, Slater (*90*) suggested that this operator should be approximated by a local operator for which one easily obtains the form (*91*):

$$
\begin{aligned}
V(x_1) &= 2e^2 \int \frac{\Gamma(x_1x_2)/\gamma(x_1)}{r_{12}}\, dx_2 \\
&= e^2 \int dx_2 \, \frac{\rho(x_2,x_2) - \rho(x_2,x_1)\rho(x_1,x_2)/\rho(x_1,x_1)}{r_{12}}, \qquad (99)
\end{aligned}
$$

where the second term in the last expression represents a local approximation of the exchange operator. By considering the free-electron model, Slater could further find an even simpler expression going from

$$V_{\text{ex}}(x_1) = -e^2 \int dx_2 \, \frac{\rho(x_2,x_1)\rho(x_1,x_2)/\rho(x_1,x_1)}{r_{12}} \qquad (100)$$

to

$$V_{\text{ex}}^{\uparrow}(x_1) = -6\alpha[(3/4\pi)\rho^{\uparrow}(x_1)]^{1/3}, \qquad (101)$$

where the vertical arrow indicates that the exchange potential is going to work only on electrons having this type of spin, and α is a coefficient that, according to Gaspar *(92)* should have the value $\alpha = 2/3$. This new SCF approach has been called the X_α method (X = exchange) and we note that, whereas the use of the form (100) leads to an approximation of the original HF scheme, the X_α method contains an adjustable parameter α, which makes it possible to incorporate not only exchange but also correlation effects into the theory; for more details, we refer to a survey paper by Slater *(93)*. In practical applications, the X_α method turns out to be an excellent "orbital generator" for all the SCF schemes and, in addition, gives theoretical results in very good agreement with experiments in the study of the electronic structure of molecules and crystals when starting from the atomic α values involved.

In numerical work, the X_α method is often combined with the multiscattering (MS) wave method to obtain further simplification of the scheme. In the MSX_α scheme, each atom is surrounded by a polyhedron as in the cellular method for treating a solid. In the so-called muffin tin model, each polyhedron is approximated by a sphere, i.e., each atom is surrounded by a sphere and the entire molecule is enclosed in an outside sphere. In some applications, even "overlapping spheres" have been used with excellent results. For more details, the reader is referred to a recent survey by Johnson *(94)*.

The X_α method has been used extensively to predict excitation energies and transition probabilities in various molecules to be compared with the results of the ESCA experiments, using "electron spectroscopy for chemical analysis," carried out by Siegbahn and co-workers *(95)*. In this connection, Slater *(96)* developed the hyper-Hartree–Fock (HHF) method based on density matrices and energy expressions of type (76) and involving also noninteger occupation numbers n_k. In treating electronic excitations from initial (i) to final (f) states, Slater considered particularly the "transition states" with half-integer occupation numbers: $n_i = n_f = \frac{1}{2}$. Later these ideas were refined and put into closer connection with the ordinary HF method by Goscinski (*97*). His transition operator method gives theoretical results in excellent agreement with experiment.

Large-scale calculations by the X_α method on inorganic metal complexes by Larsson (*98*) have given some very interesting results and have, among other things, given some new interpretations of the so-called satellites in the ESCA spectra. Other recent applications include a study of TCNQ–TTF by Herman and co-workers (*99*) and an investigation of the phenomena of catalysis and biocatalysis by Johnson and Messmer (*100*).

D. Semiempirical Methods

In all the molecular *ab initio* calculations based on the HF, the MC–SCF, the CI, or the VB methods, one is always dealing with a large number of *energy matrix elements* associated with the effective or the many-electron Hamiltonian, which are evaluated numerically starting from some given basis set. In the semiempirical methods, these matrix elements are not evaluated from "first principles" but are instead considered as parameters determined from experimental data available. The theory is then used to predict results from other experiments and becomes hence essentially a tool for correlating one set of empirical data with another.

As a rule, one tries to keep the number of parameters as small as possible, since one usually does not have too many experimental data to start from. This is often accomplished by considering some of the energy matrix elements as large and essential, whereas others are considered small and are entirely neglected. This gives a certain amount of ambiguity to the semiempirical schemes, and this means that even their most successful applications are often considered with some suspicion by the *ab initio* people, who are unable to reproduce their fine results. There is no question, however, that the semiempirical schemes are highly useful for predicting the outcomes of experiments within the limited field for which they are constructed and that they provide valuable guidance for experimentalists in looking for his next results. At the same time, the limitations are severe, and the current trend is to use the semiempirical methods only on molecular systems that are so large that they cannot conveniently be handled by the *ab initio* programs. However, many molecular scientists are certainly looking for a revival of the semiempirical approach.

We start by reviewing part of the historical development. One of the first successful semiempirical schemes was developed in connection with Slater's theory (*101*) for complex atomic spectra, which contained certain F and G integrals that could be evaluated from first principles. However, it was quickly realized that the theory would give much better results in comparison with experience if the F and G integrals were determined by certain empirical results related directly to experiments (*102*). Similarly, in the theory of molecular spectra dealing with the vibrational and rotational properties and the vibrational–rotational couplings, it was learned that the theory would give much better results if one did not determine the fundamental parameters A, B, C, . . . involved from first principles but from available experiments; for details, the reader is referred to Herzberg's book (*103*). The success of these semiempirical approaches has still to be explained.

In the study of the Kekule structures of benzene by quantum-mechanical methods, the principle of "chemical resonance" was early discovered (*104*). In the hands of Pauling (*105*), the theory of resonance between various chemical structures was later extensively developed essentially as a semiempirical form of the VB theory. It is still considered one of the most powerful tools for predicting molecular structures and properties without going into large-scale computations, but the reasons for its success remain to be fully understood.

Most of the conventional semiempirical methods for studies of molecular structure are modifications of the HF scheme, in which the matrix elements $F_{\mu\nu}$ of the effective Hamiltonian are considered as parameters to be determined from empirical data available. All of these schemes go back to Hückel's fundamental investigations (*106*) of the properties of the π electrons in aromatic molecules and conjugated systems, which led to a new understanding of their basic features. In the Hückel scheme, certain simple rules for the determination of the parameters $\alpha_\mu = F_{\mu\mu}$ and $\beta_{\mu\nu} = F_{\mu\nu}$ were developed. The Hückel scheme has been used extensively by the Pullmans in their encyclopedic studies of the conjugated systems in biochemistry (107).

In order to generalize this approach to include *all* the electrons of a molecular system, Hoffman (*108*) later developed the so-called extended Hückel method (EHM), which has turned out to be of essential practical importance. In retrospect, it is interesting to observe that the studies of the symmetry properties involved in this method form the background for the development of the famous Woodward–Hoffman rules (*109*) in the theory of organic reactions.

In an attempt to include certain correlation effects in the Hückel scheme, the PPP method (*110*) was developed. In connection with the treatment of those matrix elements in the scheme that are so small that they should be neglected, they introduced the idea of "zero-differential overlap" between the atomic orbitals involved. It was later shown (*111*) that this feature could be rather well explained if one introduced symmetrically orthonormalized atomic orbitals defined by the formula $\varphi = \phi\langle\phi|\phi\rangle^{-1/2}$. The same principle is fundamental also in the modifications of this approach known as the CNDO (complete neglect of differential overlap) method, the INDO (intermediate neglect of differential overlap) method, and the MINDO (modified intermediate neglect of differential overlap) method. All of these methods have been used to study the electronic structure of very large molecules, particularly those of importance in molecular biology.

It should be observed that also the CNDO, INDO, MINDO, etc., methods require large-scale calculations but that the computer times involved are one or two orders of magnitude less than in the corresponding *ab initio* calculations.

E. Further Computer Developments

Since 1950, there have been at least three generations of electronic computers used in molecular calculations: those based on punch cards and vacuum tubes, those based on transistors and solid-state circuits, and finally those based on minicircuits. It is characteristic that the minicomputers of today have almost the same speed and capacity as the giant computers of two generations ago at a very small fraction of the cost. The main limitation for future developments seems to be only that any electric signal (at the speed of light) travels at most 30 cm in 1 nsec (10^{-9} sec). Integrated circuits with solid-state components have certainly revolutionized the computer industry.

New minicomputers are now being developed that approach the speed and capacity of the large computers of one generation ago, at a purchase price some research groups may be able to afford. The use of such minicomputers in molecular calculations has recently been studied by Schaefer (*112*). It is evident however, that the existence of many such group minicomputers on a campus may threaten the existence of the large all-around university computer, and the question is whether such a development is desirable.

It has now also become evident that a fourth generation of computers—microcomputers—is under rapid development. They are going to have much larger and faster memory capacities than ever before, based on the use of magnetic bubbles, the Williams tube, and on silicon crystals. It is certainly still too early to forecast future developments, but it seems likely that many of the present large-scale computers at universities are going to be updated by having a series of fast microcomputers incorporated in their configurations. It seems rather certain that molecular structure calculations will benefit greatly from developments in the computer field over the next decade.

V. Some Current Problems and Aspects for the Future

A. Perturbation Theory and Partitioning Technique

In principle, the exact solution to the Schrödinger equation $H\Psi = E\Psi$ may be provided by the CI method, provided one used a complete discrete basis and could handle the linear equation system (64) and the associated secular equation of infinite order. It is clear that these problems are by no means trivial, and the question is whether one could obtain the solution in some "closed form" that would be easier to handle.

For this purpose, we use the partitioning technique (*113*). Let us introduce a normalized reference function φ having the property $\langle\varphi|\varphi\rangle = 1$, and let us consider the Schrödinger equation $H\Psi = E\Psi$ subject to the boundary condition $\langle\varphi|\Psi\rangle = 1$, which renders the energy spectrum $\{E\}_\varphi$. Let us further introduce the projector $O = |\varphi\rangle\langle\varphi|$ and the projector $P = 1 - O$ for the complement, satisfying the relations

$$\begin{gathered} O^2 = O, \qquad O^\dagger = O, \qquad \operatorname{Tr} O = 1, \\ P^2 = P, \qquad P^\dagger = P, \qquad OP = PO = 0, \qquad O + P = 1. \end{gathered} \tag{102}$$

One has further $O\varphi = \varphi$, $O\Psi = \varphi$, and finally,

$$P\varphi = 0, \qquad P\Psi = \Psi - \varphi = \phi. \tag{103}$$

In order to proceed, it is convenient to introduce a complex variable $\mathscr{E}$ and to consider the *inhomogenous* equation

$$(H - \mathscr{E})\Psi_{\mathscr{E}} = a\varphi, \qquad \langle\varphi|\Psi_{\mathscr{E}}\rangle = 1, \tag{104}$$

where the second relation is an auxiliary condition analogous to our previous boundary condition, which determines the constant a. Multiplying the first relation on the left by φ, one gets directly $a = \langle\varphi|H - \mathscr{E}|\Psi_{\mathscr{E}}\rangle = \langle\varphi|H|\Psi_{\mathscr{E}}\rangle - \mathscr{E} = \mathscr{E}_1 - \mathscr{E}$, where we have introduced the notation

$$\mathscr{E}_1 = \langle\varphi|H|\Psi_{\mathscr{E}}\rangle. \tag{105}$$

Letting P work to the left of the first relation (104), one gets further $P(H - \mathscr{E})\Psi_{\mathscr{E}} = 0$, since $P\varphi = 0$, and $(\mathscr{E} - PH)\Psi_{\mathscr{E}} = \mathscr{E}\varphi$, which gives the formal solution

$$\Psi_{\mathscr{E}} = [1 - (PH/\mathscr{E})]^{-1}\varphi = W_{\mathscr{E}}\varphi, \tag{106}$$

where $W_{\mathscr{E}} = [1 - (PH/\mathscr{E})]^{-1}$ is referred to as the "wave operator." Introducing this solution into (105), one obtains directly

$$\mathscr{E}_1 = \langle\varphi|HW_{\mathscr{E}}|\varphi\rangle = \langle\varphi|H[1 - (PH/\mathscr{E})]^{-1}|\varphi\rangle \equiv f(\mathscr{E}), \tag{107}$$

where $\mathscr{E}_1 = f(\mathscr{E})$ is a function to be further investigated. In the case when $\mathscr{E}_1 = \mathscr{E} = E$, i.e., when $E = f(E)$, one has $a = \mathscr{E}_1 - \mathscr{E} = 0$, and it is easily shown that the associated function

$$\Psi_E = [1 - (PH/E)]^{-1}\varphi \tag{108}$$

satisfies the original Schrödinger equation $(H - E)\Psi_E = 0$ with the boundary condition $\langle\varphi|\Psi_E\rangle = 1$, and vice versa. Hence the eigenvalue problem has been transformed into the "algebraic" form

$$E = f(E), \tag{109}$$

the only difference being that all multiple eigenvalues are now reduced to single ones as is shown below.

In dealing with the wave operator $W_{\mathscr{E}}$ and the operators derived from it, it is often useful to utilize the simple operator identity

$$(1 - RS)^{-1}R \equiv R(1 - SR)^{-1}, \tag{110}$$

which is easily proven by multiplying (110) to the left by $(1 - RS)$ and to the right by $(1 - SR)$, which gives $R(1 - SR) \equiv (1 - RS)R$. Using this identity and relations (106) and (105), one obtains for real $\mathscr{E}$

$$\begin{aligned} f'(\mathscr{E}) &= -\left\langle\varphi\left|H\left(1 - \frac{PH}{\mathscr{E}}\right)^{-1}\frac{PH}{\mathscr{E}^2}\left(1 - \frac{PH}{\mathscr{E}}\right)^{-1}\right|\varphi\right\rangle \\ &= -\left\langle\frac{PH}{\mathscr{E}}\left(1 - \frac{PH}{\mathscr{E}}\right)^{-1}\varphi\left|\frac{PH}{\mathscr{E}}\left(1 - \frac{PH}{\mathscr{E}}\right)^{-1}\varphi\right.\right\rangle \\ &= -\left\langle\frac{PH}{\mathscr{E}}\Psi_{\mathscr{E}}\left|\frac{PH}{\mathscr{E}}\Psi_{\mathscr{E}}\right.\right\rangle \\ &= -\langle\Psi_{\mathscr{E}} - \varphi|\Psi_{\mathscr{E}} - \varphi\rangle \\ &= -\langle\phi_{\mathscr{E}}|\phi_{\mathscr{E}}\rangle < 0, \end{aligned} \tag{111}$$

provided that the last integral exists, which means that $\mathscr{E}$ is somewhere in the "discrete" part of the spectrum. The fact that the derivative is negative has an important consequence.

Putting $\mathscr{E} = E + \varepsilon$ and $\mathscr{E}_1 = E + \varepsilon_1$, and using Lagrange's mean-value formula $f(\mathscr{E}) = f(E + \varepsilon) = f(E) + \varepsilon f'(E + \theta\varepsilon)$, where $0 \leq \theta \leq 1$, one obtains directly

$$\varepsilon_1 = \varepsilon f'(E + \theta\varepsilon), \tag{112}$$

which shows that ε and ε_1 have different signs: if $\mathscr{E}$ is an upper bound to E, $\mathscr{E}_1 = f(\mathscr{E})$ is going to be a lower bound, and vice versa. Hence there is at least one true eigenvalue E situated between $\mathscr{E}$ and $\mathscr{E}_1$, and the function $\mathscr{E}_1 = f(\mathscr{E})$ has been referred to as the "bracketing function." It is evident that it offers a new approach for finding upper and lower bounds to the energy eigenvalues E.

If one introduces the function $F(\mathscr{E}) = \mathscr{E} - f(\mathscr{E})$, the eigenvalue problem (109) may be presented in the simple form $F(E) = 0$. For the derivative, one gets directly, by using (111),

$$F'(\mathscr{E}) = 1 - f'(\mathscr{E}) = 1 + \langle \phi_{\mathscr{E}} | \phi_{\mathscr{E}} \rangle = \langle \Psi_{\mathscr{E}} | \Psi_{\mathscr{E}} \rangle > 1, \tag{113}$$

which implies that all roots to the equation $F(E) = 0$ must be single roots, since $F'(E) \neq 0$, and this proves the theorem quoted in connection with (109). It is also clear that one could try to solve the equation $F(E) = F(\mathscr{E} - \varepsilon) \approx F(\mathscr{E}) - \varepsilon F'(\mathscr{E}) \approx 0$ by the Newton–Raphson formula $\mathscr{E}_{\mathrm{NR}} = \mathscr{E} - \varepsilon$ or

$$\mathscr{E}_{\mathrm{NR}} = \mathscr{E} - \frac{F(\mathscr{E})}{F'(\mathscr{E})} = \mathscr{E} - \frac{\mathscr{E} - f(\mathscr{E})}{1 - f'(\mathscr{E})} = \mathscr{E} - \frac{\mathscr{E} - \mathscr{E}_1}{\langle \Psi_{\mathscr{E}} | \Psi_{\mathscr{E}} \rangle}, \tag{114}$$

which may serve as a basis for a well-known second-order iteration procedure. It is interesting to observe that this quantity is identical to the variational energy $\mathscr{E}_{\mathrm{VAR}}$. Using inhomogeneous equation (104) in the form $(H - \mathscr{E})\Psi_{\mathscr{E}} = (\mathscr{E}_1 - \mathscr{E})\varphi$ and the intermediate normalization $\langle \varphi | \Psi\mathscr{E} \rangle = 1$, one gets directly

$$\begin{aligned} \mathscr{E}_{\mathrm{VAR}} &= \frac{\langle \Psi_{\mathscr{E}} | H | \Psi_{\mathscr{E}} \rangle}{\langle \Psi_{\mathscr{E}} | \Psi_{\mathscr{E}} \rangle} = \mathscr{E} + \frac{\langle \Psi_{\mathscr{E}} | H - \mathscr{E} | \Psi_{\mathscr{E}} \rangle}{\langle \Psi_{\mathscr{E}} | \Psi_{\mathscr{E}} \rangle} \\ &= \mathscr{E} + \frac{\mathscr{E}_1 - \mathscr{E}}{\langle \Psi_{\mathscr{E}} | \Psi_{\mathscr{E}} \rangle} = \mathscr{E}_{\mathrm{NR}}, \end{aligned} \tag{115}$$

which proves the statement. There is hence a close connection between the partitioning technique and the variation principle, and the iteration scheme takes the simple form

$$\begin{tikzcd} \mathscr{E}, & \mathscr{E}_1, & \mathscr{E}_{\mathrm{VAR}} \arrow[ll, bend left] \end{tikzcd} \tag{116}$$

where the variational value may be calculated from $\mathscr{E}$, $\mathscr{E}_1$, and the value of the normalization integral $\langle \Psi_{\mathscr{E}} | \Psi_{\mathscr{E}} \rangle = \langle \varphi | \mathrm{W}_{\mathscr{E}}^{\dagger} \mathrm{W}_{\mathscr{E}} | \varphi \rangle$.

More generally we note that, if F is an arbitrary linear operator, one may according to (106) calculate its expectation value from

$$\langle F \rangle_{\mathrm{AV}} = \frac{\langle \Psi_{\mathscr{E}} | F | \Psi_{\mathscr{E}} \rangle}{\langle \Psi_{\mathscr{E}} | \Psi_{\mathscr{E}} \rangle} = \frac{\langle \varphi | W_{\mathscr{E}}^{\dagger} F W_{\mathscr{E}} | \varphi \rangle}{\langle \varphi | W_{\mathscr{E}}^{\dagger} W_{\mathscr{E}} | \varphi \rangle}. \tag{117}$$

Similarly, one obtains for the Nth-order density matrix

$$\Gamma_{\mathscr{E}} = | \Psi_{\mathscr{E}} \rangle \langle \Psi_{\mathscr{E}} | \Psi_{\mathscr{E}} \rangle^{-1} \langle \Psi_{\mathscr{E}} | = \frac{W_{\mathscr{E}} | \varphi \rangle \langle \varphi | W_{\mathscr{E}}}{\langle \varphi | W_{\mathscr{E}}^{\dagger} W_{\mathscr{E}} | \varphi \rangle}. \tag{118}$$

In principle, it should then be possible also to evaluate all the reduced-density matrices including $\Gamma(x_1 x_2 | x_1' x_2')$ and $\gamma(x_1 | x_1')$ and their natural representations. However, it should be observed that, even if the starting formulas are explicitly known, the evaluation of the final expressions requires some rather formidable work.

The wave operator $W = [1 - (PH/\mathscr{E})]^{-1}$ has a very simple form but, since it is not self-adjoint or normal for any values of the complex variable $\mathscr{E}$, it has a somewhat unusual structure, and it may seem desirable to reduce it to operators of a more conventional form. Using the elementary identity $(A - B)^{-1} \equiv A^{-1} + (A - B)^{-1}BA^{-1}$, one obtains directly

$$W = 1 + \left(1 - \frac{PH}{\mathscr{E}}\right)^{-1} \frac{PH}{\mathscr{E}} = 1 + (\mathscr{E} - PH)^{-1}PH = 1 + TH,$$

where $T = (\mathscr{E} - PH)^{-1}P$. Using (110), one obtains further the transformations

$$\begin{aligned} T &= (\mathscr{E} - PH)^{-1}P = P(\mathscr{E} - HP)^{-1} \\ &= P(\mathscr{E} - PH)^{-1}P = P(\mathscr{E} - HP)^{-1}P \\ &= P(\mathscr{E} - PHP)^{-1}P = P(\mathscr{E} - \bar{H})^{-1}P, \end{aligned} \tag{119}$$

where $\bar{H} = PHP$ is an auxiliary Hamiltonian. For real values of $\mathscr{E}$, the operator T is hence self-adjoint, and it is sometimes referred to as a "reduced resolvent." It becomes singular in the points $\mathscr{E} = \bar{E}_k$, i.e., for the eigenvalues of $\bar{H} = PHP$, which usually satisfy the inequalities $\bar{E}_k > E_k$ in order from below. For the wave operator W and the operator $\Omega = HW$ occurring in the bracketing function (107), one hence obtains the expressions

$$W = 1 + TH, \qquad \Omega = HW = H + HTH. \tag{120}$$

This implies that, for real $\mathscr{E}$, $\boldsymbol{\Omega}$ is also a self-adjoint operator.

B. Wave and Reaction Operators in Perturbation Theory

In the modern formulation of perturbation theory based on an arbitrary normalized reference function φ, one assumes that the Hamiltonian is the sum of two terms $H = H_0 + V$, where the first term H_0 is such that one can evaluate the associated "unperturbed" reduced resolvent

$$T_0(\mathscr{E}) \equiv (\mathscr{E} - PH_0)^{-1}P, \tag{121}$$

where $P = 1 - |\varphi\rangle\langle\varphi|$. According to (119), there are also four alternative forms of $T_0(\mathscr{E})$. Using the rule $(AB)^{-1} = B^{-1}A^{-1}$ and the identity $(A - B)^{-1} \equiv A^{-1} + A^{-1}B(A - B)^{-1}$, one gets further the following transformation of the fundamental operator T:

$$\begin{aligned} T &= [\mathscr{E} - PH]^{-1}P = [(\mathscr{E} - PH_0) - PV]^{-1}P \\ &= [1 - (\mathscr{E} - PH_0)^{-1}PV]^{-1}(\mathscr{E} - PH_0)^{-1}P \\ &= [1 - T_0V]^{-1}T_0 = T_0 + T_0V[1 - T_0V]^{-1}T_0 \\ &= T_0 + T_0tT_0, \end{aligned} \tag{122}$$

where we have used the notation $t = V[1 - T_0V]^{-1}$. The operator

$$t = V[1 - T_0V]^{-1} = [1 - VT_0]^{-1}V = [V^{-1} - T_0]^{-1}, \tag{123}$$

is usually referred to as the *reaction operator;* the last form may be used provided that V^{-1} exists. One has further $t = V + V[1 - T_0V]^{-1}T_0V = V + VTV$, which gives the reciprocity relations

$$T = T_0 + T_0tT_0, \qquad t = V + VTV. \tag{124}$$

If V^{-1} exists, the inverse of the reaction operator is given by the simple two-term formula

$$t^{-1} = V^{-1} - T_0. \tag{125}$$

Let us now consider the fundamental operators W and $\Omega = HW$. Introducing the unperturbed wave operator $W_0 = [1 - (PH_0/\mathscr{E})^{-1} = 1 + T_0H_0$, one gets directly

$$\begin{aligned} W &= \left[1 - \frac{PH}{\mathscr{E}}\right]^{-1} = \left[\left(1 - \frac{PH_0}{\mathscr{E}}\right) - \frac{PV}{\mathscr{E}}\right]^{-1} \\ &= \left(1 - \left(1 - \frac{PH_0}{\mathscr{E}}\right)^{-1} \frac{PV}{\mathscr{E}}\right]^{-1} \left(1 - \frac{PH_0}{\mathscr{E}}\right)^{-1} \\ &= [1 - T_0V]^{-1}W_0 \end{aligned} \tag{126}$$

Here the operator $W_V = [1 - T_0V]^{-1}$ is identical to the reaction operator t, except that one left-hand factor V is missing. For $\Omega = HW$, one obtains similarly

$$\begin{aligned} \Omega &= HW = (H_0 + V)[1 - T_0V]^{-1}W_0 = H_0[1 - T_0V]^{-1}W_0 + tW_0 \\ &= H_0W_0 + H_0T_0V[1 - T_0V]^{-1}W_0 + tW_0 \\ &= H_0W_0 + H_0T_0tW_0 + tW_0 = \mathbf{\Omega}_0 + (1 + H_0T_0)tW_0 = \mathbf{\Omega}_0 + W_0^{\dagger}tW_0, \end{aligned} \tag{127}$$

where we have used the fact that $W_0^{\dagger} = 1 + H_0T_0$ for real $\mathscr{E}$. Introducing the "modified reference function" $\bar{\varphi} = W_0\varphi$, which is associated with the solution of the unperturbed problem, one obtains for the solution $\Psi_{\mathscr{E}}$ to the inhomogeneous equation

$$\Psi_{\mathscr{E}} = W\varphi = (1 - T_0V)^{-1}\bar{\varphi}, \tag{128}$$

and for the bracketing function $\mathscr{E}_1 = f(\mathscr{E})$

$$\mathscr{E}_1 = \langle\varphi|\Omega|\varphi\rangle = \langle\varphi|\Omega_0|\varphi\rangle + \langle\varphi|W_0^{\dagger}tW_0|\varphi\rangle = \langle\varphi|H_0|\bar{\varphi}\rangle + \langle\bar{\varphi}|t|\bar{\varphi}\rangle. \tag{129}$$

The problem is hence essentially solved, if one can evaluate the operator $W_V = (1 - T_0V)^{-1}$ or the reaction operator $t = V(1 - T_0V)^{-1} = VW_V$.

A slight generalization of this approach is provided by the fact that one may write the Hamiltonian in the form $H = H_0' + V_0' = (H_0 + \Delta) + (V - \Delta)$, where Δ is an arbitrary linear operator. In the simplest case, one may choose Δ as a constant times the identity operator, i.e., $\Delta = (\mathscr{E} - a) \cdot 1$. For the new unperturbed resolvent T_0', one obtains

$$T_0'(\mathscr{E}) \equiv (\mathscr{E} - PH_0')^{-1}P = P(\mathscr{E} - H_0'P)^{-1} = (a - PH_0)^{-1}P, \tag{130}$$

where the last form may be derived from the identity $(a - PH_0)P = P(\mathscr{E} - H_0'P)$. Hence one has the result

$$T_0'(\mathscr{E}) = T_0(a), \tag{131}$$

and all the previous formulas are still valid provided that one also replaces V by $V' = V - (\mathscr{E} - a)$.

By putting $a = \mathscr{E}$, one gets a so-called Brillouin-type resolvent $T_0(\mathscr{E})$, whereas by putting $a = E_0$ one gets a so-called Schrödinger-type resolvent R_0.

In the conventional perturbation theories, one chooses the reference function φ to be an eigenfunction to H_0, so that $H_0\varphi_0 = E_0\varphi_0$. In such a case, one has immediately $\tilde{\varphi} = \varphi_0$, and formulas (128) and (129) take the simplified form

$$\Psi_{\mathscr{E}} = (1 - T_0V)^{-1}\varphi_0, \tag{132}$$

$$\mathscr{E}_1 = E_0 + \langle\varphi_0|t|\varphi_0\rangle, \tag{133}$$

with $t = V(1 - T_0V)^{-1}$. Since the projector $P = 1 - |\varphi_0\rangle\langle\varphi_0$ commutes with H_0, one obtains further $T_0(a) = (a - PH_0)^{-1}P = P(a - H_0P)^{-1} = (a - H_0)^{-1}P$, i.e., $T_0(a)$ may be expressed by means of the spectral resolution of the resolvent $(a - H_0)^{-1}$ involving the eigenfunctions $\Psi_k^{(0)}$ and eigenvalues $E_k^{(0)}$ to H_0:

$$T_0(a) = \sum_{k \neq 0} |\Psi_k^{(0)}\rangle\langle\Psi_k^{(0)}|/(a - E_k^{(0)}), \tag{134}$$

where one should sum over the discrete part of the spectrum of H_0 and integrate over the continuum.

In the conventional perturbation theories, one has no bracketing function $\mathscr{E}_1 = f(\mathscr{E})$, and one tries instead to find the values $\mathscr{E}_1 = \mathscr{E} = E$ directly. One further evaluates the inverse $(1 - T_0V)^{-1}$ by means of the geometric series $(A - B)^{-1} = A^{-1} + A^{-1}BA^{-1} + A^{-1}BA^{-1}BA^{-1} + \cdots$, which is certainly not a particularly strong tool for finding the inverse; in the following, we study a completely different approach to this problem.

In concluding this section, it should be observed that the conventional Brillouin- and Schrödinger-type perturbation formulas are easily derived from relations (132) and (133), which hence represent closed forms corresponding to infinite-order perturbation theories (*114*). They are thus of essential value also in studying the background of the conventional theories.

C. Inner Projections

It is interesting to observe that the computational tool we are going to study to evaluate the reaction operator t is closely related to the partitioning technique itself. For this purpose, we consider a finite quadratic matrix $\mathbf{M}$, which has been partitioned in the form

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{aa} & \mathbf{M}_{ab} \\ \mathbf{M}_{ba} & \mathbf{M}_{bb} \end{pmatrix} \tag{135}$$

Starting from the simple matrix identity

$$\begin{pmatrix} \mathbf{1}_{aa}; & -\mathbf{M}_{ab}\mathbf{M}_{bb}^{-1} \\ \mathbf{0}_{ba}; & \mathbf{M}_{bb}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{M}_{aa} & \mathbf{M}_{ab} \\ \mathbf{M}_{ba} & \mathbf{M}_{bb} \end{pmatrix} = \begin{pmatrix} \mathbf{M}_{aa} - \mathbf{M}_{ab}\mathbf{M}_{bb}^{-1}\mathbf{M}_{ba}; & \mathbf{0}_{ab} \\ \mathbf{M}_{bb}^{-1}\mathbf{M}_{ba}; & \mathbf{1}_{bb} \end{pmatrix} \tag{136}$$

taking the determinant of both members, and observing that $|\mathbf{M}_{bb}^{-1}| = |\mathbf{M}_{bb}|^{-1}$, one obtains

$$|\mathbf{M}_{aa} - \mathbf{M}_{ab}\mathbf{M}_{bb}^{-1}\mathbf{M}_{ba}| = |\mathbf{M}|/|\mathbf{M}_{bb}|. \tag{137}$$

This formula is a convenient tool for studying the background for the partitioning technique for solving eigenvalue problems in general. Here it is now used for a different purpose.

Let M be an arbitrary positive definite operator, and let $\mathbf{g} = \{g_1, g_2, \ldots, g_n\}$ be a set of linearly independent basic functions. Using (137), one obtains

$$\langle f|M|f\rangle - \langle f|M|\mathbf{g}\rangle\langle \mathbf{g}|M|\mathbf{g}\rangle^{-1}\langle \mathbf{g}|M|f\rangle = \frac{\begin{vmatrix} \langle f|M|f\rangle & \langle f|M|\mathbf{g}\rangle \\ \langle \mathbf{g}|M|f\rangle & \langle \mathbf{g}|M|\mathbf{g}\rangle \end{vmatrix}}{|\langle \mathbf{g}|M|\mathbf{g}\rangle|} \tag{138}$$

for any arbitrary function f. If f is a linear combination of the $\mathbf{g}$, the determinant in the numerator of the right-hand member is vanishing; otherwise it is positive. Since $|\langle \mathbf{g}|M|\mathbf{g}\rangle|$ is also positive, one gets the inequality

$$\langle f|M|f\rangle \geq \langle f|M|\mathbf{g}\rangle\langle \mathbf{g}|M|\mathbf{g}\rangle^{-1}\langle \mathbf{g}|M|f\rangle. \tag{139}$$

It is easily shown that, if M has exactly p negative eigenvalues and $\mathbf{\Delta} = \langle \mathbf{g}|M|\mathbf{g}\rangle$ has also p negative eigenvalues, then the quotient in the right-hand member of (138) is never negative, and the inequality is still valid. If A and B are two self-adjoint operators, one has the operator inequality $A \geq B$ if the difference $A - B$ is positive semidefinite, i.e., $\langle f|A - B|f\rangle \geq 0$, for all f in the common domain of A and B, which also means $\langle f|A|f\rangle \geq \langle f|B|f\rangle$. Relation (139) implies then the operator inequality

$$M \geq M', \tag{140}$$

where

$$\begin{aligned} M' &= M|\mathbf{g}\rangle\langle\mathbf{g}|M|\mathbf{g}\rangle^{-1}\langle\mathbf{g}|M \\ &= \sum_{k,l=1}^{n} M|g_k\rangle(\mathbf{\Delta}^{-1})_{kl}\langle g_l|M, \end{aligned} \tag{141}$$

is a sum of n^2 terms. Putting $M\mathbf{g} = \mathbf{h}$, one gets instead

$$M' = |\mathbf{h}\rangle\langle\mathbf{h}|M^{-1}|\mathbf{h}\rangle^{-1}\langle\mathbf{h}| \tag{142}$$

$$= \sum_{k,l=1}^{n} |h_k\rangle(\mathbf{\Delta}^{-1})_{kl}\langle h_l|, \tag{143}$$

This operator is referred to as the "inner projection" of M with respect to the finite subspace $\mathbf{h} = \{h_1, h_2, \ldots, h_n\}$. It is interesting to observe that, if M is positive definite or has only a finite number of negative eigenvalues, M' is a *lower bound* to M. However, even for an arbitrary linear operator M, the inner projection M' will converge toward M when the set $\mathbf{h}$ becomes infinite and complete.

The key quantity in the inner projection M' is the matrix $\mathbf{\Delta} = \langle\mathbf{g}|M|\mathbf{g}\rangle = \langle\mathbf{h}|M^{-1}|\mathbf{h}\rangle$ of order $n \times n$. Since each element of the inverse matrix $\mathbf{\Delta}^{-1}$ may be expressed as a quotient between two determinants, the formula (143) renders automatically *rational approximations* to the quantities involved, and this is sometimes highly convenient from the computational point of view.

Applying (142) to the reaction operator t defined through relation (125), one obtains directly

$$t' = |\mathbf{h}\rangle\langle\mathbf{h}|t^{-1}|\mathbf{h}\rangle^{-1}\langle\mathbf{h}| = |\mathbf{h}\rangle\langle\mathbf{h}|V^{-1} - T_0|\mathbf{h}\rangle^{-1}\langle\mathbf{h}|. \quad 144)$$

Substituting this expression into (129), one gets

$$\begin{aligned} \mathscr{E}_1' &= \langle\varphi|\Omega_0|\varphi\rangle + \langle\bar{\varphi}|t'|\bar{\varphi}\rangle \\ &= \langle\varphi|\Omega_0|\varphi\rangle + \langle\bar{\varphi}|\mathbf{h}\rangle\langle\mathbf{h}|V^{-1} - T_0|\mathbf{h}\rangle^{-1}\langle\mathbf{h}|\bar{\varphi}\rangle, \end{aligned} \tag{145}$$

and $\mathscr{E}_1'$ is then going to converge toward $\mathscr{E}_1 = f(\mathscr{E})$ as the set $\mathbf{h} = \{h_1, h_2, \ldots, h_n\}$ becomes complete. If the situation is such that the reaction operator t is positive definite or if it has only a finite negative part, the quantity $\mathscr{E}_1'$ is going to be a lower bound to the bracketing function $\mathscr{E}_1 = f(\mathscr{E})$. This approach for evaluating lower bounds to the true eigenvalues E has been successfully applied to a series of simple systems (*115*).

In all calculations based on truncated sets, it is important to estimate the remainder associated with the infinite part left out of the computations. One way to estimate the errors involved is to calculate both upper and lower bounds to the true energy eigenvalues E by using the same truncated sets in both calculations.

D. Molecular Applications

In order to apply partitioning technique or perturbation theory to molecular structure problems, one may write the electronic Hamiltonian (9) in the special form

$$H_{\mathrm{el}} = H_{\mathrm{c}} + \sum_i (H_i + u_i) - \sum_i u_i + \sum_{i<j} H_{ij}, \qquad (146)$$

where the nuclear repulsion energy

$$H_{\mathrm{c}} = e^2 \sum_{g<h} Z_g Z_h / R_{gh} \qquad (147)$$

is essentially a constant, whereas the one-electron energy H_i and the interelectronic energy H_{ij} have the form

$$H_i = (p_i^2/2m) - e^2 \sum_g Z_g/r_{ig}, \qquad H_{ij} = e^2/r_{ij}, \qquad (148)$$

and the operators u_i are potentials to be determined. Hence one may use the relations

$$H_0 = \sum_i (H_i + u_i), \qquad V = -\sum_i u_i + \sum_{i<j} H_{ij}, \qquad (149)$$

as the basis for a perturbation study of the molecular structure. It is evident that, even if all the starting formulas are explicitly known, the derivation and evaluation of the final results is going to be a formidable task. There is no question, however, that the present HF, MC-SCF, and CI calculations have to be supplemented by more deep-going studies, even if initial resistance to the new pathways of research turns out to be appreciable.

One of the first aims is to formulate the present CI method in terms of partitioning technique and inner projections (*116*). The next goal is to estimate the error involved in a CI calculation based on a truncated set, e.g., by evaluating upper and lower bounds to the true energy eigenvalues and to the expectation values $\langle F \rangle_{\mathrm{AV}}$ under consideration.

Since the Hamiltonian (146) is given as a sum of one- and two-electron terms, it is evident that, in the algebraic analysis of the partitioning technique, one is going to need "cluster expansions" of the type

$$t = \sum_i t_i + \sum_{i<j} t_{ij} + \sum_{i<j<k} t_{ijk} + \cdots, \tag{150}$$

for all the fundamental operators t, W, $\Omega = HW$, . . . involved. In treating this problem, the method of "second quantization" and the Fock space associated with all N-electrons systems for $N = 1, 2, 3, \ldots$ are certainly going to be valuable tools.

One should further observe the existence of the *separability theorem,* which says that if the Hamiltonian takes the special form

$$H = H_a + H_b + \lambda_{ab}H_{ab}, \tag{151}$$

then the energy E takes the form $E = E_a + E_b$ in the limit of $\lambda_{ab} = 0$—at least in the nondegenerate case—which corresponds to the famous "linked-cluster theorem" (*117*). However, even if the separability theorem is a fairly obvious property of the Schrödinger equation $(H - E)\Psi_E = 0$, it is by no means trivial for the inhomogeneous equation (104), which forms the basis for the partitioning technique, and a deeper study of this problem remains to be done.

The potentials u_i in the Hamiltonian (146) are still at our disposal, and the question is whether they should be chosen as Hartree–Fock, Brueckner, or some other type (*117*). It is clear that, if the molecule under consideration has a constant of motion Q that satisfies the relations $HQ = QH$, $Q^2 = Q$, $Q^\dagger = Q$, then it may be convenient to choose the potentials u_i so that $H_0Q = QH_0$, since otherwise the perturbation problem becomes exceedingly complicated (*118*). This implies that the *symmetry-restricted* SCF schemes are going to be of particular importance in choosing the unperturbed Hamiltonian H_0.

In the conventional perturbation theories for treating the Hamiltonian $H = H_0 + \lambda V$, one expresses all results in terms of power series in λ and keeps track of all the terms occurring by means of the "diagram technique," which is essentially a bookkeeping procedure. Since in the partitioning technique based on inner projections, the power series are replaced by rational approximations in the form of quotients between polynomials in λ, it seems desirable to develop a similar diagram technique for the classification of all the coefficients occurring in the numerator and the denominator. Present experience with so-called Padé approximations (*119*) seems to indicate that the rational approximations will render a much more powerful computational tool than the present power series, particularly for molecular applications.

It is evident that the concept of the inner projection

$$M' = |\mathbf{h}\rangle\langle\mathbf{h}|M^{-1}|\mathbf{h}\rangle^{-1}\langle\mathbf{h}|$$

defined by (142) plays an important role in the development of the fundamental theory. It should be observed, however, that it is an important computational tool also in other connections. For instance, if one has to evaluate the Coulomb integral

$$\left\langle f(1,2) \left| \frac{e^2}{r_{12}} \right| g(1,2) \right\rangle = e^2 \int \frac{f^*(1,2)g(1,2)}{r_{12}}\, dx_1\, dx_2, \qquad (152)$$

the operator e^2/r_{12} is certainly positive definite, and this implies that the quantity

$$e^2 \langle f(1,2)|\mathbf{h}(1,2)\rangle \langle \mathbf{h}(1,2)|r_{12}|\mathbf{h}(1,2)\rangle^{-1} \langle \mathbf{h}(1,2)|g(1,2)\rangle \qquad (153)$$

gives a strict lower bound to the integral (152) in the particular case when $f(1,2) \equiv g(1,2)$, and that in general even when $f \neq g$ it approaches the correct value as the set **h** tends to become complete. If one chooses the set $\mathbf{h}(1,2)$ to consist of a product of Gaussians, all the elements of the matrix $\langle \mathbf{h}|r_{12}|\mathbf{h}\rangle$ may be evaluated in analytic form. The formula (153) thus provides a new method for evaluating molecular integrals by means of Gaussians, but the algorithms for the computations have to be carefully studied in order to preserve the maximum number of significant figures in the final result—particularly if the set of Gaussians has any approximate linear dependencies. This approach renders it possible to use ordinary atomic orbitals as a basis for a molecular calculation and still evaluate the molecular integrals by means of Gaussians.

E. Green's Functions Technique

Another possibility of solving the inhomogeneous equation (104) is based on the existence of the "resolvent"

$$R = (\mathscr{E} - H)^{-1}, \qquad (154)$$

which is regular everywhere in the complex plane except for the eigenvalues $\mathscr{E} = E$. One gets directly $\Psi_{\mathscr{E}} = -aR\varphi$ and $\langle \varphi|\Psi_{\mathscr{E}}\rangle = -a\langle\varphi|R|\varphi\rangle = 1$, and $a = -\langle\varphi|R|\varphi\rangle^{-1}$, i.e.,

$$\Psi_{\mathscr{E}} = R\varphi/\langle\varphi|R|\varphi\rangle. \qquad (155)$$

If $\mathscr{E}$ approaches an eigenvalue E to the Hamiltonian H, this quotient has a well-defined limit, which has previously been studied by means of the partitioning technique. We note that the eigenvalues E correspond to the zero points of the function $a(\mathscr{E}) = -\langle\varphi|R|\varphi\rangle^{-1}$, i.e., to the poles of the so-called Weinstein function

$$W(\mathscr{E}) = \langle\varphi|R|\varphi\rangle. \qquad (156)$$

This function gives also a tool for determining upper and lower bounds to E, and we note the relation $a = \mathscr{E}_1 - \mathscr{E}$, which gives the connection $W(\mathscr{E}) = [\mathscr{E} - f(\mathscr{E})]^{-1}$ with the bracketing function.

The resolvent $R = (\mathscr{E} - H)^{-1}$ is often referred to as Green's operator and its kernel is called Green's function. By considering a Fock space associated with all N-electron systems for $N = 1, 2, 3, \ldots$, one may then also consider ionization processes associated with a shift of the number of electrons from N to $N \pm 1$. The properties of R are then studied in great detail by means of its spectral resolution.

The use of Green's functions in molecular calculations was pioneered by Linderberg and Ohrn (*120*) and, for a survey of the current development, we refer to some recent reviews (*121*).

F. Density Matrices and Quantum Statistics

Molecular structure calculations deal not only with the ground state of systems but also with conformational analysis, studies of scattering phenomena, intermolecular energy surfaces, and the kinetics of chemical reactions. Many of these problems involve processes that depend both on the time t and the temperature T, and they are hence not described by ordinary quantum mechanics, which deals with time-independent phenomena at $T = 0°\mathrm{K}$. In order to treat such systems, one needs *quantum statistics* based on time-dependent density operators $\Gamma = \Gamma(t)$, which are generalizations of those studied in a previous section.

Let us first consider the carrier space $A = \{\psi\}$ consisting of all the wave functions ψ represented by points in this space. If during the time interval from t_0 to t, an arbitrary element changes from $\psi(t_0)$ to $\psi(t)$, the mapping $\psi(t_0) \rightarrow \psi(t)$ is called an *evolution mapping* and the corresponding operator $S = S(t, t_0)$ *an evolution operator.* It is usually assumed that such a mapping is one to one, i.e., that S^{-1} exists so that

$$\psi(t) = S\psi(t_0), \qquad \psi(t_0) = S^{-1}\psi(t). \tag{157}$$

We further assume that the density operator Γ, which describes an assembly of physical systems, renders a mapping between elements of the carrier space that is essentially invariant under time, i.e., if $\Gamma(t_0)$ gives a mapping between $\psi_1(t_0)$ and $\psi_2(t_0)$ at the time t_0, then $\Gamma(t)$ renders a mapping between the elements $\psi_1(t)$ and $\psi_2(t)$ at any time t. This gives

$$\begin{tikzcd}
\psi_1(t) \arrow[r, "\Gamma(t)"] & \psi_2(t) \\
\psi_1(t_0) \arrow[u, "S"] \arrow[r, "\Gamma(t_0)"'] & \psi_2(t_0) \arrow[u, "S"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\psi_1(t) \arrow[r, leftrightarrow, "\Gamma(t)"] & \psi_2(t) \\
{} \arrow[u, leftrightarrow, "S^{-1}"] \arrow[r, "\Gamma(t_0)"'] & {} \arrow[u, "S"']
\end{tikzcd}$$

the relation

$$\Gamma(t) = S\Gamma(t_0)S^{-1}, \tag{158}$$

i.e., the time-evolution of Γ, is then immediately given by a *similarity transformation.* If $\Gamma(t_0)$ is normalized according to (67), one gets directly $\mathrm{Tr}\,\Gamma(t) = \mathrm{Tr}\,\Gamma(t_0) = 1$. The expectation value of a linear operator F is then given by formula (68).

In ordinary quantum mechanics, one has the time-dependent Schrödinger equation

$$-\frac{h}{2\pi i}\frac{d\psi}{dt} = H\psi, \tag{159}$$

where H is the Hamiltonian operator. Since

$$\frac{d\psi}{dt} = \frac{dS}{dt}\psi(t_0) = \frac{dS}{dt}S^{-1}\psi(t),$$

one has the connection

$$H = -\frac{h}{2\pi i}\frac{dS}{dt}S^{-1} = +\frac{h}{2\pi i}S\frac{dS^{-1}}{dt}, \tag{160}$$

and it is hence always possible to derive a Hamiltonian from the evolution operator S. It is easily shown that if and only if $S(t, t_0) = S(t - t_0, 0)$ then the Hamiltonian H is independent of time; in such a case, one says that $S = S(t, t_0)$ is homogeneous in time, and one has the simple relation

$$S = \exp[-(2\pi i/h)H(t - t_0)]. \tag{161}$$

In general, the connection between S and H is much more complicated. Taking the time derivative of (159) and using (160), one obtains further

$$-\frac{h}{2\pi i}\frac{d\Gamma}{dt} = H\Gamma - \Gamma H, \tag{162}$$

which is Liouville's equation. Taking the time derivative of the expectation value $\langle F\rangle_{\mathrm{AV}} = \mathrm{Tr}\{F\Gamma\}$, one gets also

$$\begin{aligned}\frac{d\langle F\rangle_{\mathrm{AV}}}{dt} &= \mathrm{Tr}\left\{F\frac{d\Gamma}{dt} + \frac{dF}{dt}\Gamma\right\}\\ &= \mathrm{Tr}\left\{-\frac{2\pi i}{h}FH\Gamma + \frac{2\pi i}{h}F\Gamma H + \frac{dF}{dt}\Gamma\right\}\\ &= \mathrm{Tr}\left\{+\frac{2\pi i}{h}(HF - FH)\Gamma + \frac{dF}{dt}\Gamma\right\}\\ &= \left\langle\frac{2\pi i}{h}(HF - FH) + \frac{dF}{dt}\right\rangle_{\mathrm{AV}}\end{aligned} \tag{163}$$

which corresponds to Ehrenfest's relations in quantum statistics.

It has previously been shown in connection with (82) and (85) that, if Γ consists of a single projector satisfying the relation $\Gamma^2 = \Gamma$, one has a homogeneous assembly corresponding to pure quantum mechanics. The quantity

$$\gamma = \mathrm{Tr}(\Gamma - \Gamma^2), \qquad 0 \leq \gamma \leq 1, \tag{164}$$

may now be used to measure the deviation from this situation. If $\gamma = 0$, one has an assembly corresponding to pure quantum mechanics, whereas if $\gamma = 1$, one has an assembly that closely corresponds to "classical mechanics." It should be observed that γ is independent of time and constant, which is characteristic for every experimental setup.

A system in thermal equilibrium with the temperature T is associated with a canonical ensemble described by the density operator

$$\Gamma = \exp[-(H - F)/kT], \tag{165}$$

where F is the so-called free energy defined by the normalization condition $\mathrm{Tr}\,\Gamma = 1$, which gives

$$e^{-F/kT} = \mathrm{Tr}\, e^{-H/kT}, \tag{166}$$

where the right-hand member is the famous "partition function." The *entropy* $\mathcal{S}$ is finally given through the definition

$$\mathcal{S} = -k\, \langle \log \Gamma \rangle_{\mathrm{AV}} = -k\, \mathrm{Tr}\, \Gamma \log \Gamma, \tag{167}$$

and, taking the expectation value of $\log \Gamma$ according to (165), one obtains the standard connection $F = \langle H \rangle_{\mathrm{AV}} - \mathcal{S}T$. The equation of state takes finally the form

$$p = -\partial F/\partial V. \tag{168}$$

These formulas indicate that molecular processes that depend on the time t or on the temperature T or both are conveniently handled by means of density operators Γ in the framework of quantum statistics. Even if such problems are on the research programs of many groups all over the world, computational progress has so far been comparatively small, and it seems as if we still have a long way to go from quantum mechanics to quantum statistics as far as molecular applications are concerned.

It seems as if, even in this area, Green's functions and the related propagators (*120, 121*) are going to be valuable conceptual and computational tools.

NOTE ADDED IN PROOF

Since this paper was written, a great deal has happened in computational quantum chemistry, particularly as to the so-called "unitary group approach." For references, see e.g., the Proceedings of the Nobel Symposium on Many-Body Theory in Lerum, Sweden, in June, 1979, published in *Physica Scripta* **21,** 229 (1980). For other recent developments, see also the Proceedings of the Third International Congress on Quantum Chemistry in Kyoto, 1979; the introductory papers are published in "Horizons of Quantum Chemistry" (K. Fukui and B. Pullman, eds.), Reidel, Dordrecht (1980), and the contributed papers in *Int. J. Quantum Chem.* **18,** Nos. 1–2 (1980).

ACKNOWLEDGMENTS

The author is greatly indebted to all members of the Uppsala Quantum Chemistry Group and the Florida Quantum Theory Project for valuable discussions about the subject of this chapter. Particular thanks go to Drs. Ulf Wahlgren and to Michael Hehenberger for presenting valuable aspects of the modern standard computational programs, and to the latter for help with the bibliography.

This work has been partly supported by a grant from the U.S. National Science Foundation, which is gratefully acknowledged.

REFERENCES

1. W. Heisenberg, *Z. Phys.* **38,** 411 (1926); **39,** 499 (1929); **41,** 239 (1927).
2. W. Heitler and F. London, *Z. Phys.* **44,** 455 (1927).
3. E. A. Hylleraas, *Z. Phys.* **48,** 469 (1928); **54,** 347 (1929).
4. H. M. James and A. S. Coolidge, *J. Chem. Phys.* **1,** 825 (1933).
5. C. L. Pekeris, *Phys. Rev.* **112,** 1649 (1958); **155,** 1216 (1959).
6. For a survey, see W. Kolos, *Adv. Quantum Chem.* **5,** 99 (1970).
7. For a survey of the spin-free theory, see, e.g., P. O. Löwdin and O. Goscinski, *Int. J. Quantum Chem.* **S3,** 533 (1970).
8. P. A. M. Dirac, *Proc. R. Soc. London, Ser. A* **123,** 714 (1929).
9. M. Born and J. R. Oppenheimer, *Ann. Phys. Leipzig* **84,** 457 (1927).
10. C. Eckart, *Phys. Rev.* **36,** 878 (1930).
11. H. Shull and P.-O. Löwdin, *Phys. Rev.* **110,** 1466 (1958); P. O. Löwdin, *Adv. Chem. Phys.* **2,** 207 (1959), particularly pp. 266–269; see also F. Weinhold, *Adv. Quantum Chem.* **6,** 299 (1972).
12. P.-O. Löwdin, *Int. J. Quantum Chem.* **S1,** 811 (1967).
13. P.-O. Löwdin, *Adv. Phys.* **5,** 1 (1956), particularly pp. 46–49.
14. For a survey, see P.-O. Löwdin, *Adv. Quantum Chem.* **5,** 185 (1970).
15. P.-O. Löwdin, *Phys. Rev.* **97,** 1509 (1955); *Rev. Mod. Phys.* **34,** 520 (1962).
16. For a survey, see P.-O. Löwdin, *Rev. Mod. Phys.* **36,** 966 (1964).
17. H. Weyl, "Gruppentheorie und Quantenmechanik." Hirzel, Leipzig (1931) [English translation: "Theory of Groups and Quantum Mechanics." Princeton Univ. Press, Princeton (1931)]. E. P. Wigner, "Gruppentheorie und ihre Anwendungen auf die Quantenmechanik der Atomspektren." F. Vieweg (1931) [English translation: Academic Press (1959)]. See also P.-O. Löwdin, *Rev. Mod. Phys.* **39,** 259 (1967).
18. N. Bohr, *Proc. Phys. Soc., London* **35,** 296 (1923). D. R. Hartree, *Proc. Cambridge Philos. Soc.* **24,** 89 (1928).
19. J. C. Slater, *Phys. Rev.* **34,** 1293 (1929); *Phys. Rev.* **35,** 210 (1930).
20. V. Fock, *Z. Phys.* **61,** 126 (1930).
21. P. A. M. Dirac, *Proc. Cambridge Philos. Soc.* **26,** 376 (1930); **27,** 240 (1931).
22. L. Brillouin, Actual. *Sci. Ind.* No. 71 (1933); No. 159 (1934); C. Møller and M. S. Plesset, *Phys. Rev.* **46,** 618 (1934).

23. T. Koopmans, *Physica* **1**, 104 (1933).
24. C. A. Coulson and H. C. Longuet-Higgins, *Proc. Soc. London, Ser. A* **191**, 39; **192**, 16 (1947); **193**, 447, 456; **195**, 188 (1948).
25. For a survey, see B. Pullman and A. Pullman, "Quantum Biochemistry." Wiley (Interscience), New York (1963).
26. C. C. J. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951); G. G. Hall, *Proc. R. Soc. London, Ser. A* **202**, 336 (1950); **205**, 541 (1951); **213**, 113 (1952).
27. P.-O. Löwdin, *Rev. Mod. Phys.* **35**, 496 (1963).
28. See, e.g., P.-O. Löwdin, *Adv. Chem. Phys.* **2**, 207 (1959); **14**, 283 (1969); R. K. Nesbet, *ibid.* **14**, 1 (1969); R. McWeeny, "The New World of Quantum Chemistry" (B. Pullman and R. G. Parr, eds.), p. 3. Reidel, Dordrecht (1976).
29. P.-O. Löwdin, *Phys. Rev.* **97**, 1474, 1490, 1509 (1955); "Quantum Theory of Atoms, Molecules, and the Solid State," p. 601. Academic Press, New York.
30. For a survey, see R. Pauncz, "Alternant Molecular Orbital Method." Saunders, Philadelphia (1967).
31. J. Frenkel, "Wave Mechanics, Advanced General Theory." Clarendon Press, Oxford (1934), particularly pp. 460–462; J. C. Slater, *Phys. Rev.* **91**, 528 (1953); P.-O. Löwdin, *Phys. Rev.* **97**, 1474 (1955).
32. K. Husimi, *Proc. Phys.-Math. Soc. Jpn.* **22**, 264 (1940); P.-O. Löwdin, *Phys. Rev.* **97**, 1474 (1955); R. McWeeny, *Proc. R. Soc. London, Ser. A* **223**, 63 (1954); **232**, 114 (1955).
33. P.-O. Löwdin, *Adv. Phys.* **5**, 1 (1956); particularly p. 12.
34. P.-O. Löwdin, *Phys. Rev.* **97**, 1474 (1955); *Adv. Chem. Phys.* **2**, 207 (1959), particularly p. 279. B. C. Carlson and J. M. Keller, *Phys. Rev.* **121**, 659 (1961); for a survey, see A. J. Coleman, *Rev. Mod. Phys.* **35**, 668 (1963).
35. P.-O. Löwdin and H. Shull, *Phys. Rev.* **101**, 1730 (1956).
36. E. R. Davidson, "Reduced Density Matrices in Quantum Chemistry." Academic Press, New York (1976).
37. For a survey, see "Reduced Density Matrices with Applications to Physical and Chemical Systems" (Queen's Univ. Papers in Pure and Applied Math. No. 11, A. J. Coleman and R. M. Erdahl, eds.) (1968); A. J. Coleman. *Energy, Struct., React., Proc. Boulder Summer Res. Conf. Theor. Chem., 1972.* Publ. (1973).
38. Y. Sugiura, *Z. Phys.* **45**, 484 (1927).
39. For a survey, see Kotani, Amemiya, Ishiguri, and Kimura, "Tables of Molecular Integrals," 2nd ed. Maruzen, Tokyo (1963).
40. S. O. Lundqvist and P.-O. Löwdin, *Ark. Fys.* **3**, 147 (1951).
41. Report from the Shelter Island Conference of 1951, "Quantum Mechanical Methods in Valence Theory." National Academy of Sciences and Office of Naval Research, Washington D.C. (1952).
42. *Tex. J. Sci.* **8**, 163 (1956).
43. *Rev. Mod. Phys.* **32**, 169 (1960).
44. H. Shull and P.-O. Löwdin, *J. Chem. Phys.* **23**, 1565 (1955).
45. E. Schrödinger, *Ann. Phys. Leipzig* **79**, 361 (1926); P. S. Epstein, *Phys. Rev.* **28**, 695 (1926); H. Shull and P.-O. Löwdin, *J. Chem. Phys.* **23**, 1565 (1955); *Phys. Rev.* **101**, 1730 (1956).
46. P.-O. Löwdin, *Ann. Rev. Phys. Chem.* **11**, 107 (1960); *J. Appl. Phys. Suppl.* **33**, 251 (1962), particularly pp. 264–265.
47. T. Givens, Oak Ridge National Laboratory Report, ORNL–1574 (1954).
48. P.-O. Löwdin, *J. Chem. Phys.* **19**, 1396 (1951); *J. Mol. Spectrosc.* **10**, 12 (1963); **13**, 326, 331 (1964); R. K. Nesbet, *J. Chem. Phys.* **43**, 311 (1965).
49. In principle, there are no proceedings from the Gordon Conferences.
50. *Rev. Mod. Phys.* **35**, 415 (1963); *J. Chem. Phys. Suppl.* **43**(10), (Part 2), 1965; *Int. J. Quantum Chem., Symp. Suppl.* **1–11** (1967–1977); **12–14** (1978–1980).

51. "The World of Quantum Chemistry" (R. Daudel and B. Pullman, eds.). Reidel, Dordrecht (1973). "The New World of Quantum Chemistry" (B. Pullman and R. G. Parr, eds.). Reidel, Dordrecht (1976).
52. *Int. J. Quantum Chem., Symp. Suppl.* **QB 1–4**(1974–1977); **5–7**(1978–1980).
53. J. Pople *et al.*, Quantum Chemistry Program Exchange 236.
54. E. Clementi and H. E. Popkie, Report from the IBM Research Laboratory, San Jose, California.
55. G. H. F. Diercksen and W. P. Kraemer, MUNICH Manual, Special Technical Report, Max-Planck Institut für Physik und Astrophysik, Munich (1970).
56. Neumann, Basch, Kornegay, Snyder, Moscowitz, Hornback, and Liebmann, "The POLYATOM (version 2) Systems of Programs," Quantum Chemistry Program Exchange 199.
57. Dedieu, Demuynck, Strich, and Veillard, ASTERIX-reports, CNRS, Strasbourg, France.
58. J. Almlöf, USIP Technical Reports 72-09 (1972) and 74-29 (1974), Institute of Theoretical Physics, Univ. of Stockholm, Sweden. For a symmetry extension, see also K. Pitzer, *J. Chem. Phys.* **59**, 3308 (1973).
59. E. Clementi, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 2942 (1972).
60. E. Clementi and H. E. Popkie, *J. Chem. Phys.* **57**, 1077 (1972).
61. A. Veillard, *J. Am. Chem. Soc.* **95**, 5663 (1973).
62. J. Almlöf, *Int. J. Quantum Chem.* **8**, 915 (1974).
63. H. Johansen, *Int. J. Quantum Chem.* **9**, 459 (1975).
64. H. E. Popkie and J. Kaufman, *Int. J. Quantum Chem.* **10**, 569 (1976).
65. Liu, Clementi, and Yoshimine, *J. Chem. Phys.* **64**, 2314 (1976).
66. C. C. J. Roothaan, *Rev. Mod. Phys.* **32**, 170 (1960).
67. Bagus, Liu, McLean, and M. Yoshimine; see, e.g., P. S. Bagus, IBM Research Laboratory Technical Report RJ 1077, San Jose, California (1972).
68. See P. S. Bagus and U. Wahlgren, Reports from the IBM Research Laboratory, San Jose, California.
69. N. M. Winter and T. H. Dunning, *Chem. Phys. Lett.* **8**, 169 (1971); L. L. Shipman and R. E. Christoffersen, *Chem. Phys. Lett.* **15**, 469 (1972); W. B. Neilsen, *Chem. Phys. Lett.* **18**, 225 (1973); I. Røeggen, *Chem. Phys. Lett.* **22**, 140 (1973); I. H. Hillier and V. R. Saunders, *Int. J. Quantum Chem.* **7**, 699 (1973).
70. G. Das and A. C. Wahl, *J. Chem. Phys.* **56**, 1769 (1972).
71. J. Hinze, *J. Chem. Phys.* **59**, 6424 (1973).
72. G. D. Gillespie *et al.*, *J. Chem. Phys.* **63**, 3425 (1975).
73. W. A. Goddard III and R. C. Ladner, *J. Am. Chem. Soc.* **93**, 6750 (1971); W. A. Goddard III *et al.*, *Acc. Chem. Res.* **6**, 368 (1973).
74. Hay, Dunning, and Goddard, *J. Chem. Phys.* **62**, 3912 (1975).
75. M. Yoshimine, IBM Research Laboratory, Technical Report RJ 555, San Jose, California (1973); see also I. Shavitt, *in* "Methods of Electronic Structure Theory" (H. F. Schaefer, III, ed.). Plenum Press, New York and London (1977).
76. R. K. Nesbet, *J. Chem. Phys.* **43**, 311 (1965); see also reference 48, as well as I. Shavitt, *J. Comput. Phys.* **11**, 90 (1972).
77. B. Roos, *Chem. Phys. Lett.* **15**, 153 (1972); see also paper by B. Roos, *In* "Computational Techniques in Quantum Chemistry" (Diercksen, Sutcliffe, and Veillard, eds.). Reidel, Dordrecht (1978).
78. I. Shavitt, *J. Chem. Phys.* **60**, 2865 (1972).
79. B. Liu and D. McLean, *J. Chem. Phys.* **59**, 4557 (1973).
80. G. H. F. Diercksen, W. P. Kraemer, and B. Roos, *Theor. Chim. Acta* **36**, 249 (1975).

81. Proceedings of the 1970 Conference on Potential Energy Surfaces in Chemistry in Santa Cruz (William A. Lester, Jr., ed.). IBM, San Jose Publications, San Jose, California (1971).
82. G. H. F. Diercksen *et al.*, *Astrophys. J.* **184,** 119 (1973); **205,** L97 (1976); cf. *Nature, London* **246,** 4 (1973).
83. B. Roos and P. Siegbahn, *Theor. Chim. Acta* **17,** 209 (1970); A. J. H. Wachters, *J. Chem. Phys.* **52,** 1033 (1970); F. B. van Duijneveldt, IBM Research Laboratory, Technical Report RJ 945, San Jose, California, (1971); B. Roos, A. Veillard, and G. Vinot, *Theor. Chim. Acta* **20,** 1 (1971).
84. T. H. Dunning, *J. Chem. Phys.* **53,** 2823 (1970); **55,** 716 (1971); **58,** 3958 (1972).
85. R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.* **54,** 7241 (1971).
86. D. McLean, in reference 81, p. 87, particularly p. 103.
87. G. Hall, "The New World of Quantum Chemistry" (B. Pullman and R. G. Parr, eds.), p. 137. Reidel, Dordrecht (1976), Introductory lecture at Oji International Seminar of Theories and *ab Initio* Computations of Molecular Electronic Structure, Hokkaido Univ., Tomakomai, 1976.
88. H. Kashiwagi and F. Sasaki, *Int. J. Quantum Chem.* **S7,** 515 (1973).
89. H. F. Schaefer, *Ann. Rev. Phys. Chem.* (1976); see also the book published by Plenum Press given in reference 75.
90. J. C. Slater, *Phys. Rev.* **81,** 385 (1951); **82,** 538 (1951).
91. P.-O. Löwdin, *Phys. Rev.* **97,** 1474 (1955); particularly p. 1487.
92. R. Gaspar, *Acta Phys.* **3,** 263 (1954).
93. J. C. Slater, *Adv. Quantum Chem.* **6,** 1 (1972).
94. K. Johnson, *Adv. Quantum Chem.* **7,** 143 (1973).
95. K. Siegbahn *et al.*, "ESCA, Atomic, Molecular, and Solid-State Structure Studies by Means of Electron Spectroscopy." Almqvist and Wiksell, Stockholm (1967).
96. J. C. Slater, *Adv. Quantum Chem.* **6,** 15 (1972).
97. For a survey, see O. Goscinski, *Int. J. Quantum Chem.* **S9,** 221 (1975).
98. S. Larsson, *Chem. Phys. Lett.* **32,** 401 (1975); **40,** 362 (1976); J. Electr. Spectrosc. Relat. Phenom. **8,** 171 (1976); M. Braga and S. Larsson, *Int. J. Quantum Chem.* **S11,** 61 (1977).
99. F. Herman *et al.*, *Int. J. Quantum Chem.* **S10,** 167 (1976).
100. K. Johnson and R. P. Messmer, *Int. J. Quantum Chem.* **S10,** 147 (1976).
101. J. C. Slater, *Phys. Rev.* **34,** 1293 (1929).
102. See, e.g., E. U. Condon and G. H. Shortley, "The Theory of Atomic Spectra." Cambridge Univ. Press, New York–London (1935).
103. G. Herzberg, "I. Spectra of Diatomic Molecules. II. Electronic Spectra and Electronic Structure of Polyatomic Molecules," 2nd ed. Van Nostrand-Reinhold, New York (1966).
104. J. C. Slater, *Phys. Rev.* **37,** 481 (1931); L. Pauling, *J. Chem. Phys.* **1,** 280, (1933).
105. L. Pauling, "The Nature of the Chemical Bond," 3rd ed. Cornell University Press, Ithaca (1960).
106. E. Hückel, *Z. Phys.* **60,** 423 (1930); **70,** 204 (1931); **72,** 310 (1931).
107. See reference 25.
108. R. Hoffman, *J. Chem. Phys.* **39,** 1397 (1963).
109. R. B. Woodward and R. Hoffman, *J. Am. Chem. Soc.* **87,** 395, 2046, 2511 (1965); "The Conservation of Orbital Symmetry" (Verlag Chemie GmbH, Academic Press (1970).
110. R. Pariser and R. G. Parr, *J. Chem. Phys.* **21,** 466, 767 (1953); **23,** 711 (1955); J. A. Pople, *Trans. Faraday Soc.* **49,** 1375 (1953); *Proc. Phys. Soc., London, Sect. A* **68,** 81; *J. Phys. Chem.* **61,** 6 (1957).

111. For a survey, see I. Fischer-Hjalmars, *Adv. Quantum Chem.* **2,** 25 (1965).
112. See, e.g., A. L. Robinson, *Science* **193,** 470 (1976).
113. For a survey, see, e.g., P.-O. Löwdin, *Phys. Rev. A* **139,** 357 (1965); *Int. J. Quantum Chem.* **2,** 867 (1968).
114. P.-O. Löwdin, *J. Math. Phys.* **3,** 969 (1962).
115. See a series of papers on lower bounds in the Mulliken dedicatory volume, *J. Chem. Phys. Suppl.* **43**(10)(Part 2); F. Weinhold, *Adv. Quantum Chem.* **6,** 299 (1972); C. Reid, "Quantum Science" (Calais, Goscinski, Linderberg, and Ohrn, eds.), p. 315. Plenum, New York–London (1976).
116. R. J. Bartlett and E. J. Brändas, *J. Chem. Phys.* **56,** 5467 (1972).
117. P.-O. Löwdin, *J. Math. Phys.* **3,** 1171 (1962). For a general survey, see, e.g., "The Quantum Theory of Many-Particle Systems" (H. L. Morrison, ed.). Gordon and Breach, New York–London (1962).
118. P.-O. Löwdin, *Int. J. Quantum Chem.* **2,** 867 (1968); particularly p. 918.
119. For surveys, see, e.g., G. A. Baker, Jr., *Adv. Theor. Phys.* **7,** 1 (1965); "The Pade Approximants in Theoretical Physics" G. A. Baker, Jr., and Y. L. Gammel, eds.). Academic Press, New York (1970). O. Goscinski and E. Brandas, *Int. J. Quantum Chem.* **5,** 131 (1971); "Pade Approximants and Their Applications" (P. R. Graves-Morris, ed.). Academic Press, New York (1973).
120. J. Linderberg and Y. Ohrn, "Propagators in Quantum Chemistry." Academic Press, New York (1973).
121. P. Jørgensen, *Ann. Rev. Phys. Chem.* **26,** 359 (1975); Y. Ohrn, "The New World of Quantum Chemistry" (B. Pullman and R. G. Parr, eds.), p. 57. Reidel, Dordrecht (1976).

Index